UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE DE
RIBEIRÃO PRETO
DEPARTAMENTO DE ADMINISTRAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO DE ORGANIZAÇÕES

LETÍCIA APARECIDA ORIGUELA

Estudo da influência de eventos sobre a estrutura do mercado brasileiro de ações a partir de
redes ponderadas por correlações de Pearson, Spearman e Kendall

ORIENTADOR: PROF. DR. GILBERTO
APARECIDO PRATAVIERA

RIBEIRÃO PRETO

2018

Prof. Dr. Vahan Agopyan
Reitor da Universidade de São Paulo

Prof. Dr. Dante Pinheiro Martinelli
Diretor da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto

Prof. Dr. Marcio Mattos Borges de Oliveira
Chefe do Departamento de Administração

LETÍCIA APARECIDA ORIGUELA

Estudo da influência de eventos sobre a estrutura do mercado brasileiro de ações a partir de redes ponderadas por correlações de Pearson, Spearman e Kendall

Dissertação apresentada ao Programa de Pós-Graduação em Administração de Organizações da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Mestre em Ciências. Versão Corrigida. A original encontra-se disponível na FEA-RP/USP.

ORIENTADOR: PROF. DR. GILBERTO
APARECIDO PRATAVIERA

RIBEIRÃO PRETO

2018

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Origuela, Letícia Aparecida

Estudo da influência de eventos sobre a estrutura do mercado brasileiro de ações a partir de redes ponderadas por correlações de Pearson, Spearman e Kendall / Letícia Aparecida Origuela – Ribeirão Preto, 2018.

85f.: il.; 30 cm

Dissertação apresentada ao Programa de Pós-Graduação em Administração de Organizações da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Mestre em Ciências. Versão Corrigida. A original encontra-se disponível na FEA-RP/USP. – Universidade de São Paulo

Orientador : Prata, Gilberto Aparecido

1. Redes. 2. Árvore geradora mínima. 3. Correlação de Pearson. 4. Correlação de Spearman. 5. Correlação de Kendall. 6. Mercados de ações. 7. Evento.

LETÍCIA APARECIDA ORIGUELA

Estudo da influência de eventos sobre a estrutura do mercado brasileiro de ações a partir de redes ponderadas por correlações de Pearson, Spearman e Kendall

Dissertação apresentada ao Programa de Pós-Graduação em Administração de Organizações da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Mestre em Ciências. Versão Corrigida. A original encontra-se disponível na FEA-RP/USP.

Área de Concentração: Administração de Organizações

Data de Aprovação:

____/____/____

Banca Examinadora:

Prof. Dr. Gilberto Aparecido Prata
Orientador

Professor
Avaliador 1

Professor
Avaliador 2

Professor
Avaliador 3

Aos meus pais, José e Maria, pelo apoio durante esse período de sacrifício e conquistas para a elaboração deste trabalho.

AGRADECIMENTOS

Primeiramente a Deus por todas as bênçãos a mim concedidas, pela oportunidade dos estudos, pela carreira acadêmica e profissional.

Ao meu orientador, Prof. Dr. Gilberto Ap. Prata, que me conduziu durante a realização da pesquisa, pelo conhecimento compartilhado, por toda a dedicação e paciência.

Ao Prof. Dr. Evandro Saidel Ribeiro, pelo conhecimento compartilhado sobre redes.

Aos meus pais, José e Maria, pela base sólida que me ajudou durante essa jornada.

Ao meu irmão, Leandro, pelo apoio e compreensão.

Ao meu noivo, Fábio, por acreditar nos meus sonhos, por ser o meu maior incentivador e por compreender o tempo que fiquei ausente para que esse trabalho fosse concluído.

Aos meus colegas de trabalho, em especial ao Fernando, por compreender a minha ausência para o cumprimento dos créditos e reuniões com o orientador. Isso foi essencial para a realização do mestrado!

Aos colegas de docência, pelo incentivo e troca de experiências, em especial ao Prof. Me. Luis Antonio Nogueira, por acreditar na minha capacidade e me permitir ingressar nessa carreira.

*O êxito da vida não se mede pelo caminho que
você conquistou, mas sim pelas dificuldades que
superou no caminho. Abraham Lincoln*

RESUMO

ORIGUELA, Letícia Aparecida. *Estudo da influência de eventos sobre a estrutura do mercado brasileiro de ações a partir de redes ponderadas por correlações de Pearson, Spearman e Kendall*. 2018. 85f. Dissertação (Mestrado em Administração de Organizações) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2018.

Neste trabalho foi analisada a influência de um evento sobre o mercado de ações brasileiro a partir das redes, e suas árvores geradoras mínimas, obtidas de medidas de dependência baseadas nas correlações de Pearson, de Spearman e de Kendall. O evento considerado foi a notícia da noite de 17 de maio de 2017 em que o dono da empresa brasileira JBS, Joesley Batista, gravou o então Presidente da República Michel Temer autorizando a compra do silêncio de um Deputado Federal. O dia seguinte a notícia, 18 de maio de 2017, foi definido como o dia do evento. Foram coletados dados de alta frequência de 58 ações do Ibovespa no período de 11 a 25 de maio de 2017. As alterações nas redes das ações do mercado foram analisadas comparando-se o período anterior e posterior ao evento em duas escalas de tempo: (1) Redes diárias: cinco pregões antes do evento, o dia do evento e, cinco pregões depois do evento, com cotações a cada 15 minutos; (2) Agrupadas em antes e depois: agrupando os dados dos 5 dias antes e dos 5 dias depois do evento. O estudo das redes diárias indicou mudança de tendência nas suas propriedades no decorrer do período que contém o evento, com cotações a cada 15 minutos. Isto sugeriu que análise do efeito médio contido nos dados agrupados antes e depois do evento poderiam tornar mais evidente as mudanças na estrutura de rede das ações. As redes antes e depois do evento apresentaram mudanças significativas nas suas métricas que ficaram mais evidenciadas nas árvores geradoras mínimas. As redes geradas pelas correlações de Kendall e Spearman apresentaram um número maior de agrupamentos antes e depois do evento e, após o evento, as árvores geradoras mínimas apresentaram uma redução do número de agrupamentos de ações para todos os tipos de correlação. As distribuições de grau ponderado após o evento indicam uma probabilidade maior de vértices com graus distante da média. As métricas das árvores geradoras mínimas por correlação de Spearman sofreram a maior variação, seguidas pelas de Kendall e Pearson, e também, indicaram que as redes após o evento ficaram mais robustas, ou seja, mais rígidas. A maior robustez das redes após o evento indica maior conectividade do mercado, tornando-o, como um todo, mais suscetível ao impacto de novos acontecimentos.

Palavras-chave: Redes. Árvore geradora mínima. correlação de Pearson. Correlação de Spearman. Correlação de Kendall. Mercado de ações. Evento.

ABSTRACT

ORIGUELA, Leticia Aparecida. *Weighted networks from Pearson, Spearman and Kendall correlations to characterize the influence of events on the Brazilian stock market structure*. 2018. 85f. Dissertação (Mestrado em Administração de Organizações) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2018.

In this work the influence of an event on the Brazilian stock market was analyzed from networks and its minimum spanning trees obtained from measures of dependence based on the Pearson, Spearman, and Kendall's correlations. The event considered was the news in the evening of May 17, 2017 in which the owner of the Brazilian company JBS, Joesley Batista, recorded the Brazilian President Michel Temer authorizing the purchase of the silence of a congress member. The day just after the news, May 18, 2017, was defined as the event day. High-frequency data from 58 Ibovespa shares were collected from 11 to 25 May 2017. Changes in the stocks networks were analyzed comparing the period before and after the event in two time scales: (1) Daily networks: five trade sections before the event, the day of the event and, five trade sections after the event, with price every 15 minutes; (2) Grouped before and after do evento: grouping data from 5 days before and 5 days after event. The study of the daily networks indicated a change of trend in their properties during the period that contains the event, with quotations every 15 minutes. The study of daily networks indicated a change of trend in their properties during the period containing the event. This suggested that analysis of the mean effect of grouped data before and after the event could highlight the changes in the network structure. The networks before and after the event showed significant changes in their metrics, which became more evident from the minimum spanning trees. After the event, the minimum spanning trees for grouped data got a smaller number of clusters in the networks for all kind of correlations. The networks generated by Kendall and Spearman correlations presented a larger number of clusters before and after the event. The weighted degree distributions after the event suggest a power law decay tail for all the correlations considered and indicates a higher probability of vertices with weighted degrees far away from the mean weighted degree. The minimum spanning tree metrics generated by Spearman correlation suffered the greatest variation, followed by those of Kendall and Pearson; and their values indicates that after the event the networks became more robust, that is, more rigid. The increase in the networks robustness after the event indicates a higher market connectivity, making it as a whole, more susceptible to the impact of new events.

Keywords: Networks. Minimum spanning tree. Pearson's Correlation. Spearman's Correlation. Kendall's Correlation. Stock Markets. Event.

LISTA DE ILUSTRAÇÕES

Figura 1 – As pontes de Königsberg	21
Figura 2 – Grafo para o problema pontes de Königsberg	21
Figura 3 – Exemplo de rede	22
Figura 4 – Rede com arestas múltiplas e auto arestas	22
Figura 5 – Rede ponderada	24
Figura 6 – Rede direcional	25
Figura 7 – Exemplo de rede tipo árvore	26
Figura 8 – Propriedades básicas de uma árvore.	26
Figura 9 – Exemplo de árvore geradora mínima	27
Figura 10 – Árvore geradora mínima via algoritmo de Kruskal.	28
Figura 11 – Etapas do algoritmo de Kruskal.	28
Figura 12 – Árvore geradora mínima pelo algoritmo de Prim.	29
Figura 13 – Algoritmo de Prim ilustrado por etapas.	29
Figura 14 – Grau de entrada e grau de saída	30
Figura 15 – O impacto da remoção do vértice	33
Figura 16 – Distância Euclidiana entre os pontos no plano.	36
Figura 17 – Variação do Ibovespa em maio de 2017.	39
Figura 18 – Período de coleta e análise dos dados.	39
Figura 19 – Variação do Ibovespa no período de 11/05/2017 a 25/05/2017.	40
Figura 20 – <i>Boxplot</i> do módulo dos coeficientes de correlação.	45
Figura 21 – Distribuição acumulada complementar do grau ponderado das redes completas para os cinco dias antes do evento.	46
Figura 22 – Grau ponderado médio, desvio-padrão e coeficiente de variação diário da rede completa durante todo o período analisado.	47
Figura 23 – Distribuição acumulada complementar do grau ponderado das árvores geradoras mínimas para os cinco dias antes do evento.	49
Figura 24 – Grau ponderado médio, desvio-padrão e coeficiente de variação da árvore geradora mínima durante todo o período analisado.	50
Figura 25 – Comprimento do caminho médio e diâmetro da árvore geradora mínima durante todo o período analisado.	51
Figura 26 – Distribuição acumulada complementar do grau ponderado das redes completas para o período agrupado dos cinco dias antes e cinco dias depois do evento.	52
Figura 27 – Grau ponderado médio, desvio padrão e coeficiente de variação da rede completa para os períodos agrupados antes e depois do evento.	53
Figura 28 – Árvore geradora mínima para correlação de Pearson.	54

Figura 29 – Árvore geradora mínima para correlação de Spearman.	54
Figura 30 – Árvore geradora mínima para correlação de Kendall.	55
Figura 31 – Distribuição acumulada complementar do grau ponderado das árvores geradoras mínimas para o período agrupado dos cinco dias antes e cinco dias depois do evento.	56
Figura 32 – Grau ponderado médio, desvio-padrão e coeficiente de variação da árvore geradora mínima para os períodos agrupados antes e depois do evento. . . .	57
Figura 33 – Comprimento do caminho médio e diâmetro da árvore geradora mínima para o período agrupado em antes e depois do evento.	58
Figura 34 – Grau ponderado $\times$ Grau simples das ações na árvore geradora mínima de (a) Pearson, (b) Spearman e Kendall para os dados agrupados antes e depois do evento.	60
Figura 35 – Parâmetro de Molloy-Reed e robustez das árvores mínimas geradoras para antes e depois do evento.	61
Figura A1 – Árvore geradora mínima de Pearson do quinto dia anterior ao evento.	69
Figura A2 – Árvore geradora mínima de Spearman do quinto dia anterior ao evento. . . .	70
Figura A3 – Árvore geradora mínima de Kendall do quinto dia anterior ao evento.	70
Figura A4 – Árvore geradora mínima de Pearson do quarto dia anterior ao evento.	71
Figura A5 – Árvore geradora mínima de Spearman do quarto dia anterior ao evento. . . .	71
Figura A6 – Árvore geradora mínima de Kendall do quarto dia anterior ao evento.	72
Figura A7 – Árvore geradora mínima de Pearson do terceiro dia anterior ao evento. . . .	72
Figura A8 – Árvore geradora mínima de Spearman do terceiro dia anterior ao evento. . .	73
Figura A9 – Árvore geradora mínima de Kendall do terceiro dia anterior ao evento. . . .	73
Figura A10–Árvore geradora mínima de Pearson do segundo dia anterior ao evento. . . .	74
Figura A11–Árvore geradora mínima de Spearman do segundo dia anterior ao evento. . .	74
Figura A12–Árvore geradora mínima de Kendall do segundo dia anterior ao evento. . . .	75
Figura A13–Árvore geradora mínima de Pearson do primeiro dia anterior ao evento. . . .	75
Figura A14–Árvore geradora mínima de Spearman do primeiro dia anterior ao evento. .	76
Figura A15–Árvore geradora mínima de Kendall do primeiro dia anterior ao evento. . . .	76
Figura A16–Árvore geradora mínima de Pearson do dia do evento.	77
Figura A17–Árvore geradora mínima de Spearman do dia do evento.	77
Figura A18–Árvore geradora mínima de Kendall do dia do evento.	78
Figura A19–Árvore geradora mínima de Pearson do primeiro dia após o evento.	78
Figura A20–Árvore geradora mínima de Spearman do primeiro dia após o evento.	79
Figura A21–Árvore geradora mínima de Kendall do primeiro dia após o evento.	79
Figura A22–Árvore geradora mínima de Pearson do segundo dia após o evento.	80
Figura A23–Árvore geradora mínima de Spearman do segundo dia após o evento.	80
Figura A24–Árvore geradora mínima de Kendall do segundo dia após o evento.	81
Figura A25–Árvore geradora mínima de Pearson do terceiro dia após o evento.	81

Figura A26–Árvore geradora mínima de Spearman do terceiro dia após o evento.	82
Figura A27–Árvore geradora mínima de Kendall do terceiro dia após o evento.	82
Figura A28–Árvore geradora mínima de Pearson do quarto dia após o evento.	83
Figura A29–Árvore geradora mínima de Spearman do quarto dia após o evento.	83
Figura A30–Árvore geradora mínima de Kendall do quarto dia após o evento.	84
Figura A31–Árvore geradora mínima de Pearson do quinto dia após o evento.	84
Figura A32–Árvore geradora mínima de Spearman do quinto dia após o evento.	85
Figura A33–Árvore geradora mínima de Kendall do quinto dia após o evento.	85

LISTA DE QUADROS

Quadro 1 – Ações da carteira teórica Ibovespa – maio a agosto de 2017, atualizada em 28 de agosto de 2017	41
Quadro 2 – Ações com maior grau ponderado em cada grupo (de acordo com a modularidade) antes e depois do evento. As células com as letras <i>AD</i> indicam as ações que permaneceram nas redes antes e depois do evento para cada correlação, como a principal de um grupo. As ações com um asterisco (*) são aquelas que aparecem em todas as redes antes do evento como a principal de um grupo (maior grau ponderado). As ações com dois asteriscos (**) são as que aparecem em todas as redes após o evento como a principal de um grupo	55
Quadro 3 – Ações com maior grau ponderado em cada grau simples antes e depois do evento. As células com as letras <i>AD</i> indicam as ações que permaneceram nas redes antes e depois do evento para cada correlação como a principal de um grupo. As ações com um asterisco (*) são aquelas que aparecem em todas as redes antes do evento como a principal de um grupo. As ações com dois asteriscos (**) são as que aparecem em todas as redes após o evento como a principal de um grupo	59

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Problema de Pesquisa	19
1.2	Objetivos	19
1.2.1	Gerais	19
1.2.2	Específicos	19
1.3	Justificativa	19
1.4	Estrutura do trabalho	20
2	CONCEITOS BÁSICOS DE TEORIA DE REDES	21
2.1	Redes Ponderadas	23
2.2	Redes Direcionadas	24
2.3	Redes tipo árvore	25
2.3.1	Árvore Geradora Mínima	25
2.4	Principais métricas de redes	30
2.4.1	Grau	30
2.4.2	Comprimento do caminho médio	31
2.4.3	Proximidade	31
2.4.4	Intermediação	31
2.4.5	Modularidade	32
2.5	Componente gigante e robustez	33
3	CORRELAÇÕES E DISTÂNCIAS	34
3.1	Correlação	34
3.2	Medidas de distância	36
4	METODOLOGIA	38
4.1	Comentários sobre o período de análise	38
4.2	Base de dados	39
4.3	A estrutura de rede	40
5	RESULTADOS	44
5.1	Rede diárias: a rede completa	44
5.2	Redes diárias: Árvore geradora mínima	48
5.3	Rede completa para dados agrupados antes e depois do evento	50
5.4	Árvore geradora mínima para dados agrupados antes e depois do evento	52
6	CONSIDERAÇÕES FINAIS	62
	REFERÊNCIAS	63

1 INTRODUÇÃO

Recentemente, a abordagem de redes envolvendo relações entre indivíduos, empresas, países, ações, tem se tornado um tema relevante de pesquisa. As técnicas de teoria de redes fornecem um método alternativo ou complementar de pesquisa que tem se mostrado eficiente para caracterizar e revelar padrões em sistemas sociais, políticos, econômicos e financeiros.

