



UNIVERSIDADE DE BRASÍLIA
Faculdade de Economia, Administração e Contabilidade - FACE

Análise Topológica de Redes Financeiras Embaralhadas

Mateus Hiro Nagata

Brasília, Maio de 2021

Mateus Hiro Nagata

Análise Topológica de Redes Financeiras Embaralhadas

Monografia apresentada à Faculdade de Economia, Contabilidade e Administração (FACE) como requisito à obtenção do grau de Bacharel em Ciências Econômicas.

Universidade de Brasília - UnB

Orientador: Daniel Oliveira Cajueiro

Brasília

Maio de 2021

Mateus Hiro Nagata

Análise Topológica de Redes Financeiras Embaralhadas

Monografia apresentada à Faculdade de Economia, Contabilidade e Administração (FACE) como requisito à obtenção do grau de Bacharel em Ciências Econômicas.

Trabalho aprovado. Brasília, maio de 2021:

Daniel Oliveira Cajueiro
Orientador

Bernardo Pinheiro Machado Mueller
Convidado 1

Brasília
Maio de 2021

*À minha família que se estende mais do que os laços de sangue.
Àqueles os quais os os sorrisos são mais importantes que os tesouros que ajuntei nessa
terra.*

Agradecimentos

Imensa gratidão àqueles que fazem parte da minha vida e são responsáveis por manter em pé esse homem que depende mais da sorte e do destino que o seu egocentrismo consegue ver. Abaixo escrevo humildemente aos meus amigos que merecem mais que um parágrafo numa folha de papel.

À primeira dama Kim Juyoung, responsável por colorir não só as redes usadas documento, mas toda a minha vida com cores em tom pastel, por comemorar cada vitória minha com flores secas e por me dar força ao acordar e alegria ao dormir.

Ao orientador professor Cajueiro, o qual me agraciou com comentários perspicazes e objetivos, além de calorosas conversas acadêmicas noturnas.

Ao professor Raul por me ensinar que a vida é estocástica e os sistemas estão fora de equilíbrio, mas que acima de tudo existe um motivo maior por trás de tudo isso.

À professora Andrea que transformou esse ingênuo garoto num economista com ensinamentos precisos, exemplos diversos e vivência que não só das quatro-paredes da universidade e, por conseguinte, cada membro do PET. Assim como aos calorosos Economigos que me acompanharam na seriedade da universidade, e nas conversas mais lúdicas.

Ao LAMFO, responsável por ampliar vastos horizontes que existem entre o mundo dos números binários, dos agentes racionais e dos mais agradáveis sábados matutinos.

À igreja que me considera um filho e a qual atribuo cada parcela de quem eu sou. Cada minuto com vocês valeram a pena e espero retribuir de igual tamanho.

Aos Esquilos que, pelo seu *nonsense* hilário e conversas descontraídas, trazem doçura aos dias mais amargos.

在臺灣見的各位.包括昌賢, 一夫, 瑞琪, 艾超, 健維, 佳任, 鴻明, 主安, 志豪, 永平, 佳琪, 國展, 馨庭, 昀昕, 育良, 林加, 秀玲.

따뜻하고 활기찬 김가족을 위해. 영백상, 선미상, 하영누나, 주영누나, 소영누나, 준하.

祖父祖母に幼いころから守り養ってくれたて感謝。多々なミスも許し、限らないやさしさ、長年の付き合いで教えてくれた人生の知恵をありがたく思います。

Aos meus pais e irmãs que aturaram e criaram-me por 22 anos e por fim começa a jornada acadêmica. Agradeço pela convivência que permitiu amadurecimento que incluiu gargalhadas e conversas sérias, chás da tarde e doces de fim de semana.

Finalmente, ao Pai Celestial o qual é dono de todo aplauso. Que seja esse trabalho o começo de vários outros que atribuem glória àquele que merece.

*“A soberba precede a ruína,
altivez do espírito precede a queda.”
(Provérbios 16:18)*

Resumo

Aplico o algoritmo small-shuffle surrogate (SSS) nas ações do S&P 500 e estudo as séries de tempo embaralhadas como uma grande rede conectada, não-ponderada e não-direcionada com técnicas de “stock network correlation”. Analiso, então, as propriedades topológicas de lei de potência, eficiência, robustez topológica.

