

Escalonamento multidimensional baseado em redes de árvore geradora mínima como proposta à análise e visualização de dados: uma investigação sobre as ações brasileiras

Alex Quintino Barbi¹, Gilberto Aparecido Prativiera²

¹Faculdade de Economia, Administração e Contabilidade
Universidade de São Paulo (USP)
Ribeirão Preto – SP – Brasil

²Faculdade de Economia, Administração e Contabilidade
Universidade de São Paulo (USP)
Ribeirão Preto – SP – Brasil

alex.barbi@usp.br, prativiera@usp.br

Abstract. *The empirical study based on network theory has greatly contributed as a methodology of financial market visualization and analysis. In this work, we propose to apply multidimensional scaling technique on network elements of a minimum spanning tree type as a technique for financial networks analysis and visualization. The method was applied in Brazilian stock market data. The analysis of the results showed that the use of multidimensional scaling on minimum spanning tree elements succeeded at simplifying the visualization of information in the data, beyond what would be obtained only processing the minimum spanning tree. On the observed dimensions (factors) of the market, we were able to identify some special segments of the Brazilian economy. In addition, the analysis of the minimum spanning tree showed that central stocks in the network followed more closely the behavior of the local general market index, while the more peripheral ones had a quite different behavior. This dynamic seems to be related to the so-called volatility paths, visual structures in which cascading effects might appear in stock returns. This study may be useful for investors and managers interested in network techniques for viewing and analysing stock market properties.*

Resumo. *O estudo empírico baseado em teoria de redes tem contribuído muito como metodologia de visualização e análise do comportamento do mercado financeiro. Neste trabalho propõe-se aplicar técnica de escalonamento multidimensional sobre elementos de redes do tipo árvore geradora mínima como metodologia de análise e visualização de redes financeiras. O método foi aplicado em dados do mercado brasileiro de ações. A análise dos resultados mostrou que a utilização do escalonamento sobre os elementos da árvore geradora mínima conseguiu ainda simplificar a visualização de informação nos dados, além do que seria obtido somente pelo processamento desta árvore. Nas dimensões observadas do mercado pudemos identificar alguns segmentos especiais da economia brasileira. Além disso, a análise da árvore geradora mínima mostrou que as ações mais centrais na rede seguiram mais de perto o comportamento do índice geral do mercado local, enquanto que, as ações mais periféricas, tiveram comportamento bastante distinto. Essa dinâmica parece ter relação com os*

chamados aqui de caminhos de volatilidade, estruturas visuais nas quais podem ocorrer fenômenos do tipo efeito cascata nos retornos das ações. Este estudo pode ser útil para investidores e gestores interessados em técnicas de redes para a visualização e investigação das propriedades do mercado de ações.

1. Introdução

Uma rede (ou grafo) é uma representação da estrutura de ligação de objetos conectados por algum tipo de relação, sendo que esse tipo de representação é útil para mapear e compreender o comportamento de sistemas complexos [Barabasi 2016]. Em particular, a teoria de redes tem ajudado a caracterizar interdependências no estudo de vários componentes do mercado financeiro, em especial o acionário. Estudos de redes obtidas a partir da correlação entre ações apontam diferentes estruturas para cada cenário do mercado, e neste caso, indicam que seria possível antever eventos extremos a partir da composição destas redes [Bonanno et al. 2001].

Já os retornos de diferentes ações no mercado financeiro têm sido estudados como medida de informação econômica, evidenciando os diversos setores da economia [Mantegna 1999, Micciché et al. 2003, Bonanno et al. 2004]. Com alguma medida de similaridade entre dois objetos (neste caso, ações), consegue-se obter uma representação visual por algumas metodologias de visualização e análise de dados, tais como o escalonamento multidimensional (EMD), a análise de agrupamentos, e mais recentemente, a análise de redes.

Estudos usando técnicas de redes procurando entender a estrutura do mercado de ações têm atraído pesquisadores das mais diversas áreas como físicos, cientistas da computação, economistas e administradores, na busca em saber qual a estrutura empírica do mercado de ações, especialmente para o desenvolvimento de novos modelos, em geral, de risco financeiro [Mantegna 1999, Micciché et al. 2003, Bonanno et al. 2004, Sensoy and Tabak 2014]. Um dos caminhos para tratar de redes financeiras é a procura por arranjos hierárquicos que podem estar relacionados aos movimentos dos preços dos ativos, pois as séries temporais dos ativos financeiros carregavam valiosas informações econômicas [Mantegna 1999]. Outros estudos evidenciaram que as ligações entre os ativos estavam longe de serem árvores randômicas, e que era possível que os arranjos hierárquicos encontrados pudessem se repetir em diversas janelas de tempo [Brida and Rizzo 2008]. Concluiu-se que *deveria haver uma regra mais profunda que controlasse as propriedades estatísticas do sistema financeiro global tanto em dias típicos como durante eventos extremos* [Bonanno et al. 2001, p. 26]. Tais estudos contribuíram para que as pesquisas em redes no mercado de ações fossem intensificadas.