Embora o estudo das redes tenha uma longa história, com raízes na teoria dos grafos e na sociologia, o capítulo moderno da ciência em rede surgiu apenas na primeira década do século XXI. As primeiras pesquisas sobre redes foram desenvolvidas nas décadas de 1950 e 1960 por Solomonoff e Rapoport (1951), Rapoport (1977) e Erdős e Rényi (1959), Erdős e Rényi (1960), Erdős e Rényi (1960 b). Após este período, outras pesquisas surgiram de forma interdisciplinar. Alguns exemplos são, nas redes sociais as conexões entre os indivíduos podem indicar, por exemplo, relações de amizade (RAPOPORT; HORVATH, 1961; GRANOVETTER, 1973), citações em trabalhos científicos (PRICE, 1965) afiliação de órgãos sociais e diretores (ROBINS; ALEXANDER, 2004), a ligação entre os criminosos e os crimes praticados (FRANK; CARRINGTON, 2007), a relação entre atores e equipes de trabalho (GUIMERA; SALES-PARDO; AMARAL, 2007), os *hiperlinks* entre páginas da *web* (EASLEY; KLEINBERG, 2010). Na área de logística, por exemplo, os voos que operam entre aeroportos (BAGLER, 2008). Uma área de aplicação recente é a área da saúde, como por exemplo a associação entre doenças (GOH et al., 2007; BARABÁSI; GULBAHCE; LOSCALZO, 2011), mapeamento do cérebro humano (SPORNS; TONONI; KÖTTER, 2005) e de animais (OH et al., 2014).

Os trabalhos de Erdős e Rényi (1959), que marcou o estudo dos grafos aleatórios e Granovetter (1973), o artigo sobre redes sociais mais citado, foram altamente considerados dentro de sua disciplina, mas tiveram impacto limitado fora de seu campo. No século XXI esses artigos tiveram um crescimento considerável de citações, mostrando o surgimento da ciência em rede, atraindo uma atenção nova e interdisciplinar para essas publicações clássicas. O artigo de Watts e Strogatz (1998) sobre as redes de pequeno-mundo e Barabási e Albert (1999) em redes de livre escala também estão entre os trabalhos mais citados da teoria de redes. A natureza comum de muitas questões levou a uma interdisciplinaridade de ferramentas e ideias, por exemplo, o conceito de centralidade de intermediação surgido na literatura de redes sociais nos anos 70, hoje é importante para identificar nós de alto tráfego na Internet, da mesma forma, algoritmos que foram desenvolvidos inicialmente por cientistas da computação, hoje, são úteis na identificação de módulos de doenças na medicina ou na detecção de comunidades dentro de grandes redes sociais.

O estudo de uma rede complexa aplicada ao mercado de ações pode ajudar a entender onde uma falha ou ruptura nas conexões podem se espalhar pela rede, formando o efeito cascata e, conseqüentemente, uma crise financeira generalizada (EASLEY; KLEINBERG, 2010). No mercado financeiro, o trabalho de Mantegna (1999) foi pioneiro, introduzindo a árvore geradora mínima para o estudo da estrutura do mercado de ações com suas interações mediadas por correlações das séries temporais dos preços das ações. A pesquisa foi realizada com os retornos

das ações das empresas que compõem os índices Dow Jones Industrial Average (DJIA) e Standard and Poor's 500 (S&P500) da Bolsa de Nova York. Os resultados mostraram uma estrutura hierárquica do agrupamento das ações por setores que pode ser útil para explicar fatores econômicos que afetam grupos específicos de ações. Estudos semelhantes foram realizados para o mercado de ações brasileiro (TABAK; SERRA; CAJUEIRO, 2010), o europeu (GILMORE; LUCEY; BOSCIA, 2008), o americano (POZZI; MATTEO; ASTE, 2008; ZENG et al., 2016; HUANG et al., 2017; TUMMINELLO et al., 2007b), o coreano (JUNG et al., 2006; NOBI et al., 2015), o chinês (HUANG; ZHUANG; YAO, 2009), de vários países (COELHO et al., 2007) e o de mercadorias (SIECZKA; HOŁYST, 2009).

Outras pesquisas começaram a surgir utilizando a árvore geradora mínima. Vandewalle et al. (2001) construíram uma árvore geradora mínima para 6358 observações de preços de ações norte-americanas durante o ano de 1999. Bonanno et al. (2003) analisaram a estrutura de uma árvore geradora mínima de um grupo de ações negociados na Bolsa de Valores de Nova York durante um período de 12 anos. Miccichè et al. (2003) investigaram árvores geradoras mínimas obtidas utilizando um procedimento de agrupamento baseado em correlação de séries temporais retorno e de volatilidade de ativos financeiros.

Peralta e Zareei (2016) mostraram a relação negativa entre a centralidade dos ativos da árvore geradora mínima de mercados financeiros americano e europeu e suas proporções segundo a teoria de formação de carteiras de Markowitz (MARKOWITZ, 1952). Onnela et al. (2003) estudaram as propriedades da árvore geradora mínima do mercado financeiro em relação a formação de carteiras de investimento, resultando no fato de que os ativos da carteira obtida pela teoria clássica de Markowitz estão localizados na periferia da árvore (ONNELA et al., 2003).

Ortega e Matesanz (2006) utilizaram dados de uma amostra de 28 países para classificar as consequências das crises cambiais com base na análise de agrupamentos da série histórica do câmbio real. A análise forneceu informações relevantes sobre ligações entre países e uma visão significativa sobre o fenômeno de contágio de crises financeiras. Garas e Argyrakis (2007) estudaram as propriedades de três diferentes carteiras negociadas na Bolsa de Valores de Atenas (ASE) no período 1987-2004. Os autores utilizaram a árvore geradora mínima para analisar evolução temporal das carteiras e do mercado como um todo e identificaram as mudanças nas carteiras de investimento associadas a uma crise no mercado. Bonanno et al. (2004) investigaram carteira de ações do mercado de capitais americano, em períodos diferentes, índices financeiros e séries temporais de volatilidade e mostraram que informações econômicas significativas podem ser extraídas de matrizes de correlação. Nobi et al. (2015) estabeleceram uma estrutura de rede do mercado de ações coreano, com uma árvore geradora mínima elaborada, também, a partir da matriz de correlação. O artigo de Nobi et al. (2015) considerou os preços das ações no mercado de ações coreano durante a crise, concentrando-se em três períodos: antes, durante e depois da crise, e observou que coeficiente de correlação aumentou durante a crise, indicando um aumento na estrutura durante este período e um comportamento similar entre as empresas.

O artigo de Naylor, Rose e Moyle (2007) utilizou a árvore geradora mínima e análise de agrupamentos para extrair um mapa de influência das principais moedas para o período de 1995-2001. Kwapien, Gworek e Drozd (2009) estudaram a estrutura e evolução temporal da moeda estrangeira usando teoria de redes, com base em taxas de câmbio de um conjunto de 46 moedas. Sieczka e Hołyst (2009) analisaram as dependências nos mercados de *commodities*, investigando a correlação dos contratos futuros entre os anos de 1998 e 2007 através da árvore geradora mínima.

Para o mercado brasileiro, Tabak, Serra e Cajueiro (2009) investigaram estrutura de rede da taxa de juros brasileira. Posteriormente, Tabak, Serra e Cajueiro (2010) investigaram as propriedades da árvore geradora mínima, utilizando a matriz de correlação para ações de diferentes setores. Os diversos setores também foram estudados no artigo de Pozzi, Matteo e Aste (2008) utilizando a árvore geradora mínima gerada a partir de correlações entre os 300 títulos de ações da Bolsa de Valores de Nova Iorque (NYSE) mais capitalizados, no período de 2001 a 2003, para estudar a estrutura e a distribuição da rede em relação aos setores econômicos.

Para a elaboração da rede tipo árvore geradora mínima, é necessária uma matriz de distâncias entre objetos. Pesquisas com dados financeiros costumam definir uma medida de distância a partir do coeficiente de correlação de Pearson calculado entre os logaritmos dos preços diários de ações (MANTEGNA, 1999). Em geral, os estudos de redes financeiras são baseados no coeficiente de correlação de Pearson, que mede a dependência linear entre pares de variáveis. Entretanto, existem outras medidas de correlação – Spearman e Kendall – baseadas em postos, que podem ser alternativas interessantes para o estudo de redes e são menos sensíveis a *outliers*¹, pois delimitam o *outlier* ao valor de seu posto. As correlações baseadas em postos verificam a semelhança entre a ordem dos dados, é igual à correlação de Pearson entre os valores dos postos de duas variáveis, ou seja, de suas posições ordenadas.

Assim, faz-se importante o estudo e a comparação da estrutura de rede gerada por diversas medidas de correlação.

Além disso, existem outros aspectos do mercado financeiro que podem ser explorados com técnicas de redes. Neste trabalho em particular, estamos interessados em obter informações sobre a estrutura do mercado de ações brasileiro quando sofre a influência de um evento. Um evento é um acontecimento que pode ser observado em um determinado momento, como por exemplo, um desastre natural, uma guerra, um acontecimento político, uma especulação de mercado, uma operação financeira ou até mesmo uma notícia. O evento analisado neste trabalho é o efeito de uma notícia sobre o mercado de ações, pois notícias de cunho econômico-político podem influenciar na variação do preço das ações (CUTLER; POTERBA; SUMMERS, 1988; SILVA; CARVALHO; NUNES, 2013). Assim, foi definido como evento um fato ocorrido em 2017, quando o Jornal O Globo publicou na noite de 17 de maio de 2017, uma reportagem dizendo que o dono da JBS, Joesley Batista, gravou o atual Presidente da República, Michel Temer, autorizando a compra do silêncio de Eduardo Cunha (JARDIM, 2017). A reportagem

¹ Em matemática, *outlier* é definido como valor diferente ou atípico em um conjunto de dados, é uma observação que apresenta um grande afastamento das demais.

diz, ainda, que a gravação também inclui o ex candidato a presidência da república, Aécio Neves, pedindo dinheiro ao empresário Joesley Batista. O dia do evento foi definido como o dia 18 de maio, devido às mudanças causadas no mercado de ações brasileiro. O objetivo é verificar como as propriedades da rede são alteradas, comparando-se um período antes e um período depois do evento e se este imprime alguma característica nas redes obtidas que identifiquem mudanças estruturais do mercado.

1.1 Problema de Pesquisa

O problema de pesquisa deste trabalho está relacionado à construção de redes, em particular, de árvore geradora mínima de ações, para explorar a influência de um determinado evento sobre a estrutura do mercado de ações brasileiro. Além disso, a pesquisa pretende comparar diferenças na estrutura das redes obtidas por medidas de dependência baseadas nas correlações de Pearson, de Spearman e de Kendall.

1.2 Objetivos

1.2.1 Gerais

O objetivo geral desta pesquisa é obter e comparar redes ponderadas por correlações de Pearson, de Spearman e de Kendall para o mercado financeiro brasileiro, e também, analisar como estas podem revelar o efeito de um evento sobre a estrutura do mercado, comparando-se um período antes e um período depois do evento.

1.2.2 Específicos

- Obter as redes completas e as redes do tipo árvore geradora mínima dos retornos de ações das empresas da BM&FBovespa a partir das medidas de dependência baseada nas correlações de Pearson, de Spearman e de Kendall.
- Caracterizar as redes obtidas através das métricas de teoria de redes.
- Identificar os agrupamentos das redes.
- Comparar os resultados obtidos para as redes formadas com as medidas de dependência consideradas.
- Explorar as alterações nas redes sob influência de um evento.

1.3 Justificativa

Em pesquisas de redes financeiras, o coeficiente de correlação de Pearson tem sido usado como medida da intensidade da ligação entre ações de empresas que fazem parte do mercado fi-

nanceiro (TABAK; SERRA; CAJUEIRO, 2010; DJAUHARI; GAN, 2015; DJAUHARI; GAN, 2016; HUANG et al., 2017). Entretanto, a correlação de Pearson é adequada como medida de uma dependência linear. Existem outras medidas de correlação pouco exploradas e que podem ser interessantes no estudo de redes, pois não requerem a suposição de que a relação entre os dados seja linear. As correlações por postos de Spearman e Kendall analisam a relação, entre duas variáveis, não linear mas monótona (se uma aumenta a outra tem sempre tendência a aumentar, ou a diminuir).

Miccichè et al. (2003) utilizaram o coeficiente de correlação de Spearman para a construção da árvore geradora mínima entre retornos de preços e volatilidade dos preços de ações. Entretanto, até o vai o conhecimento desta pesquisadora, nenhuma comparação entre redes obtidas com correlações de Kendall, de Spearman e de Pearson foi realizada.

Justifica-se, portanto, a importância desta análise, pois grande parte das pesquisas utilizam a correlação de Pearson para a construção de redes. Assim, pretende-se estudar a existência de diferenças na estrutura de rede obtidas considerando-se as medidas de correlação que permitam definir a estrutura de mercado sob a influência de um evento.

1.4 Estrutura do trabalho

Esta dissertação está organizada da seguinte maneira: no Capítulo 2 são abordados os conceitos básicos de teoria de redes. No Capítulo 3 são apresentadas as correlações de Pearson, Kendall e Spearman e a definição de medidas de distância baseada em correlações. O Capítulo 4 apresenta os métodos utilizados no desenvolvimento desta pesquisa. Os resultados obtidos são apresentados no Capítulo 5. Finalmente, o Capítulo 6 apresenta as considerações finais.

2 CONCEITOS BÁSICOS DE TEORIA DE REDES

O estudo de redes teve início no século dezoito, quando o matemático Leonhard Euler resolveu o problema das pontes de Königsberg. A cidade da Prússia tinha sete pontes que atravessavam o rio Pregel ligando duas ilhas. A situação está esquematizada na Figura 1 onde as sete pontes ligam a ilha A aos lados B, C e D de terra. O problema consistia em partir de um determinado local e retornar a ele mesmo atravessando as sete pontes sem cruzar uma mesma ponte duas vezes (NEWMAN, 2010; BARABÁSI, 2016), até que Euler demonstrou, matematicamente, a impossibilidade de cruzar as pontes uma única vez para o caso considerado.