Os resultados indicam que a rede é, de fato, eficiente e livre de escala, ainda frágil topologicamente contra ataques direcionados, mas robusta contra erros aleatórios; isso tudo é conhecido da literatura.

A análise indica que existe uma dinâmica de curto prazo no movimento de preço das ações, o qual é responsável pelo movimento correlacionado de ações do mesmo setor e quando anulado essa dinâmica, os setores perdem a sua aglutinação. Todavia, as características topológicas se mantêm independente dessa dinâmica, estipulando que o mercado acionário tem fundamentos que não se limitam às flutuações de curto prazo.

Palavras-chaves: Small-Shuffle Surrogate, Stock Network Correlation, Redes, Leis de Potência, Eficiência, Robustez Topológica.

Abstract

I apply a small-shuffle surrogate (SSS) algorithm on the S&P 500 stocks and study the shuffled time-series as a big undirected, unweighted and connected network with stock network correlation techniques. Then, I analyze the topological properties such as power law, efficiency, and topological robustness.

The outcome shows that, in fact, the network is efficient and scale-free, yet topologically fragile against directed attacks and topologically robust against random errors; which is all known in the literature.

The analysis indicates the existence of a short-term dynamic in the price movements of the stocks, which are responsible for the correlated movement of the stocks of the same sector, and when this dynamic is canceled, the sectors lose agglutination. However, the topological characteristics remain the same regardless of that dynamic, stipulating that the stock market has some fundamentals that is not limited to short-term fluctuations.

Keywords: Small-Shuffle Surrogate, Stock Network Correlation, Networks, Power Law, Efficiency, Topological Robustness.

1 Introdução

Nesse artigo, represento as ações do índice americano S&P 500 como uma grande rede para analisar suas propriedades topológicas. Em seguida, aplicamos o algoritmo small-shuffle surrogate (SSS), de forma a quebrar a estrutura local e capturar tendências de longo prazo (NAKAMURA; SMALL, 2005a). Então analiso a mudança nas características da rede, tais como a lei de potência, a eficiência global, a robustez e o limiar de quebra da rede de acordo com a escala do embaralhamento do algoritmo SSS. O objetivo é analisar quais características são mantidas quando as flutuações de curto-prazo são ignoradas e se considera apenas as relações de longo prazo dos dados. Surpreendentemente, O resultado é que as características topológicas não se alteram à medida que a amplitude varia, mas apenas o agrupamento dos setores das ações é alterado.

O mercado financeiro (especificamente aqueles resultantes do algoritmo MST) é tipicamente uma rede com interação de vários participantes que agem de maneira independente e não-linear (BOCCALETTI et al., 2006; GAO; SMALL; KURTHS, 2017; CALDARELLI et al., 2004; YAN; XIE; WANG, 2014). Grande parte da literatura é responsável por estudar as propriedades topológicas dessa rede e auferir sentido econômico e compreender as estruturas (TABAK; SERRA; CAJUEIRO, 2010; TABAK et al., 2014; CHAKRABORTI et al., 2013). Alguns trabalhos recentes focam no quanto essa estrutura é mantida ou destruída em momentos de quebra-estrutural e de grandes choques (ZHANG; ZHUANG, 2019; YAN; XIE; WANG, 2014; NOBI et al., 2014; SENSOY; TABAK, 2014; HEIBERGER, 2014).

Para selecionar as conexões com maior relevância entre as ações, implemento o minimal spanning tree (MST) (MANTEGNA, 1999). Nesse algoritmo, das $N(N - 1)/2$ conexões originais, selecionamos apenas as $N - 1$ conexões tais que minimizam a soma das distâncias entre os nós, fornecendo um grafo que é eficiente em termos de processamento e passagem de informação (LATORA; MARCHIORI, 2003). Uma proposta alternativa é o planar maximally filtered graph (PMFG) apresentado por (TUMMINELLO et al., 2005), semelhante no intuito, mas que retém $3(N - 2)$ conexões, ou seja, mais informação do que o MST. Escolhemos o MST para o nosso estudo pois, além de ser o modelo tradicional nessa literatura (MANTEGNA, 1999), não temos nenhuma motivação de estudar redes mais complexas, então a escolha desse algoritmo é uma consequência direta da Navalha de Ockham (JEFFERYS; BERGER, 1992).