Os estudos em análise de redes no mercado de ações usam técnicas e estatísticas de redes na análise da topologia (arranjos) de diferentes mercados [Mantegna 1999, Coelho et al. 2007, Dias 2013, Gilmore et al. 2010, Sensoy and Tabak 2014], e na aplicação de diferentes métricas de similaridade e informação entre as ações [Tumminello et al. 2010, Yang et al. 2014, Fiedor 2014].

Os primeiros estudos em redes financeiras datam de 1999, com o surgimento do artigo *Hierarchical structure in financial markets* [Mantegna 1999], que, utilizando uma medida de distância baseada na correlação de Pearson, mostrou pela técnica de árvore geradora mínima que as ações se arranjavam de maneira hierárquica, for-

necendo informações acerca da taxonomia dos setores econômicos. Demais estudos começaram também a discutir o papel da variação temporal da volatilidade dos mercados através de medidas de rede, como por meio de sua distribuição de grau [Bonanno et al. 2004]. Outras pesquisas [Coelho et al. 2007, Sensoy and Tabak 2014] indicaram que a estrutura da rede possa se alterar de acordo com o risco existente no mercado acionário. Fato semelhante é observado para outros ativos financeiros, como taxas de juros e moedas [Dias 2013, Gilmore et al. 2010]. Por outro lado, alguns trabalhos avançaram na avaliação do estudo da dependência no mercado financeiro, adotando outras métricas de similaridade, como é o caso da chamada informação mútua [Yang et al. 2014, Fiedor 2014], e que pode revelar estruturas não lineares entre as ações, uma alternativa à análise por correlação linear de Pearson [Dionisio et al. 2004]. Apesar dos diversos estudos elaborados, a falta de trabalhos desenvolvidos para o mercado financeiro local abre uma lacuna para novos estudos empíricos nesta área, bem como sua avaliação frente a outras alternativas de visualização e análise de dados.

Neste trabalho é realizado um estudo da estrutura da árvore mínima geradora para o mercado brasileiro de ações a partir de uma distância definida por correlações entre os retornos das ações. Métricas de centralidade da árvore mínima são obtidas e analisadas. Também propomos uma alternativa para visualização de dados por meio de um escalonamento multidimensional baseado apenas nas distâncias contidas na árvore mínima geradora.

O artigo está organizado da seguinte forma: na Seção 2 são apresentados os dados estudados, os métodos de análise baseados na árvore mínima geradora e no escalonamento multidimensional. Na Seção 3 apresentamos os resultados contendo a análise da árvore mínima geradora e do escalonamento multidimensional obtido a partir desta. Finalmente, na Seção 4 são apresentadas as conclusões.

2. Métodos

Nossa base de dados é composta por séries temporais de 99 ações durante 386 pregões (negociações de período de um dia). Para entrar na base, uma ação deve ter pelo menos uma vez participado do Ibovespa (Índice Bovespa) no período em análise, que variou de janeiro de 2014 a dezembro de 2014 (período de baixa) e de janeiro de 2015 a julho de 2015 (período de alta). O período escolhido foi selecionado para evitar viés de alta ou baixa do mercado. Foram calculados os log retorno, R_t , das ações da seguinte maneira

$$R_t = \ln \left(\frac{p_t}{p_{t-1}} \right), \quad (1)$$

onde p_t é o preço atual da ação no instante t , e p_{t-1} é o preço no instante imediatamente anterior.

Uma medida típica de dependência entre objetos é o coeficiente de correlação de Pearson. No nosso caso, os objetos são ações do mercado e quantificadas pelos seus retornos. A partir do coeficiente de correlação podemos definir a seguinte medida de distância [Mantegna and Stanley 1999]

$$d_{ij} = \sqrt{2(1 - r_{ij})}, \quad (2)$$

em que r_{ij} é a correlação de Pearson entre os retornos R_t das ações i e j . Dessa maneira, se a correlação for -1, devemos obter distância igual a 2 (máxima), já no caso da correlação ser igual a 1, a distância é 0 (mínima). Penaliza-se, assim, ações com intercorrelações negativas, isolando-as das comunidades.