Figura 1 – As pontes de Königsberg
Fonte: Adaptado de Barabási (2016, p.24)

Assim, foi introduzindo um grafo para representar o problema. Um grafo é um diagrama contendo vértices e arestas. Euler imaginou quatro vértices (pontos) representando as áreas de terra da cidade e sete arestas (linhas), representando as pontes (NEWMAN, 2010; BARABÁSI, 2016). A Figura 2 representa o grafo elaborado por Euler.

Figura 2 – Grafo para o problema das Pontes de Königsberg
Fonte: Barabási (2016, p.24)

Uma rede é um conjunto de elementos discretos chamados vértices (ou nós) e um conjunto de arestas (ou ligações) que ligam os vértices (NEWMAN, 2010; BARABÁSI, 2016). A Figura 3 mostra uma possível representação de uma rede, ou grafo, como é conhecido em Matemática.

Os vértices podem representar os mais variados objetos e as arestas representam a ligação entre eles. Assim, numa rede de amizades, os vértices poderiam representar pessoas e as arestas, a existência de amizade entre elas; numa rede de negócios os vértices podem representar os ativos de uma empresa (ações) e as arestas, a existência de relações comerciais entre elas.

A rede representada na Figura 3 é chamada de rede simples, pois tem arestas simples entre dois vértices. Em alguns casos pode haver mais do que uma aresta entre o mesmo par

Figura 3 – Exemplo de uma rede composta de oito vértices e dez arestas
Fonte: Elaborado pela autora.

de vértices, que são chamadas de “arestas múltiplas”. Existem redes que apresentam “auto arestas”, isto é, arestas conectando vértices consigo mesmo. A Figura 4 mostra uma rede com arestas múltiplas e auto arestas.

Figura 4 – Exemplo de rede com arestas múltiplas e auto arestas
Fonte: Elaborado pela autora.

A estrutura de uma rede é obtida se existe ou não ligação entre dois vértices. Assim, a rede pode ser representada matematicamente pela matriz de adjacência. Para uma rede não direcionada, se há uma ligação entre dois vértices i e j , atribui-se o valor 1, e quando não há ligação, atribui-se o valor 0. Portanto, a matriz de adjacência A , para uma rede simples, é uma matriz com elementos

$$A_{ij} = \begin{cases} 1 & \text{se houver aresta entre } i \text{ e } j; \\ 0 & \text{se não houver aresta entre } i \text{ e } j. \end{cases} \quad (2.1)$$

Por exemplo, para o grafo da Figura 3 a matriz de adjacência tem a seguinte forma

$$\mathbf{A}_{Ex} = \begin{pmatrix} & V_1 & V_2 & V_3 & V_4 & V_5 & V_6 & V_7 & V_8 \\ V_1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ V_2 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ V_3 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ V_4 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ V_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ V_6 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ V_7 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ V_8 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix} \quad (2.2)$$

onde os vértices são indicados como $V_1, V_2, V_3, V_4, V_5, V_6, V_7$ e V_8 .

A matriz de adjacência também pode ser usada para definir redes com arestas múltiplas e auto arestas. Na matriz das redes multigrafo, A_{ij} corresponde ao número de arestas entre os vértices i e j . Para as redes com auto arestas, o elemento corresponde a diagonal de A e é igual a 2, pois de i a i existem 2 extremidades ligadas ao vértice i . Ou seja, as arestas simples aparecem duas vezes na matriz de adjacência, significando que A_{ij} e A_{ji} são iguais a 1. Portanto, para contar as auto arestas da mesma maneira, estas também devem aparecer duas vezes, e existindo apenas um elemento diagonal A_{ii} , o elemento corresponderá a duas ligações (NEWMAN, 2010). Para o exemplo da Figura 4 a matriz de adjacência é dada por

$$A = \begin{pmatrix} & V_1 & V_2 & V_3 & V_4 & V_5 & V_6 \\ V_1 & 0 & 1 & 0 & 0 & 3 & 0 \\ V_2 & 1 & 2 & 2 & 1 & 0 & 0 \\ V_3 & 0 & 2 & 0 & 1 & 1 & 1 \\ V_4 & 0 & 1 & 1 & 0 & 0 & 0 \\ V_5 & 3 & 0 & 1 & 0 & 0 & 0 \\ V_6 & 0 & 0 & 1 & 0 & 0 & 2 \end{pmatrix} \quad (2.3)$$

2.1 Redes Ponderadas

As redes simples são representadas por matrizes de adjacência com elementos 0 ou 1, isto é, que existe ou não ligação entre vértices. Todavia, pode ser interessante considerar que a ligação entre os vértices possui um grau de intensidade. Neste caso, temos as redes ponderadas que são representadas por uma matriz de adjacência cujos elementos indicam a intensidade da ligação entre os vértices (EASLEY; KLEINBERG, 2010; NEWMAN, 2010; BARABÁSI, 2016). A equação 4 apresenta um exemplo de matriz de adjacência de uma rede ponderada (NEWMAN, 2010).

Figura 5 – Rede ponderada baseada na matriz de adjacência

Fonte: Elaborada pela autora.

$A =$

$$\begin{pmatrix} & V_1 & V_2 & V_3 \\ V_1 & 0 & 3 & 1.5 \\ V_2 & 3 & 0 & 1 \\ V_3 & 1.5 & 1 & 0 \end{pmatrix} \quad (2.4)$$

A Figura 5 representa o grafo da rede ponderada a partir da matriz de adjacência 2.4. Note que é possível indicar o grau de ligação com a espessura das arestas. Assim, a espessura das arestas indica que a conexão entre os vértices 1 e 2 é duas vezes mais forte que a conexão entre 1 e 3, que por sua vez, é mais forte que a conexão entre 2 e 3. A intensidade da conexão em uma rede ponderada, também é conhecida como peso. Os números atribuídos são geralmente números inteiros positivos, mas não há impedimento para que o número seja negativo. Em uma rede social, por exemplo, os pesos positivos representam a amizade e os negativos, a inimizade (NEWMAN, 2010).

Redes ponderadas aparecem em diversas situações. Numa rede social, o grau das ligações pode ser caracterizado por uma combinação de tempo, emoção, intimidade e reciprocidade (GRANOVETTER, 1973). Em uma rede de neurônios, o grau de ligação pode ser a quantidade de sinapse (WATTS; STROGATZ, 1998). Numa rede de aeroportos, o número de voos que operam entre aeroportos (BAGLER, 2008). No mercado financeiro, a proporção total do mercado de ativos controlados por um banco através de relações societárias (PECORA; SPELTA, 2015), as taxas inflação de um período (KANDA et al., 2016) ou ainda os coeficientes de correlação entre variáveis (MANTEGNA, 1999; BONANNO; LILLO; MANTEGNA, 2001b; BONANNO; LILLO; MANTEGNA, 2001a; VANDEWALLE et al., 2001; ONNELA et al., 2003; ONNELA; KASKI; KERTÉSZ, 2004; BONANNO et al., 2004; JUNG et al., 2006; COELHO et al., 2007; TUMMINELLO et al., 2007a; TUMMINELLO et al., 2007b; HUANG; ZHUANG; YAO, 2009; TABAK; SERRA; CAJUEIRO, 2010; DJAUHARI; GAN, 2015; DJAUHARI; GAN, 2016; DEVIREN; DEVIREN, 2016; HUANG et al., 2017).

2.2 Redes Direcionadas

Nas seções anteriores foram apresentadas redes em que suas arestas não tinham direção alguma, com uma ligação simples entre os vértices. Nesta seção, será apresentado o conceito

de redes direcionadas.

Redes direcionadas são redes em que cada aresta está direcionada a um vértice, ou seja, possui setas que apontam a direção para um determinado vértice (NEWMAN, 2010; BARABÁSI, 2016), conforme pode ser observado na Figura 6.

Figura 6 – Exemplo de rede direcionada

Fonte: Elaborada pela autora.

A matriz de adjacência A de uma rede direcionada tem elementos

$$A_{ij} = \begin{cases} 1 & \text{se houver aresta de } j \text{ para } i; \\ 0 & \text{se não houver aresta de } j \text{ para } i. \end{cases} \quad (2.5)$$

As redes dirigidas podem representar os voos de um aeroporto para outro (BAGLER, 2008), os *hyperlinks* em uma página da internet que direcionam para uma outra página da *web* (EASLEY; KLEINBERG, 2010), redes de citação de autores (YU; SHI, 2015) ou até mesmo, redes de investidores ou controladores i de uma determinada empresa j (PECORA; SPELTA, 2015).

2.3 Redes tipo árvore

Uma árvore consiste em um grafo de rede conexo, não direcionado, que não possui ciclos e há exatamente um caminho entre qualquer par de vértices (NEWMAN, 2010; BARABÁSI, 2016). A Figura 7 mostra o grafo do tipo árvore, que pode ser desenhado de modo ramificado, com um vértice raiz no topo e uma estrutura de ramificação no sentido inferior, e os vértices, na parte de baixo que estão conectados apenas a um outro vértice, são chamados de folhas.

As árvores apresentam duas propriedades básicas ilustradas na Figura 8 (SEDGEWICK; WAYNE, 2011): a adição de uma aresta na árvore cria um ciclo, enquanto que a remoção de uma aresta, quebra a rede em duas partes.

2.3.1 Árvore Geradora Mínima

As árvores geradoras criam um subgrafo que reflete o grafo original, ou seja, uma árvore geradora de um grafo G é um subgrafo de G que contém todos os vértices do grafo original conectados entre si por uma única aresta (WU; CHAO, 2004).

Figura 7 – Exemplo de rede tipo árvore
Fonte: Elaborada pela autora.

Figura 8 – Propriedades básicas de uma árvore. (a) Adição de aresta. (b) Remoção de aresta.
Fonte: Adaptado de Sedgewick e Wayne (2011).

As árvores podem ser utilizadas em muitas aplicações, como por exemplo, em um mapa aéreo, onde as arestas ponderadas representam as distâncias ou as tarifas de voo, ou em um circuito elétrico, onde as arestas podem representar o comprimento do fio, o custo ou, o tempo de propagação de sinal (SEDGWICK; WAYNE, 2011).

Para se obter um custo, tempo ou comprimento mínimo, a otimização pode ser feita através da obtenção de uma árvore geradora mínima (*Minimum Spanning Tree* - MST, em inglês). A árvore geradora mínima de uma rede ponderada é uma rede não direcionada, com todos os vértices conectados por uma única aresta, a qual representa o menor peso possível (WU; CHAO, 2004; SEDGWICK; WAYNE, 2011).

A Figura 9-(a) apresenta uma rede com 250 vértices e 1273 arestas, enquanto que, a Figura 9-(b) mostra a árvore geradora mínima com as ligações destacadas.

A árvore geradora mínima é um dos problemas mais conhecidos de otimização combinatória, onde os métodos para sua solução desempenharam um papel importante na construção dos algoritmos computacionais ao longo dos anos, desenvolvidos para encontrar a árvore mínima geradora de forma eficiente (GRAHAM; HELL, 1985; JUNGnickel, 2008). O problema teve início na década de 1920 quando Boruvka identificou e resolveu o problema da eletrificação da área rural da região de Morávia, na República Tcheca (JAYAWANT; GLAVIN et al., 2009). Assim, o desafio era juntar os pontos (municípios) através de uma rede de modo que quais-

Figura 9 – (a) Figura de uma rede com 250 vértices e 1273 arestas. (b) árvore geradora mínima da rede (a).

Fonte: Adaptado de Sedgewick e Wayne (2011)

quer dois pontos sejam unidos diretamente ou indiretamente por um único caminho, e que, o comprimento e o custo total da rede fossem o menor possível (MILKOVÁ, 2007). Nessa rede ponderada, os vértices representavam as cidades, as arestas representavam a ligação entre as cidades e os pesos representam os custos ou comprimento dos fios (GRAHAM; HELL, 1985).

Na década de 1950, muitas pessoas contribuíram para o problema da árvore geradora mínima, entre eles Kruskal e Prim, cujos algoritmos são usados para a construção da árvore geradora mínima (JAYAWANT; GLAVIN et al., 2009).

O algoritmo de Kruskal se baseia no teorema de que há apenas uma árvore de cobertura do grafo, cuja soma dos pesos das arestas seja a menor possível. Para implementar este algoritmo, utiliza-se a “matriz de distâncias” do grafo, ou seja, a matriz com elementos a_{ij} onde a_{ij} é a distância (peso) da aresta entre os vértices i e j (KRUSKAL, 1956).

A Figura 10 ilustra, de forma geral, a formação da árvore geradora mínima pelo algoritmo de Kruskal. O algoritmo de Kruskal busca as arestas por toda a extensão do grafo e seleciona as que tenham a menor distância (peso), em ordem crescente, até que todos os vértices estejam conectados e assim, a árvore esteja completa.

O algoritmo de Kruskal calcula a árvore geradora mínima de qualquer grafo ponderado, somando as menores arestas sucessivamente, até que todos os vértices estejam conectados pela árvore geradora mínima. As arestas são selecionadas em ordem crescente, desde que a próxima aresta a ser considerada não forme um ciclo com outras arestas já selecionadas (EISNER, 1997). A Figura 11 ilustra a seleção das arestas de maneira detalhada, conforme o algoritmo de Kruskal. Na Figura 11(a) as arestas em vermelho indicam que são as próximas a serem selecionadas e as arestas em preto já estão integradas ao grafo. A Figura 11(b) mostra os pesos das arestas em ordem crescente, os pesos em preto são das arestas que foram integradas ao grafo, os pesos em cinza foram descartados, pois formavam um ciclo.

Outro algoritmo para obter a árvore mínima é o algoritmo de Prim, que constrói uma árvore geradora mínima identificando uma aresta com peso mínimo incidente no vértice inicial e espalhando a árvore. Em cada passo, o algoritmo encontra e acrescenta a aresta de menor peso

Figura 10 – Formação da árvore mínima geradora pelo algoritmo de Kruskal.
Fonte: Adaptado de Sedgewick e Wayne (2011).

Figura 11 – Algoritmo de Kruskal ilustrado por etapas. (a) Ilustração da seleção das arestas, conforme o algoritmo. (b) Peso das arestas.

Fonte: Adaptado de Sedgewick e Wayne (2011).

que liga um vértice já pertencente à árvore geradora mínima a um outro vértice que ainda não pertence. O procedimento é concluído uma vez que o conjunto de vértices da árvore é igual ao conjunto de vértices do grafo original, ou seja, a árvore abrange todos os vértices (JAYAWANT; GLAVIN et al., 2009). De forma diferente do algoritmo de Kruskal, o algoritmo de Prim forma a árvore geradora mínima de forma ramificada. Conforme pode ser observado na Figura 12, as arestas são selecionadas conforme o menor peso, entretanto são selecionadas apenas as arestas vizinhas aos vértices já selecionados.

A Figura 13 representa a construção do algoritmo de Prim. Cada desenho representa

Figura 12 – Formação da árvore mínima geradora pelo algoritmo de Prim.
Fonte: (SEdgeWICK; WAYNE, 2011).

o grafo e ao lado a lista de arestas por ordem crescente de peso. O algoritmo é construído da seguinte forma: Primeiramente é selecionado um vértice aleatoriamente para ser a raiz da árvore. O passo seguinte é selecionar os vértices vizinhos à raiz por ordem crescente de peso das arestas, depois, seleciona-se o próximo vértice vizinho à raiz ou ao vértice já selecionado, e assim sucessivamente até que todos os vértices da rede sejam selecionados (CORMEN, 2009).

Figura 13 – Algoritmo de Prim ilustrado por etapas.
Fonte: Sedgewick e Wayne (2011).

2.4 Principais métricas de redes

Esta seção apresenta algumas métricas para a caracterização de uma rede. As métricas de centralidade buscam medir o grau de importância de um elemento da rede.