A análise de correlações entre ativos financeiros em forma de rede é conhecido na literatura como “stock network correlation” e a vantagem desse método é simplificar a estrutura de correlação entre as ações. Como essa configuração permite encontrar quais são

as ações que são mais influenciadas diretamente enquanto separa as que sofrerão apenas efeitos indiretos, a presença de clusters e hubs fica aparente graficamente.

Dessa análise gráfica, não só é possível destilar o comportamento dos grupos, mas também é fácil de identificar quais são os ativos fora do grupo que podem transmitir um grande choque no grupo. Além disso, permite estudar características topológicas que podem ser usadas para explicar o comportamento do mercado acionário quanto à fragilidades sistêmicas e ao processamento de informação.

A capacidade de processamento de informação de forma não custosa é capturado pela medida de eficiência do grafo. Um grafo eficiente tem a característica de “small-world”. O termo é normalmente associado ao experimento de (MILGRAM, 1967) o qual propõe que as relações sociais das pessoas formam uma rede com comprimento médio pequeno. No caso de Milgram, seis passos de distância. Podemos pensar no nosso caso que é o quanto a rede é sensível à mudança de preço de um ativo.

Outra característica de interesse é a resistência que a rede tem contra erros e ataques. Redes livres de escala são extremamente robustos contra erros aleatórios. Albert, Jeong e Barabási (2000) dizem que características topológicas como diâmetro da rede costumam ser estáveis mesmo em frente a um número elevado de erros (aleatórios), sendo que mesmo com a falha de 5% dos nós, a comunicação da rede costuma não ter problemas. No caso de análise de ações, a interpretação é a dependência que reside em um ativo.

O motivo para tal resiliência, ainda de acordo com Albert, Jeong e Barabási (2000) é a heterogeneidade entre as conexões. A lei de potência indica que alguns poucos nós tem extrema conectividade enquanto muitos nós tem conectividade insignificante e como esses são numerosos, há maior chance de esses serem escolhidos em remoção aleatória. Dessa forma, a retirada desses não impacta de modo significativa a topologia da rede.

Por outro lado, a rede é extremamente vulnerável contra ataques direcionados à nós centrais. Isso se deve à presença de poucos nós super conectados, enquanto a maioria dos nós tem pouquíssimas conexões. Mencionado anteriormente, “livre de escala” quer dizer que não há um número padrão para a quantidade de graus do nó que esperamos observar ao escolher um elemento aleatório da rede. Muitos nós tem pouquíssimas conexões, vários desvios-padrão abaixo da média, enquanto alguns nós são super conectados, puxando a média para um patamar muito mais elevado que a maioria dos nós da rede (BARABÁSI; BONABEAU, 2003; GAO; SMALL; KURTHS, 2017; DONG et al., 2013). A definição de rede livre de escala é que os graus da rede, ou seja a quantidade de arestas o qual um nó está conectado, é distribuído de acordo com uma lei de potência.

A contribuição desse documento é especificamente combinar a análise de redes e de grafos complexos à técnica small-shuffle surrogate. No capítulo 2 detalho a motivação por trás do uso de um método de embaralhamento nos dados e descrevo os dados que

usarei no trabalho. No capítulo 3, discuto a importância de cada propriedade topológica estudada. Os resultados empíricos estão no capítulo 4 e concluo no capítulo 5.

2 Metodologia

Uma forma de aferir a relação entre dois ativos é a correlação e essa estatística consegue capturar semelhanças em tendências presentes em longos períodos de tempo. Todavia, mesmo quando sinais não são similares, é possível que haja uma estrutura de correlação entre elas (ou seja, eles estão conectados de alguma forma). Dessa forma, o movimento deles não seriam aleatórios e independentes, mas coordenados por uma mesma força ou são influenciados um pelo outro (NAKAMURA; HIRATA; SMALL, 2006).

Há motivos plausíveis para acreditar que a dinâmica do mercado financeiro é determinístico, uma vez que toda compra e toda venda é motivada por um preço. Ademais, Nakamura e Small (2006) citam Plerou et al. (1999) que diz que no mercado de câmbio, os agentes vislumbram até 3 minutos anteriores para fazer a negociação (PLEROU et al., 1999) e que existe uma janela temporal de 10 minutos que separa dinâmicas de puro ruído (PLEROU et al., 2002). Portanto, há indícios de um certo determinismo nas decisões tomadas no mercado financeiro.