2.1. Escalonamento multidimensional via árvore geradora mínima

Para se encontrar o escalonamento multidimensional (EMD) via rede baseada na árvore geradora mínima (AGM) foram definidos de acordos com os passos a seguir:

1. Obter a correlação r_{ij} entre os retornos das ações i e j .
2. Obter uma matriz de distâncias a partir da Equação (2).
3. Implementar o algoritmo da árvore geradora mínima (AGM) pelo método de Kruskal [Kruskal 1956]. O procedimento consiste em ligar dois elementos com a menor distância na rede até que não haja mais elementos, dada a condição de que não se criem ciclos fechados. Encontra-se, assim, uma árvore que conecta todos os elementos com $n - 1$ ligações, mediante a minimização da soma de distâncias das arestas (ligações) da rede.
4. Aplicar o escalonamento multidimensional usando somente as distâncias dos objetos ligados na árvore geradora mínima. O algoritmo de EMD tenta plotar os objetos em R^n , preservando ao máximo possível as distâncias originais [Cormen et al. 2001]. Então, cada objeto é designado em coordenadas para cada N -dimensão, no caso aqui, $N = 2$ dimensões. O algoritmo de Kruskal [Kruskal 1964] é um dos mais utilizados na técnica do EMD.

Nessa metodologia, podemos visualizar as coordenadas em um espaço bidimensional xy , onde o escalonamento multidimensional utiliza somente as distâncias das ligações encontradas pelo algoritmo de Kruskal [Kruskal 1964]. Esse passo facilita visualizar somente as ligações que têm peso baixo (altas correlações), eliminando as com peso muito alto (baixas correlações), de forma a extrair as informações mais úteis da rede. A simplificação da AGM tem como problema o fato de esta poder excluir algumas ligações ainda de certa forma relevantes. Porém seu uso é uma maneira interessante de verificar o ganho de informação para a análise, quando eliminado boa parte do ruído. O tópico seguinte trata da árvore geradora mínima e algumas métricas associadas a ela.

2.2. Métricas da árvore geradora mínima

A partir da medida de distância definida pela Eq. (2), podemos obter a AGM e calcular algumas métricas para caracterizá-la. Uma delas é o algoritmo chamado de modularidade (Q), que é utilizado para se encontrar comunidades (grupos) em redes. O algoritmo de modularidade é designado para medir a força da divisão dos módulos (grupos), assim, redes com alta modularidade têm densas conexões entre os nós dentro do mesmo módulo e esparsas conexões entre diferentes módulos. Formalmente, é a fração de ligações que se distribuem dentro de dado grupo menos o que seria esperado caso as ligações fossem distribuídas em grafos aleatórios com a mesma distribuição de grau da rede em análise [Newman 2010, Barabasi 2016]. O valor a ser maximizado (Q) varia de -1/2 (pior) a 1 (melhor). Matematicamente, pode-se calcular Q por

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta_{(C_i, C_j)}, \quad (3)$$

em que m é o número de vértices, A_{ij} representa o total de número de vértices que se ligam entre i e j , $\frac{k_i k_j}{2m}$ representa o número esperado de arestas entre i e j , e $\delta_{(C_i, C_j)}$ é o delta de Kronecker, definido como sendo 1 para elementos pertencentes ao mesmo componente, e 0, para quando estes pertencem a diferentes componentes. A análise dos grupos pode nos evidenciar estruturas de redes fragmentadas ou então bastante densas, de forma a evidenciar grupos de volatilidade.

Outra medida, útil para caracterizar individualmente as ações é centralidade, pois denota-se quais ações são mais centrais naquela rede. Uma ação ser muito central significa que, uma vez atingida por alguma volatilidade extrema do mercado, volatilidade esta medida pela variância do retorno, ela será a intermediadora que espalha, por efeito cascata, esse efeito ao resto do mercado. Dentre muitas medidas, para este trabalho exploramos a centralidade baseada em autovetores (Cav). A Cav x_i de um nó (ação) i é proporcional a somatória da centralidade de seus i 's vizinhos. A centralidade baseada em autovetor tem limites entre 0, o menos central, e 1, o mais central. Um indivíduo em uma rede social, por exemplo, pode ser importante por conhecer muitas pessoas, ou simplesmente por conhecer algumas pessoas que são realmente muito importantes. Matematicamente, a centralidade x_i é dada por [Newman 2010]

$$x_i = \gamma_1^{-1} \sum_j A_{ij} x_j, \quad (4)$$

na qual γ_1^{-1} representa o inverso do maior autovalor de A , sendo que A é a matriz de adjacência ponderada somente pelas correlações dos elementos que compõem a AGM, e x_j é o grau do nó j .