2.4.1 Grau

O grau mede o número de arestas conectadas a um vértice, ou seja, o grau k_i de um vértice i , indica as conexões de i com outros vértices da rede (NEWMAN, 2010; BARABÁSI, 2016). Para um grafo não direcionado, com n vértices o grau pode ser escrito a partir dos elementos da matriz de adjacência A_{ij} pela expressão

$$k_i = \sum_{j=1}^n A_{ij}. \quad (2.6)$$

O grau pode ser usado como uma medida de importância. Interpreta-se que o vértice com o maior grau é o mais importante da rede, porque tem o maior número de conexões.

Para as redes direcionadas é diferente, pois cada vértice tem dois graus. Neste caso, define-se o grau de entrada (*in-degree*) k_i^{in} que representa o número de arestas que apontam para o vértice i , e o grau de saída (*out-degree*) k_j^{out} que representa o número de arestas que apontam do vértice i para outros vértices (NEWMAN, 2010; BARABÁSI, 2016).

Figura 14 – Representação do grau de entrada e grau de saída em redes direcionadas.

Fonte: Elaborado pela autora.

Para um grafo direcionado com n vértices e matriz de adjacência com elementos $A_{ij} = 1$ se existe uma aresta de j até i , o grau de entrada e o grau de saída são dados pelas expressões

$$k_i^{in} = \sum_{j=1}^n A_{ij} \quad (2.7)$$

e

$$k_j^{out} = \sum_{i=1}^n A_{ij}, \quad (2.8)$$

respectivamente.

Assim, interpreta-se que, o vértice com o maior grau de entrada é o mais importante da rede, porque tem o maior número de conexões de outros vértices com ele.

Note que para redes binárias o grau é uma soma de pesos com valor 1. No caso de redes ponderadas o conceito de grau é generalizado para incluir a intensidade, isto é, o grau de um dado vértice é a soma de intensidades, e em geral, é chamado de intensidade de vértice (NEWMAN, 2004; PARK; LAI; YE, 2004; BARTHELEMY et al., 2005). A partir dos

graus é possível obter a distribuição de grau de uma rede, isto é, a probabilidade $P(k)$ de um vértice ter grau k . No caso de redes ponderadas a distribuição de grau é uma distribuição de intensidade $P(A_{ij})$ das ligações de um dado vértice (NEWMAN, 2004; PARK; LAI; YE, 2004; BARTHELEMY et al., 2005).

2.4.2 Comprimento do caminho médio

O caminho com menor número de ligações possíveis entre dois vértices i e j é chamado de caminho geodésico, ou também, distância geodésica, e é denotado por d_{ij} (NEWMAN, 2010). Considerando um determinado vértice i , a média sobre todos os vértices j , desta distância geodésica entre i e j , é dada por

$$\ell_i = \frac{1}{(n-1)} \sum_{j(\neq i)} d_{i,j}. \quad (2.9)$$

Este resultado indica uma distância média do vértice i aos outros vértices da rede e pode ser utilizado para obter uma medida de proximidade. Ao calcular a média de todos os valores de ℓ_i , obtém-se o comprimento do caminho da rede, definido por

$$\ell = \frac{1}{n} \sum_i \ell_i. \quad (2.10)$$

Esta medida pode ser utilizada para obter-se o grau de separação médio entre dois vértices na rede.

2.4.3 Proximidade

A proximidade (*Closeness Centrality*, em inglês) C_i é uma medida de centralidade frequentemente utilizada no estudo de redes. A proximidade é a média do inverso das distâncias entre todos os vértices, em que a distância a partir de um vértice i para outro j é definido como o comprimento, medido pelo número de ligações do caminho geodésico de um vértice para o outro, que é dado por

$$C'_i = \frac{1}{(N-1)} \sum_{j \neq i}^N \frac{1}{d_{ij}} \quad (2.11)$$

Note que na equação 2.11 exclui-se o vértice i da soma. Assim, interpreta-se que, quanto maior a proximidade de um vértice, mais influente ele é.

2.4.4 Intermediação

A medida de intermediação (*Betweenness*, em inglês), foi inicialmente utilizada no trabalho de Bavelas (1948), mas se tornou conhecida com o trabalho de Freeman (1977). A intermediação mede o quanto um vértice encontra-se nos caminhos entre outros vértices, assim, os

vértices com alta centralidade de intermediação podem ter uma influência considerável dentro de uma rede devido ao controle sobre a informação que passa para outros vértices (NEWMAN, 2010).

Matematicamente, considera-se n_{st}^i igual a 1 para um determinado vértice i , se este vértice está no caminho geodésico entre os vértices s e t , e igual a 0 se não está no caminho geodésico (NEWMAN, 2010; BARABÁSI, 2016). Assim, a intermediação x_i é definida por

$$x_i = \sum_{st} n_{st}^i \quad (2.12)$$

A intermediação pode ser interpretada como a quantidade de vezes que o vértice i aparece nos caminhos geodésicos. Quanto maior a intermediação, maior a centralidade do vértice.

De modo geral, para calcular a intermediação de um vértice i em uma rede dirigida, onde n_{st}^i é o número de caminhos geodésicos de s até t que passa por i , e g_{st} é o número total de caminhos geodésicos de s até t será

$$x_i = \sum_{st} \left(\frac{n_{st}^i}{g_{st}} \right) \quad (2.13)$$

A medida de intermediação também pode ser aplicada em redes direcionadas. Neste caso, o caminho mais curto entre dois vértices depende da direção seguida. Por exemplo, o menor caminho do vértice A ao vértice B pode ser diferente do menor caminho do vértice B ao vértice A , e ainda, poderá existir caminho em uma direção e não existir na direção contrária (NEWMAN, 2010; BARABÁSI, 2016). A intermediação não mede o quão ligado o vértice está, mas mede o quanto o vértice aparece entre os outros. Assim, um vértice com baixo grau, que pode estar conectado a outros com baixo grau, e até mesmo ter um caminho longo, e ainda pode ter alta intermediação (NEWMAN, 2010).

2.4.5 Modularidade

A modularidade é frequentemente utilizada para identificar comunidades na estrutura de redes. A modularidade é definida como a fração das arestas dentro das comunidades menos o valor esperado da fração de arestas inseridas aleatoriamente. Matematicamente, a modularidade Q é calculada pela expressão

$$Q = \frac{1}{2m} \sum_{ij} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta_{c_i, c_j} \quad (2.14)$$

onde A_{ij} representa o número de arestas entre os vértices i e j , o produto $k_i k_j$ representa o número esperado de arestas entre i e j , m é o número total de arestas e δ_{c_i, c_j} é a chamada delta de Kronecker, que é um se os elementos pertencem a mesma comunidade e zero se não pertencem a mesma comunidade. A modularidade é a medida mais usada para separar comunidades em redes (NEWMAN, 2006).

2.5 Componente gigante e robustez

Uma rede tem uma componente gigante se em média cada vértice tem mais de uma ligação. Podemos então perguntar quantos vértices devem ser removidos de uma rede para que ela se fragmente em vários grupos. A importância dessa questão fica clara se for considerada a pergunta para um caso real: “qual fração dos roteadores da internet deve parar de funcionar para que a rede se transforme em grupos de computadores incomunicáveis?”. A Figura 15 mostra uma fragmentação gradual de uma pequena rede seguindo a divisão de vários vértices. Em cada painel foi removido um novo vértice (destacado), juntamente com as suas arestas. A imagem indica que a remoção do primeiro vértice (a) tem um pequeno impacto, a remoção do segundo vértice (b) isola dois pequenos grupos do resto da rede e a remoção do terceiro vértice (c) fragmenta a rede, quebrando-a em cinco grupos incomunicáveis entre si (d). Obviamente, quanto mais vértices forem removidos, maiores são as chances de danificar uma rede.

Figura 15 – O impacto da remoção do vértice.

Fonte: (BARABÁSI, 2016).

A existência de uma componente gigante pode ser obtida pelo critério de Molloy-Reed (MOLLOY; REED, 1995; BARABÁSI, 2016) baseado na razão entre o grau quadrático médio $\langle k^2 \rangle$ e o grau médio $\langle k \rangle$ obtidos da distribuição de grau (binário). O critério de Molloy-Reed para uma rede ter componente gigante é dado por

$$k_o = \frac{\langle k^2 \rangle}{\langle k \rangle} > 2. \quad (2.15)$$

A partir do critério de Molloy-Reed é possível definir a robustez de uma rede por

$$f_c = 1 - \frac{1}{k_o - 1}, \quad (2.16)$$

e que representa a fração crítica de vértices que devem ser removidos para que uma componente gigante seja desfeita (BARABÁSI, 2016). A robustez pode ser pensada como a rigidez ou densidade da rede.

3 CORRELAÇÕES E DISTÂNCIAS

A primeira etapa no estudo de redes é a regra que gera a estrutura e diz como objetos são ou não são conectados, ou seja, como eles se interagem. O conceito básico é o de matriz de adjacência, com elementos um ou zero, representando a ligação ou não-ligação entre dois objetos, respectivamente. Entretanto, alternativas ponderadas da matriz de adjacência podem ser mais adequadas em certos estudos e o conceito de “ligado/não-ligado” é substituído pelo conceito de intensidade da ligação (NEWMAN, 2010; BARABÁSI, 2016).

Para estudos financeiros, como por exemplo, a ligação entre ações de um mercado, uma medida típica da interação é o coeficiente de correlação de Pearson. Por exemplo, no modelo de Markowitz (1952) de carteiras de investimento é a correlação de Pearson que é usada como medida de dependência entre os retornos das aplicações.

Entretanto, a correlação de Pearson é adequada como medida de uma dependência linear. Em particular, no caso de dados cuja distribuição de probabilidade conjunta é uma Gaussiana (distribuição normal) ela é exata. Existem outras medidas de correlação pouco exploradas no estudo de redes e que podem ser interessantes, principalmente em situações em que a distribuição normal não seja o melhor modelo.

Na sequência, apresenta-se o coeficiente de correlação de Pearson e duas medidas alternativas de correlação, os coeficientes de Spearman e de Kendall, que serão explorados para o estudo de redes realizado neste trabalho.

3.1 Correlação

O conceito de correlação já aparece no trabalho de Bravais de 1846 (BRAVAIS, 1846), entretanto, foi Pearson (1896) quem descreveu a fórmula conhecida como “Coeficiente de Pearson” (MARI; KOTZ, 2001). O coeficiente de correlação de Pearson é uma medida de intensidade e direção de uma relação linear entre duas variáveis x e y . O coeficiente de correlação de Pearson é obtido pela expressão (JOHNSON; BHATTACHARYYA, 2009)

$$\rho_P = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (3.1)$$

onde (x_i, y_i) são pares de n observações das variáveis x e y com médias $\bar{x}$ e $\bar{y}$, respectivamente. O valor numérico de ρ varia de -1 a $+1$. Quando ρ estiver próximo de -1 , indica uma correlação negativa forte e, próximo de $+1$, uma correlação positiva forte. Um valor de ρ próximo a 0 indica que há correlação fraca, ou ainda, que não há correlação (JOHNSON; BHATTACHARYYA, 2009).

No caso em que as variáveis x e y são descritas por uma distribuição conjunta normal a correlação nula implica que as variáveis são independentes linearmente (JOHNSON; BHATTACHARYYA, 2009). Entretanto, uma correlação nula (ou muito pequena) pode implicar uma relação não linear entre as variáveis. Além disso, o coeficiente de correlação de Pearson é sen-

sível a presença de *outliers*, isto é, pontos muito distantes da maioria dos pontos de um conjunto de dados.

Uma outra medida de correlação é o coeficiente de correlação de Spearman. Ele é análogo a correlação de Pearson, exceto que nele é substituído os valores das variáveis x e y pelos seus postos, ou seja, pela posição dos dados dispostos na forma ordenada.

O coeficiente de correlação por postos de Spearman é definido por (JOHNSON; BHATTACHARYYA, 2009)

$$\rho_S = \frac{\sum_{i=1}^n (R_i - \frac{n+1}{2})(S_i - \frac{n+1}{2})}{\frac{n(n^2-1)}{12}}, \quad (3.2)$$

onde R_i e S_i são os postos das variáveis e n é o número de observações.

O coeficiente de Spearman também varia de -1 a $+1$. Entretanto, o coeficiente pode indicar uma associação não necessariamente linear entre as variáveis, mas indica uma relação crescente ou decrescente (JOHNSON; BHATTACHARYYA, 2009). Assim, no caso em que os dados seguem uma relação monotônica e não linear, o coeficiente de correlação de Spearman pode ser uma medida mais adequada da relação entre variáveis.

A vantagem principal do método baseado em postos é a redução de erros acidentais, pois em uma curva de frequência normal, os casos excepcionais, também conhecidos como *outliers*, ficam muito mais afastados do que aqueles próximos à média, podendo causar, assim, um efeito grande sobre o resultado da correlação. Outra vantagem, é que a classificação em postos elimina qualquer disparidade entre as duas características comparadas, em relação à sua distribuição, assim, uma série com uma curva de distribuição normal pode ser comparada a outra série com curva de distribuição diferente (SPEARMAN, 1904).

Uma outra alternativa de medida de correlação é o coeficiente de correlação de Kendall. O coeficiente de correlação de Kendall também utiliza os dados ordenados em postos. Embora sejam parecidos, os coeficientes de Spearman e de Kendall apresentam interpretações diferentes. O coeficiente de Spearman é considerado similar ao coeficiente de Pearson em proporção a variância explicada, enquanto Kendall representa uma probabilidade, uma diferença entre a probabilidade de que os dados observados estão na mesma ordem e a probabilidade de que não estejam na mesma ordem (HAUKE; KOSSOWSKI, 2011).

O coeficiente de correlação de Kendall é definido por (BHATTACHARYYA; JOHNSON, 1977)

$$\rho_K = \sum_{\text{pares}(i,j), i < j} \frac{\text{sign}(R_i - R_j) \text{sign}(S_i - S_j)}{\frac{n(n^2-1)}{2}}, \quad (3.3)$$

onde " $\text{sign}(\)$ "significa o sinal da diferença entre parênteses, isto é, $\text{sign}(a - b) = +1$ se $a > b$; ou -1 se $a < b$. Dois pares (X_1, Y_1) e (X_2, Y_2) são concordantes se $(X_1 - X_2)$ e $(Y_1 - Y_2)$ tem o mesmo sinal e discordantes se tiverem sinais contrários (BHATTACHARYYA; JOHNSON, 1977). O coeficiente de correlação de Kendall é um estimador não enviesado da

diferença $P_c - P_d = 2P_c - 1$ entre a proporção P_c de valores concordantes e P_d de valores discordantes.

3.2 Medidas de distância

Este estudo se propõe a elaborar uma matriz de correlação para definir a relação entre os preços de um par de ações e obter a rede do conjunto de ações. Para elaborar a árvore geradora mínima, as arestas precisam ser ponderadas com a distância entre os vértices, assim, é necessário introduzir uma medida de distância entre os vértices.

Há várias maneiras de se calcular a distância entre dois pontos, a mais conhecida é a chamada distância Euclidiana. Considere os pontos $i = (x_1, y_1)$ e $j = (x_2, y_2)$ em um plano como ilustrado na Figura 16. A distância Euclidiana d_{ij} entre os pontos i até j é, de acordo com o teorema de Pitágoras, dada por (JOHNSON; WICHERN, 2002)

$$d_{ij} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}. \quad (3.4)$$

Figura 16 – Distância Euclidiana entre os pontos i e j de um plano.

Fonte: Elaborada pela autora.