Há também motivos plausíveis para acreditar que o mercado financeiro é estocástico. O grande número de transações e de agentes tornam o sistema complexo e de alta-dimensão de tal forma que o sistema pode ser considerado puramente aleatório (NAKAMURA; SMALL, 2005b).

Nakamura (2006) advoca que existem certas dinâmicas determinísticas que contém informações valiosas para a compra e venda de ativos, ou seja, o movimento não é puramente estocástico.

Para avaliar a existência ou não de tais dinâmicas, o método random shuffle surrogate (RSS) foi proposto por (THEILER et al., 1992) que é apropriado para séries temporais exclusivamente compostos por flutuações irregulares, enquanto o método small-shuffle surrogate (SSS) proposto por (NAKAMURA; SMALL, 2005a) consegue lidar com dados com tendências e periodicidades também.

2.1 Small-Shuffle Surrogate

O método small-shuffle surrogate (SSS) pode ser aplicado para dados com tendências de longo-prazo, pois preserva esse tipo de dinâmica, enquanto destrói estruturas locais de correlação (NAKAMURA; HIRATA; SMALL, 2006). Como o nome sugere, a ideia é embaralhar a ordem dos dados da série de tempo em escala “pequena”. O método permite checar se as flutuações irregulares são ou não aleatórias.

Aplico o método SSS nos log-retornos. Seja o dado original $x(t)$ e seja $i(t)$ o índice

de $x(t)$. Ou seja, $i(t) = t$ e $x(i(t)) = x(t)$. Seja $g(t)$ uma variável aleatória gaussiana $\mathcal{N}(0, 1)$. Com essa variável vamos perturbar o índice $i(t)$ e atualizar com o índice perturbado. O procedimento é o seguinte:

1. Obtenha o número embaralhado $i'(t) = i(t) + Ag(t)$, sendo A a amplitude. Dessa forma, $i(t)$ será uma sequência de números inteiros, enquanto $i'(t)$ será esse mesmo número somado de uma variável aleatória real;
2. Ordene as variáveis pelo $i'(t)$ gerado. Vamos chamar de $i'(t)$ a ordem pós-perturbação;
3. Obtenha a série de tempo embaralhada $s(t) = x(i'(t))$.

O método não depende da distribuição dos dados para sua eficácia. De fato, a nova série de tempo tem a mesma distribuição dos dados originais.

O intento desse método é manter certas propriedades de interesse e destruir outras, de forma a ser consistente com dada hipótese nula (NAKAMURA; SMALL, 2006). Além disso, a nova série deve ser suficientemente semelhante à original. Se há uma diferença significativa entre elas, a hipótese nula é rejeitada. A premissa é que se os dados não são aleatórios, existe um sistema dinâmico oculto que gera as tendências e as periodicidades que contaminam os dados. Na presença de tal dinâmica, correlações de curtos períodos de tempo são presentes e então a ordem o qual os dados aparecem são profundamente importantes. Quando o índice então é perturbado, os dados não refletem mais a dinâmica original. O objetivo do método é distinguir flutuações aleatórias com ou sem dinâmica (NAKAMURA; SMALL, 2005a).

Os autores mostram que a manutenção de componentes de longo prazo e destruição de componentes de curto prazo dependem da escolha de uma amplitude A adequada. Em muitos casos a amplitude necessária é $A = 1.0$ (NAKAMURA; HIRATA; SMALL, 2006). A contribuição desse trabalho é averiguar o quanto as estruturas de rede do mercado financeiro são preservadas com o tamanho da amplitude. Dessa forma, faremos um grid de $A = 0.0, 0.1, \dots, 2.0$ e ver o quanto a rede se modifica ao longo do grid.

O benefício prático disso é claro. Ao analisarmos uma rede em que as estruturas locais são destruídas, retendo apenas tendências de largos períodos de tempo, o investidor de longo prazo se beneficiaria em considerar as flutuações de curtíssimo prazo como ruído e assim, analisar as correlações que perduram num horizonte temporal mais longo.