3. Resultados e discussões

Esta seção dedica-se a explorar os resultados obtidos pela estrutura da AGM e pela posterior análise por EMD via AGM.

3.1. Explorando a estrutura da árvore geradora mínima

Nesta seção apresentamos a AGM elaborada e suas propriedades. Para a visualização da AGM usamos o algoritmo *Force Atlas 2*, que repulsa os nós dissimilares por meio de um procedimento de gravidade, empurrando os nós mais centrais para a periferia do grafo [Jacomy et al. 2014]. A Figura 1 apresenta a estrutura da AGM.

A Tabela 1 apresenta as centralidades baseadas em autovetor das ações da rede analisada das ações mais centrais para a rede, bem como as mais periféricas (menos centrais). Esse tipo de centralidade pode estar ligada a efeitos cascata nas volatilidades das ações [Di Matteo et al. 2010]. Volatilidade alta nas ações com maior centralidade induz maior volatilidade ao mercado como um todo, propagando-se a outras ações por aquelas mais centrais, a saber, as melhores ranqueadas da Tabela 1. Por outro lado, alta volatilidade em ações como EMBR3 e DASA3 (periféricas) não afetam o mercado como um todo, pelo menos de imediato.

Pela rede financeira elaborada podemos verificar alguns caminhos para transmissão mais rápida da volatilidade. Na Figura 1, vemos que BBDC4 (Banco Bradesco S/A, em rosa) se relaciona com ITSA4 (em azul claro, controladora do Banco Itaú S/A)