Para um espaço de dimensão p , a distância euclidiana entre os pontos i e j com coordenadas $i = (x_1, x_2, x_3, \dots, x_p)$ e $j = (y_1, y_2, y_3, \dots, y_p)$ é a generalização da Equação 3.4 dada por

$$d_{ij} = \sqrt{\sum_{i=1}^p (x_i - y_i)^2}. \quad (3.5)$$

A distância Euclidiana satisfaz os seguintes requisitos:

$$d_{ij} = 0, \quad \text{se } i = j, \quad (3.6)$$

$$d_{ij} = d_{ji}, \quad (3.7)$$

$$d_{ij} \leq d_{ik} + d_{kj}. \quad (3.8)$$

As propriedades 3.6, 3.7 e 3.8 são usadas para definir uma medida de distância de modo geral. Assim, qualquer função que relaciona dois pontos satisfazendo as propriedades 3.6, 3.7 e 3.8 é

uma medida de distância entre pontos. O coeficiente de correlação é um parâmetro muito usado para indicar a relação entre variáveis. A partir da correlação é possível definir uma medida de distância (GOWER, 1966; JOHNSON; WICHERN, 2002) pela expressão

$$d_{ij} = \sqrt{2(1 - \rho_{ij})}. \quad (3.9)$$

onde ρ_{ij} é o coeficiente de correlação de Pearson, Spearman ou Kendall. A Equação 3.9 tem valor máximo igual a 2, indicando máxima distância entre variáveis quando estas são maximamente anticorrelacionadas ($\rho = -1$) e o valor mínimo igual a 0 quando as variáveis são maximamente correlacionadas ($\rho_p = +1$).

A distância a partir de correlações, conforme equação 3.9 tem sido utilizada para elaborar redes e MST em mercados financeiros (MANTEGNA, 1999; BONANNO; LILLO; MANTEGNA, 2001b; BONANNO; LILLO; MANTEGNA, 2001a; VANDEWALLE et al., 2001; ONNELA et al., 2003; ONNELA; KASKI; KERTÉSZ, 2004; BONANNO et al., 2004; JUNG et al., 2006; COELHO et al., 2007; TUMMINELLO et al., 2007a; TUMMINELLO et al., 2007b; TABAK; SERRA; CAJUEIRO, 2010; DJAUHARI; GAN, 2015; DJAUHARI; GAN, 2016; DEVIREN; DEVIREN, 2016; HUANG et al., 2017).

Neste trabalho será utilizada uma medida de distância baseada no valor absoluto da correlação e definida como

$$d_{ij} = 1 - |\rho_{ij}| \quad (3.10)$$

e varia de 0 até 1. A justificativa para o uso da Equação 3.10 é que em uma rede a ligação é medida por uma intensidade. Assim, um alto valor de correlação, mesmo sendo negativo, indica uma forte relação entre variáveis. No caso de Pearson, uma forte relação linear. Do ponto de vista de uma medida de dependência entre variáveis, esta deve ser um número real não negativo (MALEVERGNE; SORNETTE, 2006).

4 METODOLOGIA

Nesta seção apresenta-se a base de dados estudada, bem como as técnicas de análises utilizadas na pesquisa.

4.1 Comentários sobre o período de análise

Além de comparar redes obtidas com diferentes tipos de correlações, este trabalho estudou o efeito de evento sobre a estrutura de redes do mercado financeiro. A escolha do evento para estudo foi motivada pela instabilidade política e econômica do Brasil nos últimos anos. Essa instabilidade política, e em particular algumas notícias disseminadas, produzem efeitos sobre o mercado de ações. O evento corresponde a notícia publicada pelo jornal O Globo na noite de 17 de maio de 2017 dizendo que o dono da empresa brasileira JBS, Joesley Batista, gravou o então Presidente da República Michel Temer autorizando a compra do silêncio do Deputado Federal, com mandato cassado em 2016 (CALGARO; RAMALHO; GARCIA, 2016), Eduardo Cunha (JARDIM, 2017). A reportagem diz, ainda, que a gravação também inclui o ex candidato a presidência da república, Aécio Neves, pedindo dinheiro ao empresário Joesley Batista. A notícia é resultado das investigações da chamada da Operação “Lava Jato”. A operação “Lava Jato” é a maior investigação de corrupção e lavagem de dinheiro que o Brasil já teve. O nome atribuído à investigação decorre do uso de uma rede de postos de combustíveis e lava a jato de automóveis para movimentar recursos ilícitos pertencentes a uma das organizações criminosas inicialmente investigadas e, embora a investigação tenha avançado para outras organizações criminosas, o nome inicial foi mantido. A divulgação desta notícia, caiu como uma “bomba” sobre o país, se espalhando rapidamente pelos noticiários, internet e redes sociais. Por consequência, o mercado também reagiu. No dia 18 de maio de 2017 a Bovespa teve uma considerável queda e o dólar fechou em alta (GODOY; SILVA, 2017; TREVIZAN, 2017; ADVFN, 2017). A Figura 17 mostra a variação do Ibovespa durante o mês de maio de 2017, onde observa-se que há uma variação anormal no dia 18, ou seja, no pregão imediatamente seguinte à divulgação da notícia.

Para o estudo do efeito do evento sobre a estrutura do mercado foram coletados o retorno das ações de cinco pregões anteriores e cinco posteriores ao dia 18 de maio de 2017. No dia 18 ocorreu a maior queda do preço das ações, portanto, este foi fixado como o dia do evento, conforme indicado na Figura 18. Foram elaboradas as redes completas e as árvores geradoras mínimas correspondentes às ações de empresas do Ibovespa com interligações ponderadas por correlações de Pearson, Spearman e Kendall do retorno do preço de suas ações. As redes foram estudadas em duas escalas de tempo:

- (1) redes diárias durante 5 dias anteriores ao evento (de 11 a 17 de maio) e 5 dias posteriores ao evento (de 19 a 25 de maio);
- (2) agrupando os dados dos 5 dias antes e dos 5 dias depois do evento. Assim, foi possível estudar o efeito médio contido nas correlações dos dados agrupados antes e depois do

Figura 17 – Variação do Ibovespa em maio de 2017.

Fonte: Elaboração própria baseados em dados da BM&FBOVESPA (2017).

evento.

Figura 18 – Período de coleta e análise dos dados.

Fonte: Elaborada pela autora.

A Figura 19 mostra com mais detalhes a variação do Ibovespa no período de análise.

4.2 Base de dados

A base de dados da pesquisa são as ações de empresas que pertencem à carteira teórica do Índice Bovespa (Ibovespa) da BM&FBovespa. As carteiras teóricas agrupam ações de empresas que possibilitam o investimento teórico representado por um índice. O índice Ibovespa é o mais importante para representar o desempenho médio das cotações dos principais ativos (ações) do mercado de ações brasileiro.

O Ibovespa indica as variações nos preços dos ativos integrantes, bem como o impacto da distribuição de proventos (dividendos ou juros sobre capital próprio) dessas companhias. O Ibovespa é composto pelas ações e *units* (ativos compostos por um conjunto de ações ordinárias e preferenciais) exclusivamente de ações de companhias listadas na BM&FBOVESPA, exceto Certificados de Depósitos de Ações (BDRs - *Brazilian Depositary Receipts*¹) e ativos de com-

¹ São valores mobiliários, emitidos no Brasil por instituições depositárias, que possuem como lastro valores mobiliários emitidos no Exterior, como por exemplo, ações de companhias estrangeiras.

Figura 19 – Variação do Ibovespa no período de 11/05/2017 a 25/05/2017. Os números indicam a variação percentual em relação ao dia anterior.

Fonte: Elaboração própria baseados em dados da BM&FBOVESPA (2017).

panhias em recuperação judicial ou extrajudicial, regime especial de administração temporária, intervenção ou que sejam negociados em qualquer outra situação especial de listagem.

As ações que integram o Ibovespa devem seguir alguns critérios, dispostos em sua metodologia, como, presença em pregão² de 95%, participação maior ou igual a 0,1% no mercado à vista, em termos de volume financeiro; não ter ação com cotação abaixo de R\$1,00 (as chamadas *Penny Stocks*). As companhias que atendam a esses critérios são ordenadas conforme o Índice de Negociabilidade (IN), que considera a quantidade de negociações e o volume financeiro gerado, sendo incluídas, assim, as ações que representem, de forma cumulativa, 85% das negociações ocorridas no período.

O Ibovespa, no período de maio a agosto de 2017, era composto por 58 ações, representando 13,39% do total de ações da BM&FBOVESPA, excluindo os BDRs. O Quadro 1 mostra todos os ativos quem compõem o Ibovespa durante o período de análise desta pesquisa.

Os dados foram coletados utilizando um código em R para baixar dados em alta frequência das ações da Bovespa a partir do pacote GetHFDData (PERLIN; RAMOS, 2017). Foram coletados preço das ações cotados a cada 15 minutos, durante o período de 11 a 25 de maio de 2017, totalizando 21.054 observações.

4.3 A estrutura de rede

Os dados utilizados para a elaborar as matrizes de correlação foram os logaritmos da variação dos preços cotados a cada quinze minutos das 58 ações das empresas selecionadas no índice Bovespa. Seja $P(t)$ e $P(t - \tau)$ os preço de um ativo no instante t e $t - \tau$. O retorno $R(t)$

² Pregão é o nome dado ao dia de trabalho na BM&FBovespa, no qual há pessoas fazendo ofertas de compra e venda.

Quadro 1 – Ações da carteira teórica Ibovespa – maio a agosto de 2017, atualizada em 28 de agosto de 2017

Código	Ação	Setor	Tipo*	Part. (%)
ABEV3	AMBEV S/A	BEBIDAS	ON	7,337
BBAS3	BRASIL	BANCO	ON NM	3,356
BBDC3	BRADESCO	BANCO	ON N1	1,636
BBDC4	BRADESCO	BANCO	PN N1	8,519
BBSE3	BBSEGURIDADE	PREVIDÊNCIA E SEGUROS	ON NM	1,575
BRAP4	BRADESPAR	MINERAÇÃO	PN N1	0,492
BRFS3	BRF SA	ALIMENTOS PROCESSADOS CARNES E DERIVADOS	ON NM	2,785
BRKM5	BRASKEM	PETROQUÍMICOS	PNA N1	0,842
BRML3	BR MALLS PAR	EXPLORAÇÃO DE IMÓVEIS	ON NM	0,963
BVMF3	BMFBOVESPA	SERVIÇOS FINANCEIROS	ON NM	3,884
CCRO3	CCR SA	EXPLORAÇÃO DE RODOVIAS	ON NM	1,688
CIEL3	CIELO	SERVIÇOS FINANCEIROS	ON NM	2,135
CMIG4	CEMIG	ENERGIA ELÉTRICA	PN N1	0,609
CPFE3	CPFL ENERGIA	ENERGIA ELÉTRICA	ON NM	1,070
CPLE6	COPEL	ENERGIA ELÉTRICA	PNB N1	0,256
CSAN3	COSAN	PETRÓLEO E GÁS	ON NM	0,504
CSNA3	SID NACIONAL	SIDERURGIA	ON	0,440
CYRE3	CYRELA REALT	CONSTRUÇÃO CIVIL	ON NM	0,264
ECOR3	ECORODOVIAS	EXPLORAÇÃO DE RODOVIAS	ON NM	0,176
EGIE3	ENGIE BRASIL	ENERGIA ELÉTRICA	ON NM	0,613
ELET3	ELETROBRAS	ENERGIA ELÉTRICA	ON N1	0,441
ELET6	ELETROBRAS	ENERGIA ELÉTRICA	PNB N1	0,428
EMBR3	EMBRAER	MATERIAL DE TRANSPORTE	ON NM	1,071
ENBR3	ENERGIAS BR	ENERGIA ELÉTRICA	ON NM	0,299
EQTL3	EQUATORIAL	ENERGIA ELÉTRICA	ON NM	1,064
ESTC3	ESTACIO PART	SERVIÇOS EDUCACIONAIS	ON NM	0,660
FIBR3	FIBRIA	PAPEL E CELULOSE	ON NM	0,808
GGBR4	GERDAU	SIDERURGIA	PN N1	0,841
GOAU4	GERDAU MET	SIDERURGIA	PN N1	0,288
HYPE3	HYPERMARCAS	COMÉRCIO MEDICAMENTOS	ON NM	1,005
ITSA4	ITAUSA	BANCO	PN N1	3,404
ITUB4	ITAUNIBANCO	BANCO	PN N1	11,086
JBSS3	JBS	ALIMENTOS PROCESSADOS CARNES E DERIVADOS	ON NM	1,180
KLBN11	KLABIN S/A	PAPEL E CELULOSE	UNT N2	0,674
KROT3	KROTON	SERVIÇOS EDUCACIONAIS	ON NM	2,150

continuação

Código	Ação	Setor	Tipo*	Part. (%)
LAME4	LOJAS AMERIC	COMÉRCIO PRODUTOS DIVERSOS	PN	1,017
LREN3	LOJAS RENNER	COMÉRCIO - VESTUÁRIO	ON NM	1,858
MRFG3	MARFRIG	ALIMENTOS PROCESSADOS CARNES E DERIVADOS	ON NM	0,240
MRVE3	MRV	CONSTRUÇÃO CIVIL	ON NM	0,314
MULT3	MULTIPLAN	EXPLORAÇÃO DE IMÓVEIS	ON N2	0,503
NATU3	NATURA	COMÉRCIO PRODUTOS DE USO PESSOAL	ON NM	0,429
PCAR4	P.ACUCAR-CBD	COMÉRCIO DE ALIMENTOS	PN N1	0,971
PETR3	PETROBRAS	PETRÓLEO E GÁS	ON	3,356
PETR4	PETROBRAS	PETRÓLEO E GÁS	PN	4,898
QUAL3	QUALICORP	SERVIÇOS MÉDICO-HOSPITALARES	ON NM	0,723
RADL3	RAIADROGASIL	COMÉRCIO DE MEDICAMENTOS	ON NM	1,200
RAIL3	RUMO S.A.	TRANSPORTE FERROVIÁRIO	ON NM	0,664
RENT3	LOCALIZA	ALUGUEL DE CARRO	ON NM	0,788
SANB11	SANTANDER BR	BANCO	UNT	0,919
SBSP3	SABESP	ÁGUA E SANEAMENTO	ON NM	0,902
SMLE3	SMILES	PROGRAMA DE FIDELIZAÇÃO	ON NM	0,322
SUZB5	SUZANO PAPEL	PAPEL E CELULOSE	PNA N1	0,590
TIMP3	TIM PART S/A	TELECOMUNICAÇÕES	ON NM	0,791
UGPA3	ULTRAPAR	PETRÓLEO E GÁS	ON NM	2,455
USIM5	USIMINAS	SIDERURGIA	PNA N1	0,273
VALE3	VALE	MINERAÇÃO	ON N1	9,513
VIVT4	TELEF BRASIL	TELECOMUNICAÇÕES	PN	1,731
WEGE3	WEG	INDÚSTRIA - MOTORES	ON NM	1,000
Total				100,000
Total de ações: 58				

*Ordinária Nominativa (ON): dá direito a voto em assembleias; Preferencial Nominativa (PN): tem preferência na distribuição de dividendos, pode ser classificados em A, B, C; Segmento de Listagem Novo Mercado (NM); Segmento de Listagem Nível 1 de Governança Corporativa (N1); Segmento de Listagem Nível 2 de Governança Corporativa (N2); Certificado de Depósito de Ações (UNT): ativos compostos por um conjunto de ações ordinárias e preferenciais

Fonte: (BM&FBOVESPA, 2017)

é definido por (MANTEGNA, 1999)

$$R(t) = \ln P(t) - \ln P(t - \tau) = \ln \left(\frac{P(t)}{P(t - \tau)} \right). \quad (4.1)$$

Com os valores empíricos dos retornos de cada ativo obtém-se a matriz de correlação, que tem por elementos as correlações entre pares de ações de empresas i e j . A matriz de correlação é simétrica e com elementos 1 na diagonal. Para a obtenção da rede foi definida a matriz adjacência ponderada por correlações que tem por elementos o valor absoluto das correlações e atribuindo-se zero aos elementos da diagonal. A matriz de distâncias para a elaboração da árvore mínima geradora foi obtida a partir da medida de distância definida pela Equação 3.10. Para o estudo da árvore geradora mínima definimos a matriz de adjacência ponderada por correlações entre os elementos da árvore. Foi realizado um estudo das distribuições de grau ponderado e as respectivas métricas de rede.