Até o valor de $A = 2.0$, os dados se comportam semelhantemente aos dados sem embaralhamento e à medida que A cresce, os dados se tornam cada vez mais estocásticos. Todavia, se os dados contém tendências, então a amplitude não deve ser alta, pois nesse caso a influência das tendências contaminam as flutuações irregulares. Na maioria dos casos, $A = 1.0$ parece apropriado. Quando $A < 1.0$ a função de autocorrelação e a média

da informação mútua não captam diferença entre os dados originais e os embaralhados. Quando $A = 1.0$, cerca de 50% dos dados estão em suas posições originais (NAKAMURA; SMALL, 2005a).

Dada a importância do índice em séries determinísticas (NAKAMURA; SMALL, 2006), seria plausível que na presença de tal dinâmica, as redes resultantes do embaralhamento possam ter características muito diferentes do processo original e é isso que queremos analisar.

2.2 Dados

Uso nesse estudo as ações do índice S&P 500, do período de 2015-01-02 até 2021-03-26, totalizando 1568 dias úteis. Descarto as ações do índice que não existem durante todo o período. A variável de interesse é o log-retorno diário r do ativo i no dia t ,

$$r_i(t) = \ln(p_i(t+1)) - \ln(p_i(t)). \quad (2.1)$$

Calculo, então, a correlação entre dois ativos dado pelo coeficiente de correlação de Pearson,

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}. \quad (2.2)$$

Para construir a matriz de distâncias, usaremos o seguinte conceito de distância comum na literatura

$$d_{X,Y} = \sqrt{2(1 - \rho_{X,Y})}. \quad (2.3)$$

Claramente, a definição atende os critérios de métrica (distância):

1. $d(x, y) = 0 \Leftrightarrow x = y$ (nula para pontos coincidentes);
2. $d(x, y) = d(y, x)$ (simetria);
3. $d(x, z) \geq d(x, y) + d(y, z)$ (desigualdade triangular);
4. $d(x, y) \geq 0$ (positivamente definida).

Como todas as distâncias são não-nulas, é necessário extrair as informações relevantes para formar o grafo das nossas 488 ações. O algoritmo minimum spanning tree (MST) escolhe apenas $N - 1$ arestas das $N(N - 1)/2$ tais que tornam a rede totalmente conexa e que minimizam a soma dos pesos das arestas. O algoritmo de Prim é utilizado para encontrar o MST (TABAK; SERRA; CAJUEIRO, 2010).

O método é interessante pois retém as conexões mais relevantes. De acordo com Mantegna (MANTEGNA, 1999), a escolha dessa topologia permite isolar grupos com sentido econômico apenas pelos dados de preço, ou seja, com apenas um tipo de dado, é possível auferir estruturas topológicas e conexões importantes entre as ações. Outra propriedade interessante é que, se as correlações entre todas as ações seja pelo menos ligeiramente diferente, o que é bastante razoável, o MST resultante é único.

No final temos um grafo não-direcionado e não-ponderado, cujas distâncias entre um nó e outro são medidos por meio de passos necessários para ir desse nó até o outro e cujas arestas retém apenas as conexões mais importantes, sendo assim, um formato bastante conciso de apresentar como as ações estão correlacionadas.

3 Propriedades

3.1 Distribuição

Redes que aparecem na literatura de finanças são tipicamente livres de escala (LILLO; MANTEGNA, 2003; BONANNO et al., 2004; KIM et al., 2002a; KIM et al., 2002b; TABAK et al., 2014). Isso implica a presença de jogadores importantes, com graus que excedem bastante a quantidade média amostral. Em outras palavras, alguns nós tem uma influência mister na estrutura da rede. Lin e Li (2010) argumenta que redes livres de escala são eficientes em termos de transferência de conhecimento.

Essa propriedade, no nosso contexto de redes de ações conectadas por correlação e filtrados pelo MST, implica que o movimento de uma ação com influência mais importante na rede pode ser um fator importante em dirigir o otimismo ou o pessimismo no mercado de ações.

O efeito dessa propriedade na escolha de portfólio é grande. Se há indícios que algumas ações são as mais importantes na questão de propagação de euforia no mercado, então essa ação consegue explicar grande parte da variação dos retornos de um portfólio. Essa realidade é mais perceptível pela escolha do nosso algoritmo de filtro MST. Como o algoritmo seleciona as conexões mais relevantes, então sabemos que as ações conectadas diretamente a esse ativo com maior relevância serão diretamente e mais fortemente impactados por um choque nos seus preços. Dito de outra forma, um portfólio composto pela ação mais relevante e as ações conectadas a ela tem uma espécie de “risco sistêmico local” elevada e a escolha de portfólio deve levar isso em consideração.