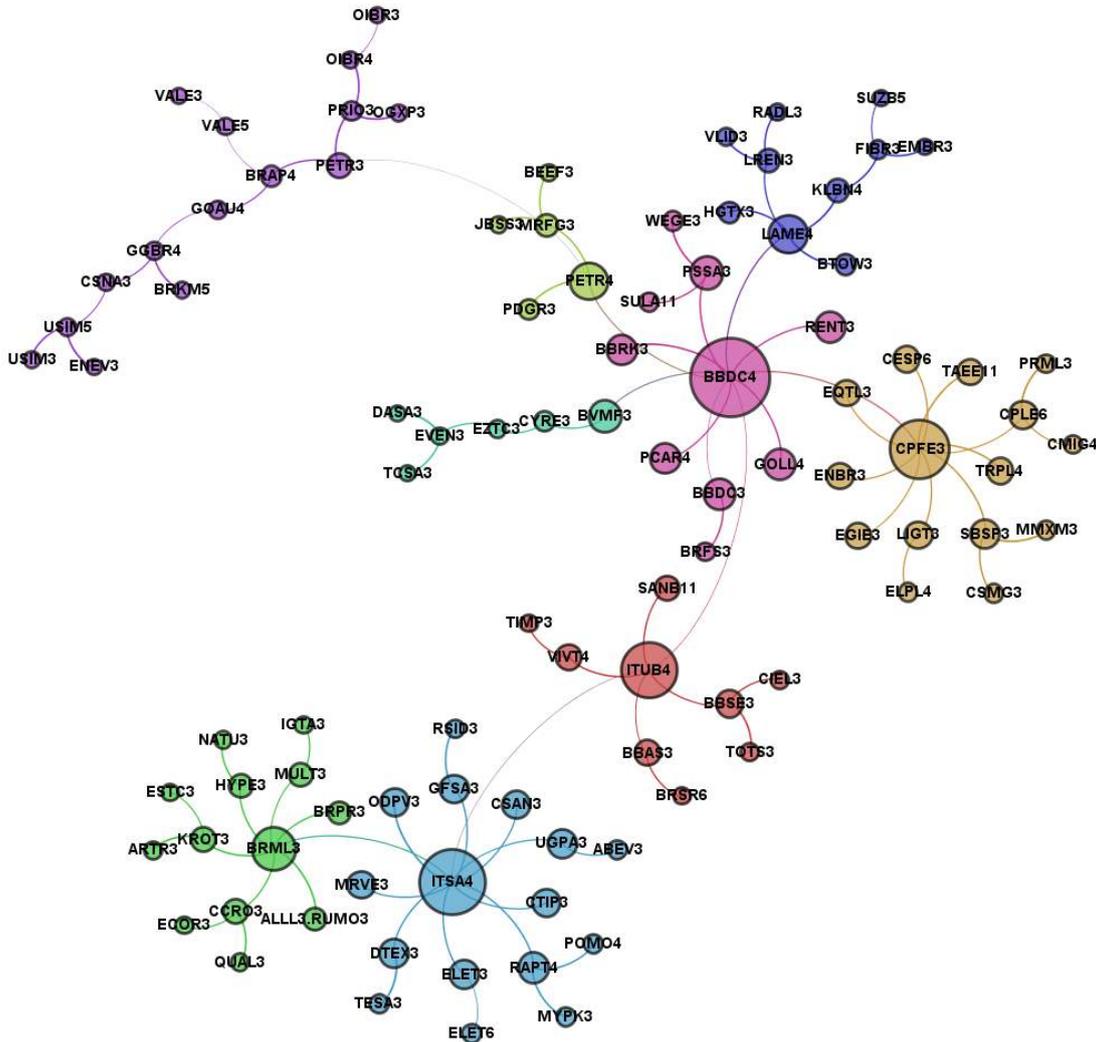


Figura 1. Rede baseada em árvore mínima geradora com os grupos formados pela modularidade (9 comunidades). Nessa rede, o peso total da estrutura é 94,37 e a medida de Q é 0,82, indicando considerável formação de grupos distintos. Nessa apresentação, pode-se verificar os principais agrupamentos de volatilidade (ações mais similares entre si). Já a medida de peso total pode ser considerada posteriormente quando da avaliação de várias composições de redes. Por exemplo, um peso total menor denota maior estrutura de dependência entre aqueles ativos. No caso da rede proposta, tem-se um peso relativamente alto. Neste caso, verificamos que a rede é razoavelmente esparsa, no sentido de que esta possui caminhos muito longínquos e menor estrutura geral de dependência baseada na correlação de Pearson. Nota: cada grupo (módulo) é ilustrado por uma cor diferente.

e ITUB4 (em vermelho, Banco Itaú S/A), onde também, neste caso, várias outras ações possuem ligações bastante próximas àquelas principais. Assim, podemos dizer que este é um ‘caminho de volatilidade’, quando, por exemplo, ITUB4 ou BBDC4 for afetado por algum evento extremo, rapidamente as ações mais próximas sofrerão impacto mais imediato, num efeito típico em cascata. Já em relação a VALE5 (em roxo), o mesmo não pode

Tabela 1. Centralidade de ações: as mais centrais, como ITUB4 e PETR4 são referência para outras ações, enquanto que as menos centrais, como DASA3 e EMBR3 são ações bastante específicas, não sendo referência para o mercado. Pela tabela notamos que, dentre as ações, aquelas do primeiro grupo (1-10) têm desempenho conjunto mais semelhante àquele ocorrido no índice geral de mercado, o IBOVESPA. Já as ações do grupo (90-99) são bastante distintas do índice em termos de retorno total R_T , médio (μ) e desvio padrão (σ).

Ordem	Ativos	Observações	Centralidade	R_T	μ	σ
1	BBDC4	Banco Bradesco PN	1.0000	-11%		
2	ITSA4	Holding ITAUSA	0.8031	6%		
3	CPFE3	CPFL Energia	0.7086	-11%		
4	ITUB4	Banco Itaú	0.6458	10%		
5	BRML3	Holding BRMalls	0.4517	-31%	-2.35%	31.18%
6	LAME4	Lojas Americanas	0.3841	60%		
7	PETR4	Petrobrás	0.3622	-57%		
8	PSSA3	Porto Seguro Seguros	0.2923	13%		
9	BVMF3	Ações Bovespa	0.2770	14%		
10	BBDC3	Banco Bradesco ON	0.2686	-15%		
90	OGXP3	Oléo Gás Partip. X	0.0295	-87%		
91	BRKM5	Braskem	0.0244	48%		
92	EMBR3	Embraer	0.0243	61%		
93	SUZB5	Suzano Papel	0.0243	105%		
94	ENEV3	MPX Energia	0.0209	-92%	-27.53%	72.27%
95	USIM3	Usiminas	0.0209	-65%		
96	DASA3	DASA - Análises clínicas	0.0206	-29%		
97	TCSA3	Tecnisa - Construtora	0.0206	-66%		
98	VALE3	Companhia Vale Rio Doce	0.0184	-58%		
99	OIBR3	Oi Brasil	0.0170	-93%		
I	IBOV	Índice Bovespa	-	-16.58%	-16.58%	7.35%

ser dito. Movimentações nesta ação não costuma ter como consequência muitos destes efeitos, ao menos no mercado como um todo.