Os cálculos estatísticos e visualizações foram elaborados com o programa estatístico R-SOFTWARE e o programa *software* Gephi versão 0.9.1 que é específico para visualização e análise de redes e elabora as árvores a conforme o algoritmo de Kruskal.

5 RESULTADOS

Este capítulo apresenta os resultados obtidos para as redes elaboradas a partir dos coeficientes de correlação de Pearson, Spearman e Kendall e que a partir deste momento serão chamadas simplesmente de rede de Pearson, rede de Spearman e rede de Kendall; e árvore geradora mínima de Pearson, árvore geradora mínima de Spearman e árvore geradora mínima de Kendall, respectivamente.

Inicialmente, apresenta-se o comportamento da distribuição do valor absoluto das correlações entre as ações ao longo do tempo. Para isso foi utilizado o chamado diagrama de caixa (*box-plot* em inglês). O *box-plot* é um diagrama que contém, para um conjunto de dados ordenados, o valor mínimo, o 1º quartil, a mediana (2º quartil), o 3º quartil e o valor máximo. Ele contém a divisão dos valores dos dados ordenados em 4 partes e cada uma destas partes corresponde a um acúmulo de 25% dos valores. Na caixa central do *box-plot*, entre o 1º e 3º quartil, corresponde a 50% dos valores centrais dos dados. A distância entre o 1º e 3º quartil é uma medida da variabilidade dos dados, chamada intervalo interquartil. A simetria do *box-plot* indica a simetria da distribuição de frequências da variável considerada (JOHNSON; BHATTACHARYYA, 2009).

No *box-plot*, as barras verticais que estão ligadas a caixa central indicam os valores máximo e mínimo das observações, excluindo-se os *outliers*. Os *outliers* são indicados pelos pontos acima da barra e aparecem quando estão abaixo ou acima do limite de detecção de *outliers*, que é construído utilizando o intervalo interquartil, dado pela distância entre o primeiro e o terceiro quartil. Sendo assim, os limites inferior e superior de detecção de *outliers* são dados por: Limite Inferior = Primeiro Quartil – (1,5 * (Terceiro Quartil – Primeiro Quartil)); Limite Superior = Terceiro Quartil + (1,5 * (Terceiro Quartil – Primeiro Quartil)).

A Figura 20 mostra o *box-plot* para o módulo dos coeficientes de correlação de Pearson, Spearman e Kendall ao longo do tempo. Podemos observar que nos dois dias posteriores a notícia houve um deslocamento da distribuição, com um aumento do valor médio (mediana) das correlações, mais evidentes para as correlações de Pearson e Spearman. A distribuição de correlações de Kendall é a mais estável ao longo do tempo. A correlação de Pearson é a que, no dia da notícia, sofre maior variação em relação aos dias anteriores e posteriores.

Nas seções seguintes serão apresentados os resultados para as redes obtidas para as duas escalas de tempo consideradas, iniciando com as redes obtidas diariamente.

5.1 Rede diárias: a rede completa

A rede completa é obtida com a ligação de todas as ações via valor absoluto das correlações. Ela é completa pois é improvável que os dados empíricos apresentem uma correlação que seja exatamente igual a zero. Assim, as redes são totalmente conectadas. Não apresentaremos o desenho dessas redes pois eles correspondem a um conjunto de pontos com arestas ligando todos eles e são visualmente confusas. Se considerarmos ligações binárias (0 ou 1) entre os vér-

Figura 20 – *Boxplot* do módulo dos coeficientes de correlação.

Fonte: Elaborada pela autora.

tices verifica-se que todos tem o mesmo grau $n - 1$ e as diferenças entre as correlações usadas pouco seriam evidenciadas. Daí a importância do estudo das estatísticas da rede considerando os graus ponderados por correlação. Também serão apresentadas apenas algumas estatísticas básicas para as redes completas.

A Figura 21(a),(b) e (c) mostra a distribuição acumulada complementar do grau ponderado das redes completas em escala log-log para os dias anteriores ao evento, o dia do evento e os dias posteriores ao evento. A distribuição complementar é dada por $1 - P(X \leq x)$, em que $P(X \leq x)$ é a probabilidade acumulada até $X = x$. Para os dias anteriores ao evento a distribuição do grau ponderado das redes de Pearson e Spearman tendem se aproximar no quinto e no primeiro dia anterior ao evento, enquanto no quarto dia, as distribuições estão mais distantes. A distribuição do grau ponderado da rede elaborada com a correlação de Kendall permanece mais distante das demais, refletindo o valor mais baixo do coeficiente desta correlação, seguido por Spearman e depois Pearson, sendo este o coeficiente mais alto em todas as redes. Para os dias posteriores ao evento, a distribuição do grau ponderado das redes de Pearson e Spearman tendem se aproximar no primeiro, terceiro e quinto dia após o evento.

Para evidenciar as diferenças entre as distribuições, será apresentado o comportamento ao longo do tempo de algumas estatísticas básicas associadas as distribuições de grau das redes. A Figura 22(a) mostra o grau médio ponderado ao longo do tempo para o período analisado. Os graus médios ponderados apresentam o mesmo comportamento ao longo tempo para as redes. As redes de Pearson tem sempre um valor maior para o grau médio ponderado, seguido pelas redes Spearman e Kendall. Observa-se uma variação maior nos dois primeiros dias após o evento. A partir do terceiro dia tem-se indicação que o mercado está voltando ao normal, ou seja, parecido com o momento anterior ao evento. Se analisar os períodos anteriores e posteriores como um todo, observa-se uma tendência crescente de grau médio das redes.

Figura 21 – Distribuição acumulada complementar do grau ponderado das redes completas diárias: (a) cinco dias antes do evento, (b) o dia do evento e (c) cinco dias após o evento

Fonte: Elaborada pela autora.

A Figura 22(b) mostra a dispersão da distribuição de grau ponderado, medido pelo desvio-padrão (σ) dos graus ponderados, da rede completa ao longo do tempo. A rede de Pearson apresenta novamente a maior dispersão na distribuição de grau, seguida pelas redes Spearman e Kendall, respectivamente. Novamente observa-se uma tendência crescente na dispersão dos graus ponderados do período como um todo. Para complementar, é interessante mostrar a variação da dispersão em relação à média, medido pelo coeficiente de variação. O coeficiente de variação é a razão do desvio padrão pelo grau médio, $CV = \sigma / \langle k \rangle$. A Figura

22(c) mostra o coeficiente de variação de grau ponderado da rede completa ao longo do tempo. Algumas diferenças entre as redes ficam mais evidentes ao observar os 3 dias anteriores e os 3 dias posteriores, onde o coeficiente de variação da rede de Pearson pouco se altera quando comparado com as redes de Spearman e Kendall.

Figura 22 – (a) Grau ponderado médio, (b) desvio-padrão de grau ponderado e (c) coeficiente de variação diário da rede completa durante todo o período analisado.

Fonte: Elaborada pela autora.

5.2 Redes diárias: Árvore geradora mínima

Nesta seção são apresentadas as árvores geradoras mínimas diárias para as correlações de Pearson, Spearman e Kendall. A árvore geradora mínima parece ser mais interessante para um estudo detalhado das métricas de rede, pois contém as ligações selecionadas em ordem de importância, segundo uma medida de distância, e assim, deve conter menos informações redundantes quando comparada à rede completa. Como nas escalas de tempo diárias temos muitas estruturas (desenhos) das árvores geradoras mínimas, estas foram apresentadas no apêndice A, onde as ações foram separadas em grupos por modularidade, com ações de mesma cor pertencente ao mesmo grupo, o tamanho do vértice é proporcional ao seu grau ponderado, e as arestas tem espessura proporcional ao valor absoluto das correlações. Nesta seção, são apresentadas as distribuições dos graus ponderados por correlações e as várias métricas de rede associadas a árvore mínima geradora ao longo do tempo.

As Figuras 23(a), (b) e (c) mostram a distribuição acumulada complementar do grau ponderado das árvores geradoras mínimas em escala log-log para os dias anteriores ao evento, o dia do evento e os dias posteriores ao evento, respectivamente. A distribuição do grau ponderado das redes de Kendall, Pearson e Spearman estão próximas em todos os dias analisados. A distribuição da rede de Kendall no dia do evento sugere uma região de comportamento linear.

A Figura 24(a) mostra o grau ponderado médio da árvore geradora mínima ao longo do tempo. Em geral, há um aumento no dia do evento se mantendo praticamente estável nos dois dias seguintes ao evento para as árvores elaboradas a partir da correlação de Spearman e Kendall, seguida de uma queda no terceiro dia. Na rede de Pearson, a alta do grau ponderado médio é maior no dia do evento, e depois queda até três dias após o evento. Ao longo do tempo observa-se uma tendência crescente para o grau ponderado médio.

A Figura 24(b) mostra que o desvio-padrão dos graus ponderados da árvore geradora mínima tendem a seguir um comportamento parecido para as diferentes correlações utilizadas na análise. Na árvore de Pearson, o desvio padrão tem aumento expressivo em relação às demais correlações no dia do evento. A árvore de Kendall se comporta de maneira oposta às demais no terceiro dia antes do evento e, a árvore de Spearman, tem aumento no desvio-padrão no último dia do período de análise, enquanto os outros têm queda. A tendência ao longo do período é de aumento na dispersão de graus ponderados.

A Figura 24(c) mostra o coeficiente de variação de grau ponderado da árvore geradora mínima. Na rede de Pearson, o coeficiente de variação tem um maior aumento no quarto dia antes e no dia do evento e tende a se comportar de maneira oposta às redes dos demais coeficientes analisados. A tendência ao longo do período é crescente.

Na Figura 25 (a) mostra o comprimento do caminho médio em função do tempo. O comprimento do caminho médio representa a média da distância entre todos os pares de vértices na rede. O comprimento do caminho médio da árvore geradora mínima tem oscilação durante todo o período analisado e com um comportamento diferente para todas as redes. Entretanto, ao analisar o comportamento global no período, observa-se uma tendência decrescente para o

Figura 23 – Distribuição acumulada complementar do grau ponderado das árvores geradoras mínimas (a) para os cinco dias antes do evento, (b) para o dia do evento e, (c) para os cinco dias após o evento.

Fonte: Elaborada pela autora.

comprimento de caminho médio. O mesmo vale para o diâmetro da árvore geradora mínima, apresentado na Figura 25(b), que apresenta variação durante todo o período analisado e com um comportamento diferente para todas as redes. O comportamento é parecido com o comprimento do caminho médio, pois se a média entre os pares de vértices reduz, a distância máxima entre

dois vértices também tende a reduzir. A tendência geral no período é decrescente.

5.3 Rede completa para dados agrupados antes e depois do evento

O estudo do comportamento das redes diárias foi útil para indicar que os parâmetros das redes apresentam uma certa tendência ao longo do tempo. Estes resultados sugeriram que

Figura 24 – (a) Grau ponderado médio, (b) desvio-padrão de grau ponderado e (c) coeficiente de variação de grau ponderado da árvore geradora mínima durante todo o período analisado.

Fonte: Elaborada pela autora.

Figura 25 – (a) Comprimento do caminho médio e (b) diâmetro da árvore geradora mínima durante todo o período analisado.

Fonte: Elaborada pela autora.

o estudo do efeito médio das correlações para os dados agrupados antes e depois do evento poderia revelar mais claramente as alterações estruturais das redes antes e depois do evento. A seguir, são apresentados os resultados para as redes obtidas para os dados agrupados de cinco dias antes e cinco dias depois do evento. Os termos “antes” e “depois” são usados para se referir aos períodos. Apresentamos alguns resultados para a rede completa seguido de um estudo mais detalhado das árvores geradoras mínimas. Novamente, para a rede completa não apresentaremos os desenhos das redes por motivos explicados anteriormente.

A Figura 26 mostra a distribuição acumulada complementar do grau ponderado da rede completa antes e depois do evento.

Para discernir diferenças entre os diferentes tipos de correlação apresentaremos algumas estatísticas básicas da rede completa. A Figura 27(a) mostra o grau médio ponderado das redes completas no período anterior e posterior ao evento analisado. Observa-se que há um aumento

Figura 26 – Distribuição acumulada complementar do grau ponderado das redes completas para o período agrupado de cinco dias antes e cinco dias depois do evento.

Fonte: Elaborada pela autora.

do grau médio para todas as redes, com Kendall e Spearman variando praticamente na mesma proporção, enquanto Pearson tem uma variação maior. O grau médio reflete o valor dos coeficientes que ponderam as ligações, assim, a Figura 27(a) mostra que após o evento houve um aumento da intensidade média das ligações na rede, com uma variação maior para a rede de Pearson.

A Figura 27(b) mostra desvio-padrão do grau médio ponderado para as redes completas antes e depois do evento. Todas as redes tiveram um aumento no desvio-padrão de grau ponderado, indicando uma maior dispersão dos graus após o evento. A variação no desvio-padrão é maior para a rede de Pearson.

A Figura 27(c) mostra que o coeficiente de variação dos graus ponderados antes e depois do evento. Enquanto coeficiente de variação das redes de Pearson sofre um aumento, para as redes de Kendall e Spearman ocorre uma redução.

5.4 Árvore geradora mínima para dados agrupados antes e depois do evento

A seguir são apresentados os resultados para as árvores geradoras mínimas para os dados agrupados antes e depois do evento. As Figuras 28, 29 e 30 mostram as árvores geradoras mínimas obtidas para as correlações de Pearson, Spearman e Kendall, respectivamente, para os dados agrupados (a) antes e (b) depois do evento. O tamanho dos vértices é proporcional ao seu grau ponderado, os grupos foram separados por modularidade e correspondem aos vértices de mesma cor com o nome da ação correspondente e as arestas são proporcionais ao valor absoluto das correlações.

Se vasculharmos os desenhos das redes vamos observar mudanças na estrutura depois do evento e também mudanças nos agrupamentos para cada tipo de correlação. Para se ter uma ideia das posições ocupadas pelas ações antes e depois do evento de acordo com o tipo de

Figura 27 – (a) Grau médio ponderado, (b) Desvio-padrão de grau ponderado e (c) Coeficiente de variação de grau ponderado da rede completa para os períodos agrupados antes e depois do evento.

Fonte: Elaborada pela autora.

correlação são apresentados no Quadro 2, o número de grupos formados, as ações com maior grau em cada grupo (de acordo com a modularidade) para os períodos antes e depois do evento. O número de grupos, após o evento, ficou menor para todas as correlações. As células com as letras *AD* indicam as ações que permaneceram nas redes antes e depois do evento para cada correlação como a principal de um grupo (maior grau ponderado). As ações com um asterisco

Figura 28 – Árvore geradora mínima para correlação de Pearson (a) antes do evento e, (b) depois do evento.

Fonte: Elaborada pela autora.

Figura 29 – Árvore geradora mínima para correlação de Spearman (a) antes do evento e, (b) depois do evento.

Fonte: Elaborada pela autora.

Figura 30 – Árvore geradora mínima para correlação de Kendall (a) antes do evento e, (b) depois do evento.

Fonte: Elaborada pela autora.

(*) são aquelas que aparecem em todas as redes antes do evento, como a principal de um grupo, e as com dois asteriscos (**) são as que aparecem em todas as redes após o evento como a principal de um grupo. O número de grupos foi sempre maior nas correlações de Spearman e Kendall.