No outro extremo, não se pode dizer que redes aleatórias (com escala) são ótimas em questão de risco. Uma queda abrupta de ações escolhidas aleatoriamente causam menos impacto na rede livre de escala, que é resiliente contra erros, do que uma em uma rede puramente aleatória. Do ponto de vista de minimização de risco, a consciência da distribuição de graus da rede permite a escolha de uma estratégia de diversificação de riscos. Caso a lei de potência pareça razoável, o portfólio deve priorizar ações que estão afastados do ponto de vista da rede e, especificamente, mais passos longe de uma ação com grande influência. Dessa forma, tendo baixo risco sistêmico global (pela resiliência da rede contra choques aleatórios) e baixo risco sistêmico local (pela distância existente entre a ação escolhida e a ação influente).

Redes de ações filtradas por MST são, notadamente, livres de escala (CHI; LIU; LAU, 2010; YAN; XIE; WANG, 2014; BONANNO et al., 2003; VANDEWALLE; BRISBOIS; TORDOIR, 2001; ONNELA et al., 2003). Como o algoritmo SSS permite estudar

características que permanecem no longo prazo e anulam as de curto prazo, a contribuição desse artigo permite dizer se a estratégia de diversificação mencionada no parágrafo anterior faz sentido se analisadas somente as relações de longo prazo.

Uma rede livre de escala é uma tal que os graus dos nós podem ser descritos por uma lei de potência

$$P(k) \propto k^{-\alpha}. \quad (3.1)$$

A interpretação da lei de potência é o seguinte: imagine que o α encontrado seja 2, então a probabilidade de um nó ter a metade de arestas de um outro nó é aproximadamente é quatro vezes maior (BARABÁSI; BONABEAU, 2003). Dessa forma é simples imaginar que poucos nós tem grau elevado enquanto muitos tem grau diminuto.

A presença de leis de potência podem indicar que eventos raros de grande escala ocorrem com probabilidade não negligenciável (LILLO; MANTEGNA, 2003). Para o nosso caso, não só indica que poucas empresas são responsáveis por grande parte da influência na rede, mas que existe um tipo de mecanismo que torna vantajoso se conectar com tal ação, ou que seja próprio do sistema que o movimento das ações seja mais atrelado com algumas ações líderes, num mecanismo de “preferential attachment”.

3.2 Eficiência global

A eficiência global aponta para que a rede seja uma “small-world network” (WATTS; STROGATZ, 1998). Esse tipo de rede vem da propriedade ilustrada por Stanley Milgram (1967) que indica a possibilidade transmissão de informação informação (cartas, no caso de Milgram) de uma extremidade à outra da rede com pouco esforço.

Especificamente, redes com essa propriedade são altamente clusterizados, mas com caminhos médios pequenos (LATORA; MARCHIORI, 2001). Tornando-se, portanto, eficiente tanto em escala global quanto local (LATORA; MARCHIORI, 2003).

Uma rede eficiente, assim como indica a presença de lei de potência, aponta para o fato da rede ser resistente contra erros aleatórios, mas ser frágil contra ataques focalizados nos polos importantes (PRIESTER; SCHMITT; PEIXOTO, 2014). Por outro lado, a interconectividade pode ser o meio pelo qual contágio, doenças e recessão podem infectar a rede como um todo (GEORG, 2010). Uma rede “small-world” representa um sistema paralelo (LATORA; MARCHIORI, 2003).

Vale lembrar, todavia, que o contágio depende também de outros parâmetros e que é um debate que não deve ser feito considerando apenas a eficiência como o único gatilho (LÓPEZ-PINTADO, 2006).

Seja L_{ij} a menor quantidade de passos que existe entre a ação i e a ação j , a eficiência global é definida da seguinte forma

$$E = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{L_{ij}}. \quad (3.2)$$

Essa medida, então, é a média da soma dos inverso das distâncias no grafo. Seria a média, pois conta com cada combinação de nós possíveis $N(N-1)$. Uma medida de eficiência parecida seria o tamanho característico do caminho $L = \frac{1}{N(N-1)} \sum_{i \neq j} L_{ij}$, mas a definição teria problemas caso o grafo se tornasse não-conectado.