Nesse sentido, vemos que a AGM pode estabelecer uma alternativa à visualização tradicional, eliminando ligações fracas. Além disso, pode-se adotar medidas da rede como importantes características da estrutura do mercado acionário. A medida de centralidade baseada em autovetores parece ser bastante importante para se verificar quais ativos ‘se parecem’ mais com o mercado como um todo. Os ativos BBDC4, ITSA4, ITUB4, CPFE3 e PETR4, nesta ordem, são os mais centrais da análise. Um investidor que gostaria de replicar o desempenho do mercado em geral pode utilizar desse conceito para se investir nestas ações que possuem as maiores centralidades, evidenciadas pela Tabela 1 (ativos 1-10). De modo contrário, um investidor que gostaria de se proteger contra possíveis efeitos ‘cascata’ no mercado geral poderia se proteger investindo em ações mais periféricas (ativos 90-99), como Embraer e Suzano Papel.

3.2. O escalonamento bidimensional do mercado via AGM

O chamado escalonamento multidimensional (EMD) é designado para construir um diagrama mostrando as relações entre objetos, dada apenas uma matriz de distâncias [Manly 2004]. Este diagrama pode ser mapeado em n -dimensões, mas os mais comuns são duas (plano) e três dimensões (pontos no espaço). Infelizmente, para dados reais, não é conhecido *a priori* o número de dimensões requerida para esta representação. Dessa maneira, a análise via EMD faz um procedimento de redução dimensional para que possamos visualizar os pontos em duas, talvez três dimensões, incorrendo em algum tipo de função de perda, sendo esta decorrente desta própria redução. O algoritmo de Kruskal [Kruskal 1964] é um clássico procedimento para o escalonamento, e será utilizado neste trabalho. Uma medida desse desvio das distâncias originais (perda) é o chamado *stress*, dado por

$$S(x_1, x_2, x_n) = \left(\sum_{i \neq j=1, \dots, N} (d_{ij} - \|x_i - x_j\|)^2 \right)^{\frac{1}{2}}, \quad (5)$$

onde d_{ij} são as distâncias originais e $\|x_i - x_j\|$, as novas distâncias calculadas pelo algoritmo de EMD para uma n -dimensão, chamadas de disparidades. Quando o EMD produz um mapa perfeito, então $d_{ij} - \|x_i - x_j\|$ é zero para qualquer i e j , e neste caso, o *stress* é zero.

A Figura 2 mostra o mapa pela tradicional técnica de escalonamento multidimensional, com a distância definida pela equação (2), sem o uso da árvore geradora mínima como entrada. Desse modo, o mapa dessa representação apresenta o escalonamento com todas as distâncias possíveis incluídas. Já a Figura 3 mostra o mapa obtido pelo escalonamento realizado somente com as distâncias da AGM. A medida de *stress* S é igual a 0,049. Neste caso, como $S \neq 0$, a utilização do EMD sobre a AGM ainda conseguiu simplificar a visualização de informação nos dados, além do que seria obtido somente via AGM. Se o *stress* S fosse igual a zero, não haveria diferenças entre a AGM e o escalonamento em termos de mapeamento. Pelo fato de $S \neq 0$, o EMD e a AGM diferem, e nesse sentido, faz-se interessante a utilização do EMD via AGM.

O EMD baseado AGM (Figura 3) tem boa adequabilidade, uma vez que podemos separar numa primeira dimensão, uma medida de risco, representada pelas ações mais voltadas ao mercado interno e geralmente com contratos de concessão, neste caso, os grupos em azul claro e rosa claro, em relação às mais industriais, de materiais pesados ou mesmo exportadoras de materiais primários, em azul escuro. Na comparação dos métodos, alguns comentários devem ser feitos. Não há como comparar o EMD tradicional com a EMD baseado em redes de árvore geradora mínima de maneira imediata, pois a primeira utiliza todas as dissimilaridades para formação do mapa dimensional, enquanto que na segunda, a MST reduz o número de ligações, reduzindo a medida de *stress* (representa perda de informação) deste mapa. Assim, tratamos a visualização de EMD por meio de redes como uma alternativa à visualização por EMD tradicional sem utilização de redes. O fato de $S \neq 0$, quando implementado o EMD sobre a AGM, corrobora esta ideia central desenvolvida neste trabalho.