Quadro 2 – Ações com maior grau ponderado em cada grupo (de acordo com a modularidade) antes e depois do evento. As células com as letras *AD* indicam as ações que permaneceram nas redes antes e depois do evento para cada correlação, como a principal de um grupo. As ações com um asterisco (*) são aquelas que aparecem em todas as redes antes do evento como a principal de um grupo (maior grau ponderado). As ações com dois asteriscos (**) são as que aparecem em todas as redes após o evento como a principal de um grupo

Correlação	Período	N. grupos	Grupo							
			0	1	2	3	4	5	6	7
Pearson	Antes	7	VALE3	BBDC3	LAME4	KROT3	*PETR4 ^{AD}	CSAN3	USIM5	
	Depois	6	**BBAS3	**BBDC3	**CSNA3	**CMIG4	PETR4 ^{AD}	ITSA4		
Spearman	Antes	8	BBDC3 ^{AD}	VALE3	CSNA3 ^{AD}	EGIE3	*PETR4	JBSS3	ESTC3	UGPA3
	Depois	7	**BBDC3 ^{AD}	**CSNA3 ^{AD}	ELET3	SUZB5	**CMIG4	**BBAS3	VIVT4	
Kendall	Antes	9	BBDC3 ^{AD}	USIM5	BRKM5 ^{AD}	EGIE3	*PETR4 ^{AD}	FIBR3	ITUB4	ESTC3
	Depois	7	**BBAS3	**BBDC3 ^{AD}	**CSNA3	BRKM5 ^{AD}	**CMIG4	ELET3	PETR4 ^{AD}	UGPA3

Fonte: Elaborada pela autora.

Na sequência apresentamos algumas métricas da rede associadas a distribuição de grau ponderado. A Figura 31 mostra a distribuição acumulada complementar do grau ponderado das árvores geradoras mínimas. Agora temos algo interessante, após o evento temos um decaimento com tendência linear das caudas de distribuição. Isto indica que após o evento houve um au-

mento no número de correlações mais fortes e que estas foram privilegiadas na árvore mínima. O comportamento no final das caudas da distribuição com uma queda mais acentuada pode ser um efeito devido ao tamanho finito da árvore geradora. Para ter mais evidência do comportamento precisaríamos obter uma árvore geradora mínima com mais vértices, ou seja, uma carteira com um número maior de ações. O comportamento descrito após o evento indica que a estrutura de rede ficou mais conectada. Mas, para verificar esse comportamento, apresenta-se na sequência algumas métricas das redes.

Figura 31 – Distribuição acumulada complementar de grau ponderado das árvores geradoras mínimas para o período agrupado dos cinco dias antes e cinco dias depois do evento.

Fonte: Elaborada pela autora.

A Figura 32(a) mostra o grau ponderado médio das redes antes e depois do evento. Observa-se um aumento significativo do grau médio em todos os tipos de redes, com as redes de Pearson apresentando um grau médio maior, seguido pelas redes de Spearman e Kendall. A variação do grau é praticamente na mesma proporção para todas as redes. O aumento no grau ponderado médio da árvore geradora mínima indica que houve aumento do número de correlações com valores mais altos após o evento.

A Figura 32(b) mostra o desvio-padrão da distribuição de grau das árvores geradoras mínimas elaboradas no período antes e depois do evento. A dispersão de grau ponderado aumentou após o evento para todas as correlações sendo que a maior variação foi para a rede de Spearman, seguida por Kendall e Pearson. A variação percentual em relação a média é mostrada na Figura 32(c) para o coeficiente de variação dos graus ponderados da árvore geradora mínima. Observa-se que antes do evento os coeficientes de variação das redes de Pearson, Spearman e Kendall são bem próximos, enquanto que, após o evento, sofrem um aumento significativo, ficando com valores bem diferentes para cada tipo de correlação. A variação maior ocorre nas árvores mínimas de Spearman, seguidas por Kendall e Pearson, respectivamente.

A Figura 33(a) mostra o comprimento do caminho médio da árvore geradora mínima para o período agrupado em antes e depois do evento. Observa-se que o comprimento do caminho médio da árvore geradora mínima teve uma redução no período após o evento para

Figura 32 – (a) Grau ponderado médio (b) desvio-padrão de grau ponderado e (c) coeficiente de variação de grau ponderado da árvore geradora mínima para os períodos agrupados antes e depois do evento.

Fonte: Elaborada pela autora.

todas as redes analisadas, indicando que os vértices ficaram mais próximos. Como mostra Figura 33(b), a redução do caminho médio também é acompanhada pela redução do diâmetro da árvore geradora mínima pois uma redução na média entre os pares de vértices após evento deve também indicar uma redução da distância máxima entre dois vértices. As variações desses

parâmetros relativamente ao evento é maior para as redes de Spearman, seguidas pelas redes de Pearson e Kendall, respectivamente.

Figura 33 – (a) Comprimento do caminho médio e (b) diâmetro da árvore geradora mínima para o período agrupado em antes e depois do evento.

Fonte: Elaborada pela autora.

Além disso, procuramos também explorar a relação entre o grau ponderado e o grau simples (binário) das redes. Note que o grau simples médio $\langle k \rangle_{\text{simples}} = 1,9655$ é o mesmo para todas as redes. A Figura 34 mostra a dispersão do Grau ponderado $\times$ Grau simples para as ações que compõem a árvore geradora mínima de (a) Pearson, (b) Spearman e (c) Kendall para os dados agrupados antes e depois do evento. A figura, mostra também, a reta de regressão ajustada de modo que o grau ponderado nulo corresponda ao grau simples igual a zero. Verifica-se um bom ajuste linear entre o grau ponderado e o grau simples. A equação dá uma indicação do grau médio ponderado em cada grau simples. A razão entre grau ponderado e grau simples sempre aumenta após o evento. O Quadro 3 mostra as ações com maior grau ponderado em cada grau simples. Novamente, as células com as letras *AD* indicam as ações que permaneceram nas redes antes e depois do evento para cada correlação como a principal de um grupo, as ações com um asterisco (*) são aquelas que aparecem em todas as redes antes do evento como a principal

de um grupo e as com dois asteriscos (**) são as que aparecem em todas as redes após o evento como a principal de um grupo.

Quadro 3 – Ações com maior grau ponderado em cada grau simples antes e depois do evento. As células com as letras *AD* indicam as ações que permaneceram nas redes antes e depois do evento para cada correlação como a principal de um grupo. As ações com um asterisco (*) são aquelas que aparecem em todas as redes antes do evento como a principal de um grupo. As ações com dois asteriscos (**) são as que aparecem em todas as redes após o evento como a principal de um grupo

Correlação	Período	Grau Simples									
		1	2	3	4	5	6	7	8	12	14
Pearson	Antes	ELET6	ELET3	GGBR4	ITUB4 ^{AD}	*BBDC4 ^{AD}			PETR4		
	Depois	**PETR3	**BBDC4 ^{AD}	CPL6	ITUB4 ^{AD}	ITSA4	BBAS3				
Spearman	Antes	PETR3 ^{AD}	ELET6	ITUB4	*BBDC4 ^{AD}	CSNA3		BBDC3			
	Depois	**PETR3 ^{AD}	USIM5	**BBDC4 ^{AD}	PETR4			ELET3			CMIG4
Kendall	Antes	PETR3 ^{AD}	ELET3 ^{AD}	*BBDC4 ^{AD}	PETR4 ^{AD}	USIM5			BBDC3		
	Depois	**PETR3 ^{AD}	BRAP4	**BBDC4 ^{AD}	PETR4 ^{AD}	CSNA3	ELET3 ^{AD}			CMIG4	

Fonte: Elaborada pela autora.

Para finalizar, as Figuras 35(a)-(b) mostram os parâmetros de Molloy-Reed k_o e a robustez f_c definidos pelas Equações 2.15 e 2.16, respectivamente, e calculados para a distribuição de grau simples das redes antes e depois do evento. O parâmetro de Molloy-Reed é sempre maior que 2 em todos os casos e indicam a presença de uma componente gigante. Antes do evento, a robustez tem um valor próximo para todas as correlações e aumentam após o evento. O aumento da robustez indica que as redes ficaram mais rígidas após o evento. As redes de Spearman são as mais robustas após o evento, seguida pela de Kendall e Pearson, respectivamente. Se usarmos a analogia de que em um material mais rígido a velocidade de propagação do som é maior, a maior rigidez sugere que a rede de ações após o evento é mais facilmente afetada por novos acontecimentos do que a rede antes do evento. As árvores geradoras de Spearman e Kendall indicam uma rigidez ainda maior da rede, após o evento, do que a rede de Pearson. Uma possível explicação é que após o evento o número de dependências não lineares entre as ações tornaram-se relevantes, mas não detectadas pela correlação de Pearson.

Figura 34 – Grau ponderado $\times$ Grau simples das ações na árvore geradora mínima de (a) Pearson, (b) Spearman e (c) Kendall para os dados agrupados antes e depois do evento.

Fonte: Elaborada pela autora.

Figura 35 – (a) Parâmetro de Molloy-Reed e (b) robustez das árvores geradoras mínimas antes e depois do evento.

Fonte: Elaborada pela autora.

6 CONSIDERAÇÕES FINAIS

Neste trabalho exploramos as redes e suas árvores geradoras mínimas de ações do mercado financeiro brasileiro quando a medida de suas interações é baseada nas correlações Pearson, Spearman e Kendall sob o efeito de um evento. O evento de interesse foi uma notícia de cunho político publicada pelo jornal O Globo na noite de 17 de maio de 2017 sobre o então Presidente da República Michel Temer autorizando a compra do silêncio do Deputado Federal Eduardo Cunha que estava preso e teve o mandato cassado em 2016. Estudamos a estrutura de redes geradas pelas correlações e as mudanças resultantes, comparando-se as redes em um período anterior e um período posterior ao evento.

Foram coletados dados de alta frequência de 58 ações do Ibovespa no período de 11 a 25 de maio de 2017, fixando-se o dia 18 de maio de 2017 como o dia do evento, e, analisamos as redes de ações para cinco pregões antes e cinco depois do evento. As redes de ações e as respectivas árvores geradoras mínimas geradas pelas correlações de Pearson, Spearman e Kendall foram estudadas em duas escalas de tempo: (1) Redes diárias: cinco pregões antes do evento, o dia do evento e, cinco pregões depois do evento, com cotações a cada 15 minutos; (2) Agrupadas em antes e depois: agrupando os dados dos 5 dias antes e dos 5 dias depois do evento. O estudo das redes diárias indicou mudança de tendência nas suas propriedades no decorrer do período que contém o evento, com cotações a cada 15 minutos, e assim, foi possível estudar o efeito médio contido nas correlações dos dados agrupados antes e depois do evento.

O estudo das redes diárias foi útil para mostrar a mudança de tendência nas suas propriedades no decorrer do período que contém o evento. Isto sugeriu que análise do efeito médio contido nos dados agrupados antes e depois do evento poderiam tornar mais evidente mudanças na estrutura de rede de ações. As redes antes e depois do evento apresentaram mudanças nas suas métricas. As mudanças ficaram mais evidentes nas árvores geradoras mínimas, pois nestas, as correlações de maior intensidade são selecionadas em ordem de importância.

As árvores geradoras mínimas para dados agrupados antes e depois do evento apresentaram um número menor de agrupamentos das ações em todas as redes após o evento. As redes geradas pelas correlações de Kendall e Spearman apresentaram um número maior de agrupamentos tanto antes, como depois do evento em relação à rede de Pearson. As distribuições de grau ponderado após o evento indicam uma probabilidade maior de vértices com maior grau. As métricas das árvores mínimas geradoras por correlação de Spearman sofreram a maior variação, seguidas pelas de Kendall e Pearson e todas indicaram que as redes após o evento ficaram mais robustas, ou seja, mais rígidas. A maior robustez após o evento indica uma maior conectividade do mercado, tornando-o como um todo mais suscetível ao impacto de novos acontecimentos.

Para terminar, alguns pontos não considerados neste trabalho e que poderiam ser realizados em estudos posteriores são: obter dados de transações no mercado numa escala de tempo menor do que 15 minutos e comparar um número maior de escalas de tempo no período em torno do evento; introduzir nível de confiança para as correlações e estudar os seus efeitos nas redes completas variando-se o tamanho das amostras; replicar o estudo para outros casos.

REFERÊNCIAS

- ADVFN. *Bovespa: Índice Ibovespa fecha em forte baixa de -8,80quinta-feira, 18 de maio de 2017*. 2017. ADVFN News. Disponível em: <<https://br.advfn.com/jornal/2017/05/bovespa-indice-ibovespa-fecha-em-forte-baixa-de-8-80-nesta-quinta-feira-18-de-maio-de-2017>>.
- BAGLER, Ganesh. Analysis of the airport network of india as a complex weighted network. *Physica A: Statistical Mechanics and its Applications*, v. 387, n. 12, p. 2972 – 2980, 2008. ISSN 0378-4371. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378437108001131>>.
- BARABÁSI, Albert-László. *Network science*. [S.l.]: Cambridge university press, 2016.
- BARABÁSI, Albert-László; ALBERT, Réka. Emergence of scaling in random networks. *Science*, American Association for the Advancement of Science, v. 286, n. 5439, p. 509–512, 1999.
- BARABÁSI, Albert-László; GULBAHCE, Natali; LOSCALZO, Joseph. Network medicine: a network-based approach to human disease. *Nature reviews genetics*, Nature Publishing Group, v. 12, n. 1, p. 56, 2011.
- BARTHELEMY, Marc et al. Characterization and modeling of weighted networks. *Physica A*, v. 346, p. 34–43, 2005.
- BAVELAS, Alex. A mathematical model for group structures. *Human organization*, Society for Applied Anthropology, v. 7, n. 3, p. 16–30, 1948.
- BHATTACHARYYA, Gouri K.; JOHNSON, Richard A. *Statistical Concepts and Methods*. [S.l.]: John Wiley & Sons, 1977. ISBN 978-0-471-07204-1.
- BM&FBOVESPA. *Bolsa de valores, mercadorias e futuros de São Paulo*. 2017. BM&F BOVESPA. Disponível em: <http://www.bmfbovespa.com.br/pt_br/>.
- BONANNO, Giovanni et al. Topology of correlation-based minimal spanning trees in real and model markets. *Physical Review E*, APS, v. 68, n. 4, p. 046130, 2003.
- _____. Networks of equities in financial markets. *The European Physical Journal B*, Springer, v. 38, n. 2, p. 363–371, 2004.
- BONANNO, Giovanni; LILLO, Fabrizio; MANTEGNA, Rosario N. High-frequency cross-correlation in a set of stocks. Taylor & Francis, 2001.
- _____. Levels of complexity in financial markets. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 299, n. 1-2, p. 16–27, 2001.
- BRAVAIS, A. Sur les probabilités des erreurs de situation d'un point [on the probability of errors in the position of a point]. *Memoirs de l'Académie Royale des Sciences de l'Institut de France*, v. 9, p. 255–332, 1846.
- CALGARO, Fernanda; RAMALHO, Renan; GARCIA, Gustavo. *Eduardo Cunha é cassado*. 2016. G1 O portal de notícias da Globo. Disponível em: <<http://especiais.g1.globo.com/politica/2016/cassacao-de-cunha/>>.

COELHO, Ricardo et al. The evolution of interdependence in world equity markets: Evidence from minimum spanning trees. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 376, p. 455–466, 2007.

CORMEN, Thomas H. *Introduction to algorithms*. [S.l.]: MIT press, 2009.

CUTLER, David M; POTERBA, James M; SUMMERS, Lawrence H. *What moves stock prices?* [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 1988.

DEVIREN, Seyma Akkaya; DEVIREN, Bayram. The relationship between carbon dioxide emission and economic growth: Hierarchical structure methods. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 451, p. 429–439, 2016.

DJAUHARI, Maman Abdurachman; GAN, Siew Lee. Optimality problem of network topology in stocks market analysis. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 419, p. 108–114, 2015.

DJAUHARI, Maman A; GAN, Siew Lee. Network topology of economic sectors. *Journal of Statistical Mechanics: Theory and Experiment*, IOP Publishing, v. 2016, n. 9, p. 093401, 2016.

EASLEY, David; KLEINBERG, Jon. *Networks, crowds, and markets: Reasoning about a highly connected world*. [S.l.]: Cambridge University Press, 2010.

EISNER, Jason. State-of-the-art algorithms for minimum spanning trees-a tutorial discussion. Citeseer, 1997.

ERDÖS, Paul; RÉNYI, Alfréd. On random graphs, i. *Publicationes Mathematicae (Debrecen)*, v. 6, p. 290–297, 1959.

_____. Additive properties of random sequences of positive integers. *Acta Arithmetica*, v. 6, n. 1, p. 83–110, 1960.

_____. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, v. 5, p. 17–61, 1960 b.

FRANK, Ove; CARRINGTON, Peter J. Estimation of offending and co-offending using available data with model support. *Journal of Mathematical Sociology*, Taylor & Francis, v. 31, n. 1, p. 1–46, 2007.

FREEMAN, Linton C. A set of measures of centrality based on betweenness. *Sociometry*, JSTOR, p. 35–41, 1977.

GARAS, Antonios; ARGYRAKIS, Panos. Correlation study of the athens stock exchange. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 380, p. 399–410, 2007.

GILMORE, Claire G; LUCEY, Brian M; BOSCIA, Marian. An ever-closer union? examining the evolution of linkages of european equity markets via minimum spanning trees. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 387, n. 25, p. 6319–6329, 2008.

GODOY, Denyse; SILVA, Chrystiane. *Ibovespa amplia queda após discurso de Temer*. 2017. Valor Econômico online. Disponível em: <<http://www.valor.com.br/financas/4973218/ibovespa-amplia-queda-apos-discurso-de-temer>>.

GOH, Kwang-Il et al. The human disease network. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 104, n. 21, p. 8685–8690, 2007.

GOWER, John C. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, Oxford University Press, v. 53, n. 3-4, p. 325–338, 1966.

GRAHAM, Ronald L; HELL, Pavol. On the history of the minimum spanning tree problem. *Annals of the History of Computing*, IEEE, v. 7, n. 1, p. 43–57, 1985.

GRANOVETTER, Mark S. The strength of weak ties. *American Journal of Sociology*, v. 78, n. 6, p. 1360–1380, 1973.

GUIMERA, Roger; SALES-PARDO, Marta; AMARAL, Luís A Nunes. Module identification in bipartite and directed networks. *Physical Review E*, APS, v. 76, n. 3, p. 036102, 2007.

HAUKE, Jan; KOSSOWSKI, Tomasz. Comparison of values of pearson's and spearman's correlation coefficients on the same sets of data. *Quaestiones geographicae*, Versita, v. 30, n. 2, p. 87–93, 2011.

HUANG, Wei-Qiang et al. Dynamic asset trees in the us stock market: Structure variation and market phenomena. *Chaos, Solitons & Fractals*, Elsevier, v. 94, p. 44–53, 2017.

HUANG, Wei-Qiang; ZHUANG, Xin-Tian; YAO, Shuang. A network analysis of the chinese stock market. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 388, n. 14, p. 2956–2964, 2009.

JARDIM, Lauro. *Dono da JBS grava Temer dando aval para compra de silêncio de Cunha: Joesley Batista e o seu irmão Wesley confirmaram a Fachin o que falaram a PGR*. 2017. O Globo online. Disponível em: <https://oglobo.globo.com/brasil/dono-da-jbs-grava-temer-dando-aval-para-compra-de-silencio-de-cunha-21353935>.

JAYAWANT, Pallavi; GLAVIN, Kerry et al. Minimum spanning trees. *Involve: A Journal of Mathematics*, Mathematical Sciences Publishers, v. 2, n. 4, p. 439–450, 2009.

JOHNSON, Richard A.; BHATTACHARYYA, Gouri K. *Statistics: principles and methods*. [S.l.]: John Wiley & Sons, 2009.

JOHNSON, Richard A; WICHERN, Dean. *Multivariate analysis*. [S.l.]: Wiley Online Library, 2002.

JUNG, Woo-Sung et al. Characteristics of the korean stock market correlations. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 361, n. 1, p. 263–271, 2006.

JUNGNICKEL, Dieter. *Graph networks and algorithms: Algorithms and computation in mathematics*. Springer-Verlag, 2004, 2008.

KANDA, Patrick T et al. Forecasting south african inflation using non-linearmodels: a weighted loss-based evaluation. *Applied Economics*, Taylor & Francis, v. 48, n. 26, p. 2412–2427, 2016.

KRUSKAL, Joseph B. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical society*, v. 7, n. 1, p. 48–50, 1956.

KWAPIEN, Jaroslaw; GWOREK, Sylwia; DROZDZ, Stanislaw. Structure and evolution of the foreign exchange networks. *arXiv preprint arXiv:0901.4793*, 2009.

MALEVERGNE, Yannick; SORNETTE, Didier. *Extreme Financial Risks: From Dependence to Risk Management*. [S.l.]: Springer, 2006.

MANTEGNA, Rosario N. Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems*, Springer, v. 11, n. 1, p. 193–197, 1999.

MARI, Dominique Drouet; KOTZ, Samuel. *Correlation and dependence*. [S.l.]: World Scientific, 2001.

MARKOWITZ, Harry. Portfolio selection. *The journal of finance*, Wiley Online Library, v. 7, n. 1, p. 77–91, 1952.

MICCICHÈ, Salvatore et al. Degree stability of a minimum spanning tree of price return and volatility. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 324, n. 1-2, p. 66–73, 2003.

MILKOVÁ, Eva. The minimum spanning tree problem: Jarník's solution in historical and present context. *Electronic notes in discrete mathematics*, Elsevier, v. 28, p. 309–316, 2007.

MOLLOY, Michael; REED, Bruce. A critical point for random graphs with a given degree sequence. *Random structures & algorithms*, Wiley Online Library, v. 6, n. 2-3, p. 161–180, 1995.

NAYLOR, Michael J; ROSE, Lawrence C; MOYLE, Brendan J. Topology of foreign exchange markets using hierarchical structure methods. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 382, n. 1, p. 199–208, 2007.

NEWMAN, Mark. *Networks: an introduction*. [S.l.]: Oxford university press, 2010.

NEWMAN, Mark EJ. Modularity and community structure in networks. *Proceedings of the national academy of sciences*, National Acad Sciences, v. 103, n. 23, p. 8577–8582, 2006.

NEWMAN, M. E. J. Analysis of weighted networks. *Physical Review E*, v. 70, p. 056131, 2004.

NOBI, Ashadun et al. Structural changes in the minimal spanning tree and the hierarchical network in the korean stock market around the global financial crisis. *Journal of the Korean Physical Society*, Springer, v. 66, n. 8, p. 1153–1159, 2015.

OH, Seung Wook et al. A mesoscale connectome of the mouse brain. *Nature*, Nature Publishing Group, v. 508, n. 7495, p. 207, 2014.

ONNELA, J-P et al. Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E*, APS, v. 68, n. 5, p. 056110, 2003.

ONNELA, J-P; KASKI, Kimmo; KERTÉSZ, Janos. Clustering and information in correlation based financial networks. *The European Physical Journal B*, Springer, v. 38, n. 2, p. 353–362, 2004.

ORTEGA, Guillermo J; MATESANZ, David. Cross-country hierarchical structure and currency crises. *International Journal of Modern Physics C*, World Scientific, v. 17, n. 03, p. 333–341, 2006.

PARK, Kwangho; LAI, Ying-Cheng; YE, Nong. Characterization of weighted complex networks. *Physical Review E*, v. 70, p. 026109, 2004.

PEARSON, Karl. Mathematical contributions to the theory of evolution. iii. regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*, JSTOR, v. 187, p. 253–318, 1896.

- PECORA, Nicolò; SPELTA, Alessandro. Shareholding relationships in the euro area banking market: A network perspective. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 434, p. 1–12, 2015.
- PERALTA, Gustavo; ZAREEI, Abalfazl. A network approach to portfolio selection. *Journal of Empirical Finance*, Elsevier, v. 38, p. 157–180, 2016.
- PERLIN, Marcelo; RAMOS, Henrique. Package ‘GetHFDData’: Download and Aggregate High Frequency Trading Data from Bovespa. 2017. Disponível em: <<https://cran.r-project.org/web/packages/GetHFDData/index.html>>.
- POZZI, Francesco; MATTEO, Tiziana Di; ASTE, Tomaso. Centrality and peripherality in filtered graphs from dynamical financial correlations. *Advances in Complex Systems*, World Scientific, v. 11, n. 06, p. 927–950, 2008.
- PRICE, Derek J De Solla. Networks of scientific papers. *Science*, JSTOR, p. 510–515, 1965.
- RAPOPORT, Anatol. Contribution to the theory of random and biased nets. In: *Social Networks*. [S.l.]: Elsevier, 1977. p. 389–409.
- RAPOPORT, Anatol; HORVATH, William J. A study of a large sociogram. *Systems Research and Behavioral Science*, Wiley Online Library, v. 6, n. 4, p. 279–291, 1961.
- ROBINS, Garry; ALEXANDER, Malcolm. Small worlds among interlocking directors: Network structure and distance in bipartite graphs. *Computational & Mathematical Organization Theory*, Springer, v. 10, n. 1, p. 69–94, 2004.
- SEDGEWICK, Robert; WAYNE, Kevin. *Algorithms*. [S.l.]: Addison-Wesley Professional, 2011.
- SIECZKA, Paweł; HOŁYST, Janusz A. Correlations in commodity markets. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 388, n. 8, p. 1621–1630, 2009.
- SILVA, CAT; CARVALHO, CC; NUNES, DMS. O que move o preço da ação. *Uma Abordagem sobre a Influência das Notícias no Mercado Acionário. REUNIR–Revista de Administração, Contabilidade e Sustentabilidade*, v. 2, n. 3, p. 1–13, 2013.
- SOLOMONOFF, Ray; RAPOPORT, Anatol. Connectivity of random nets. *The bulletin of mathematical biophysics*, Springer, v. 13, n. 2, p. 107–117, 1951.
- SPEARMAN, Charles. The proof and measurement of association between two things. *The American journal of psychology*, JSTOR, v. 15, n. 1, p. 72–101, 1904.
- SPORNS, Olaf; TONONI, Giulio; KÖTTER, Rolf. The human connectome: a structural description of the human brain. *PLoS computational biology*, Public Library of Science, v. 1, n. 4, p. e42, 2005.
- TABAK, Benjamin M; SERRA, Thiago R; CAJUEIRO, Daniel O. The expectation hypothesis of interest rates and network theory: The case of brazil. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 388, n. 7, p. 1137–1149, 2009.
- _____. Topological properties of stock market networks: The case of brazil. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 389, n. 16, p. 3240–3249, 2010.

TREVIZAN, Karina. *Bovespa fecha na maior queda em quase 9 anos após denúncias da JBS*. 2017. G1 O portal de notícias da Globo. Disponível em: <https://g1.globo.com/economia/mercados/noticia/bovespa-fecha-em-forte-queda-de-olho-em-denuncias-sobre-temer.ghtml>.

TUMMINELLO, Michele et al. Spanning trees and bootstrap reliability estimation in correlation-based networks. *International Journal of Bifurcation and Chaos*, World Scientific, v. 17, n. 07, p. 2319–2329, 2007.

_____. Correlation based networks of equity returns sampled at different time horizons. *The European Physical Journal B*, Springer, v. 55, n. 2, p. 209–217, 2007.

VANDEWALLE, N et al. Non-random topology of stock markets. *Quantitative Finance*, Taylor & Francis, v. 1, n. 3, p. 372–374, 2001.

WATTS, Duncan J; STROGATZ, Steven H. Collective dynamics of small-world networks. *nature*, Nature Publishing Group, v. 393, n. 6684, p. 440, 1998.

WU, Bang Ye; CHAO, Kun-Mao. *Spanning trees and optimization problems*. [S.l.]: CRC Press, 2004.

YU, Dejian; SHI, Shunshun. Researching the development of atanassov intuitionistic fuzzy set: Using a citation network analysis. *Applied Soft Computing*, Elsevier, v. 32, p. 189–198, 2015.

ZENG, Zhi-Jian et al. Are stock market networks non-fractal? evidence from new york stock exchange. *Finance Research Letters*, Elsevier, v. 17, p. 97–102, 2016.

APÊNDICE A – ÁRVORES GERADORAS MÍNIMAS DIÁRIAS

Este apêndice contém as figuras das árvores geradoras mínima diárias.

Figura A1 – Árvore geradora mínima de Pearson do quinto dia anterior ao evento.

Fonte: Elaborada pela autora.

Figura A2 – Árvore geradora mínima de Spearman do quinto dia anterior ao evento.
Fonte: Elaborada pela autora.

Figura A3 – Árvore geradora mínima de Kendall do quinto dia anterior ao evento.
Fonte: Elaborada pela autora.

Figura A4 – Árvore geradora mínima de Pearson do quarto dia anterior ao evento.

Fonte: Elaborada pela autora.

Figura A5 – Árvore geradora mínima de Spearman do quarto dia anterior ao evento.

Fonte: Elaborada pela autora.

The diagram illustrates a complex network of protein interactions. The nodes are represented by colored circles, and the edges represent interactions between them. The network is organized into several distinct clusters. The leftmost cluster, primarily in shades of green and purple, includes proteins such as KROT2, EQL3, ELET3, and PETR4. The central cluster, featuring teal and blue nodes, contains HYPE3, ITSA4, CCRO3, and TIMP3. The rightmost cluster, dominated by orange and blue nodes, includes BBAS3, BBSE3, NATU3, BBDC3, GGBR4, GOAU4, CSNA3, RENT3, SANB11, SUZB5, and ENR3. The connections between these clusters suggest a highly integrated signaling pathway or regulatory network.

Figura A7 – Árvore geradora mínima de Pearson do terceiro dia anterior ao evento.
Fonte: Elaborada pela autora.

Fonte: Elaborada pela autora.

Fonte: Elaborada pela autora.

Figura A12 – Árvore geradora mínima de Kendall do segundo dia anterior ao evento.

Fonte: Elaborada pela autora.

Figura A13 – Árvore geradora mínima de Pearson do primeiro dia anterior ao evento.

Fonte: Elaborada pela autora.

Fonte: Elaborada pela autora.

Fonte: Elaborada pela autora.

Figura A16 – Árvore geradora mínima de Pearson do dia do evento.

Fonte: Elaborada pela autora.

Figura A17 – Árvore geradora mínima de Spearman do dia do evento.

Fonte: Elaborada pela autora.

Fonte: Elaborada pela autora.

Fonte: Elaborada pela autora.

Figura A22 – Árvore geradora mínima de Pearson do segundo dia após o evento.

Fonte: Elaborada pela autora.

Figura A23 – Árvore geradora mínima de Spearman do segundo dia após o evento.

Fonte: Elaborada pela autora.

Figura A25 – Árvore geradora mínima de Pearson do terceiro dia após o evento.
Fonte: Elaborada pela autora.

Figura A26 – Árvore geradora mínima de Spearman do terceiro dia após o evento.

Fonte: Elaborada pela autora.

Figura A27 – Árvore geradora mínima de Kendall do terceiro dia após o evento.

Fonte: Elaborada pela autora.

The network diagram illustrates a complex web of gene-gene interactions. The nodes are represented by colored circles of varying sizes, indicating their degree or importance in the network. The central node is ITSA4 (pink). Other prominent nodes include CMIG4 (green), USIM5 (pink), SANB11 (purple), and ABEV3 (orange). The network is highly interconnected, with many nodes having multiple connections.

Figura A29 – Árvore geradora mínima de Spearman do quarto dia após o evento.
Fonte: Elaborada pela autora.

Figura A30 – Árvore geradora mínima de Kendall do quarto dia após o evento.

Fonte: Elaborada pela autora.

Figura A31 – Árvore geradora mínima de Pearson do quinto dia após o evento.

Fonte: Elaborada pela autora.

Figura A32 – Árvore geradora mínima de Spearman do quinto dia após o evento.
Fonte: Elaborada pela autora.

Figura A33 – Árvore geradora mínima de Kendall do quinto dia após o evento.
Fonte: Elaborada pela autora.