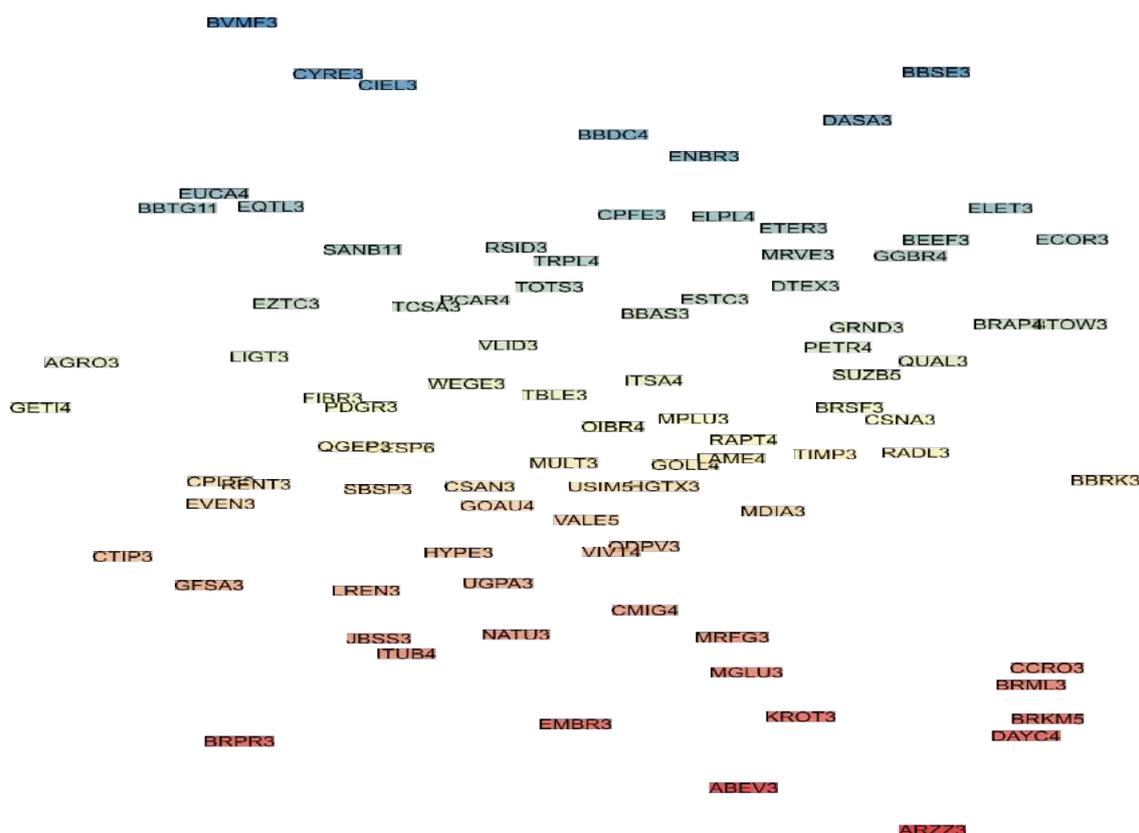


Figura 2. O EMD representado pela distância adotada pela equação (2), sem a utilização da metodologia baseada em rede. A representação do escalonamento baseada em todas as distâncias é pouco ilustrativa em termos de setores da economia, já que a medida de *stress* ficou em 0,27, representando uma razoável perda das informações das distâncias originais.

4. Conclusões

O artigo explorou, pela metodologia baseada em redes, algumas das características topológicas dessas estruturas no mercado brasileiro de ações. A análise foi baseada nas metodologias de algoritmos de redução dimensional, análise de agrupamentos e centralidade. A visualização da rede pela árvore geradora mínima e a construção do escalonamento multidimensional por métricas baseada em grafos e sua posterior visualização bidimensional mostra-se uma possibilidade interessante para o estudo de redes. Verificamos que este ilustrou satisfatoriamente a estrutura do mercado acionário brasileiro, evidenciando uma dimensão tradicional do mercado, o risco setorial. Ainda, foi mostrado como a medida de centralidade baseada em autovetores se aplica ao mercado como medida de risco geral de dependência que uma ação deve ter em relação às demais que estão ligadas àquela. À este conceito, demos o nome de 'caminhos de volatilidade'.

Alguns autores ponderam a existência de um pequeno número de fatores direcionando as dinâmicas estocásticas dos retornos dos ativos no mercado financeiro [Mantegna and Stanley 1999, Bonanno et al. 2001]. Assim, a abordagem via escalonamento multidimensional, a partir das estatísticas da rede de correlação de ativos parece ser boa alternativa à visualização de tais dimensões ou fatores.

Concluindo, a proposta aqui apresentada precisa de estudos mais profundos para



Figura 3. Representação do EMD baseado em árvore geradora mínima (AGM). Os nós foram evidenciados a fim de não se sobrepossem uns aos outros. A medida de *stress* S é igual a 0,049. Neste caso, como $S > 0$, a utilização do EMD sobre a AGM ainda conseguiu simplificar a visualização de informação nos dados, além do que seria obtido somente via AGM. As cores representam os grupos da AGM.

validar de forma mais precisa os resultados obtidos pela análise de centralidade, como a correlação entre centralidade e o retorno das ações, bem como estudar possíveis dimensões encontradas a partir do uso do escalonamento multidimensional. Também faz-se interessante o uso de outras medidas de dependência, particularmente medidas que incluam dependências não lineares, a fim de encontrar outros padrões que possam caracterizar o comportamento do mercado financeiro. Estes e outros temas estão em fase de desenvolvimento por estes autores. Por fim, esperamos que este trabalho contribua para que pesquisadores, investidores e gestores possam usar tais métodos para auxiliar na tomada de decisão de investimentos, dessa forma, interligando características estatísticas da rede com fatores econômico-financeiros das ações.

Agradecimentos: Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro.

Referências

- Barabasi, A.-L. (2016). *Network Science*. Cambridge University Press, Cambridge CB2 8BS, United Kingdom.
- Bonanno, G., Caldarelli, G., Lillo, F., Micciche, S., Vandewalle, N., and Mantegna, R. N. (2004). Networks of equities in financial markets. *The European Physical Journal B - Condensed Matter*, 38(2):363–371.
- Bonanno, G., Lillo, F., and Mantegna, R. N. (2001). Levels of complexity in financial markets. *Physica A: Statistical Mechanics and its Applications*, 299(1-2):16–27.
- Brida, J. G. and Risso, W. A. (2008). Multidimensional minimal spanning tree: The dow jones case. *Physica A: Statistical Mechanics and its Applications*, 387(21):5205–5210.
- Coelho, R., Gilmore, C. G., Lucey, B., Richmond, P., and Hutzler, S. (2007). The evolution of interdependence in world equity markets from minimum spanning trees. *Physica A: Statistical Mechanics and its Applications*, 376:455–466.
- Cormen, T., Leiserson, C., Rivest, R., and Stein, C. (2001). *Introduction to Algorithms, Second Edition*. Cambridge, Massachusetts London, England.
- Di Matteo, T., Pozzi, F., and Aste, T. (2010). The use of dynamical networks to detect the hierarchical organization of financial market sectors. *Eur. Phys. J. B*, 73(1):3–11.
- Dias, J. (2013). Spanning trees and the eurozone crisis. *Physica A: Statistical Mechanics and its Applications*, 392(23):5974–5984.
- Dionisio, A., Menezes, R., and Mendes, D. A. (2004). Mutual information: a measure of dependency for nonlinear time series. *Physica A: Statistical Mechanics and its Applications*, 344(1-2):326–329.
- Fiedor, P. (2014). Networks in financial markets based on the mutual information rate. *Phys Rev E Stat Nonlin Soft Matter Phys*, 89(5):052801.
- Gilmore, C. G., Lucey, B. M., and Boscia, M. W. (2010). Comovements in government bond markets: A minimum spanning tree analysis. *Physica A: Statistical Mechanics and its Applications*, 389(21):4875–4886.
- Jacomy, M., Venturini, T., Heymann, S., and Bastian, M. (2014). Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. *PLoS ONE*, 9(6):e98679.
- Kruskal, J. B. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem. *Proc. Amer. Math. Soc.*, 7(1):48–48.
- Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a non-metric hypothesis. *Psychometrika*, 29(1):1–27.
- Manly, B. (2004). *Multivariate Statistical Methods: A Primer, Third Edition*. Taylor & Francis.
- Mantegna, R. (1999). Hierarchical structure in financial markets. *Eur. Phys. J. B*, 11(1):193–197.

- Mantegna, R. N. and Stanley, H. E. (1999). *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge University Press, Cambridge.
- Miccichè, S., Bonanno, G., Lillo, F., and N. Mantegna, R. (2003). Degree stability of a minimum spanning tree of price return and volatility. *Physica A: Statistical Mechanics and its Applications*, 324(1-2):66–73.
- Newman, M. (2010). *Networks*. Oxford University Press.
- Sensoy, A. and Tabak, B. M. (2014). Dynamic spanning trees in stock market networks: The case of asia-pacific. *Physica A: Statistical Mechanics and its Applications*, 414:387–402.
- Tumminello, M., Lillo, F., and Mantegna, R. N. (2010). Correlation, hierarchies, and networks in financial markets. *J Econ Behav Organ*, 75(1):40–58.
- Yang, C., Zhu, X., Li, Q., Chen, Y., and Deng, Q. (2014). Research on the evolution of stock correlation based on maximal spanning trees. *Physica A: Statistical Mechanics and its Applications*, 415:1–18.