Veja que, por definição, nossa rede não poderá ser 100% eficiente, uma vez que a única rede que teria essa característica seria uma rede 100% conectada, o que não é possível, uma vez que filtramos com o MST. Dessa forma, não faz sentido comparar redes que passaram por filtros diferentes. Como vou comparar as mesmas ações, no mesmo período e com o mesmo filtro, apenas com o embaralhamento diferente, a comparação pode ser feita diretamente.

A eficiência tem implicações na escolha de portfólios assim como a lei de potência. Uma rede comparativamente mais eficiente é aquecida e desaquecida de forma mais rápida e há mais propensão dos preços das ações moverem-se de forma conjunta. Dessa forma, analisar a eficiência pode ser útil para analisar o market-timing e identificar momentos que o mercado está mais para “bull” ou mais para “bear” e as estratégias de portfólio devem incorporar tais considerações.

3.3 Estabilidade

A estabilidade, ou robustez topológica, indica o quanto o grafo é resistente à mudanças estruturais ou topológicas. Mensura-se a partir do limiar de colapso da remoção de nós (YAN; XIE; WANG, 2014). Podemos fazer essa remoção de forma aleatória ou escolhendo especificamente o nó mais conectado.

O método apresentado por (YAN; XIE; WANG, 2014) e proposto por (ALBERT; JEONG; BARABÁSI, 2000) é medir certas propriedades topológicas da rede à medida que retiramos ações da rede.

A robustez topológica indica a dependência que a rede tem em relação a certo ativo financeiro. Se a remoção de um ativo não traz muita variação na topologia da rede, então um choque nos preços dessa ação não deve alterar muito os retornos esperados de um portfólio composto por ações dessa rede. Todavia, a remoção de um ativo que destrói a rede topologicamente indica que uma grande parte dos ativos é afetada pelo choque, implicando num prejuízo abrangente. Como faremos a remoção tanto aleatoriamente quanto direcionadamente, poderemos comparar a robustez contra erros e contra ataques.

4 Resultados Empíricos

4.1 Análise de clusters

Nas Figuras 1, 2, 3 abaixo, apresento o MST gerado com SSS de Amplitude $A = 0.0$, $A = 1.0$, $A = 2.0$. Veja que no primeiro caso, as ações se aglomeram fortemente com ações do mesmo setor. No segundo caso, os nós começam a se dispersar, mas os clusters dos setores ainda é presente. No último caso, o agrupamento não é mais aparente. Vale lembrar que de acordo com (NAKAMURA; SMALL, 2005b), a amplitude adequada seria $A = 1.0$. Dessa forma, a Figura 2 reflete as tendências de longo prazo.

Veja que a estrutura do grafo não tem mudanças significativas e parece indicar que é o mesmo processo gerador que construiu as três redes. Para amplitudes diferentes, veja o Apêndice A. Contudo, pequenas mudanças principalmente na composição dos setores parece ter mudado. Há uma maior quantidade de setores que se dividiram e estão agrupados em galhos separados da rede, além de haver uma variedade intersetorial mais acentuada. Os clusters que continham apenas ações de um tipo de setor são menos frequentes.

4.2 Lei de potência

Quanto a lei de potência, vemos que não há relação entre o grau de embaralhamento da série de tempo e sua lei de potência. Os resultados podem ser vistos na Figura 4. Veja que, ao menos graficamente, a presença de uma lei de potência coerente com os nossos dados. Os detalhes estatísticos dessa análise podem ser vistos no Apêndice B.

Temos que o $\hat{\alpha}$ encontrado é, em média, 1.476. De acordo com a análise, é razoável pensar que os dados vieram de fato por uma distribuição de lei de potência, ou uma variação exponencialmente truncada, o que condiz com a presença de alguns nós com influência fenomenal e muitos nós com pouca influência. De qualquer forma, a distribuição dos graus parece a mesma independente da amplitude A usada para gerar os dados.

O $\hat{\alpha}$ encontrado indica que a distribuição não tem uma média e nem uma variância bem definida (NEWMAN, 2005), pois para o primeiro momento, seria necessário $\alpha > 2$ e para o segundo momento, seria necessário $\alpha > 3$.

4.3 Eficiência

A eficiência encontrada se encontra na Figura 5. Veja que, embora a eficiência oscile um pouco, não parece haver nenhuma relação explícita entre a amplitude e a eficiência.

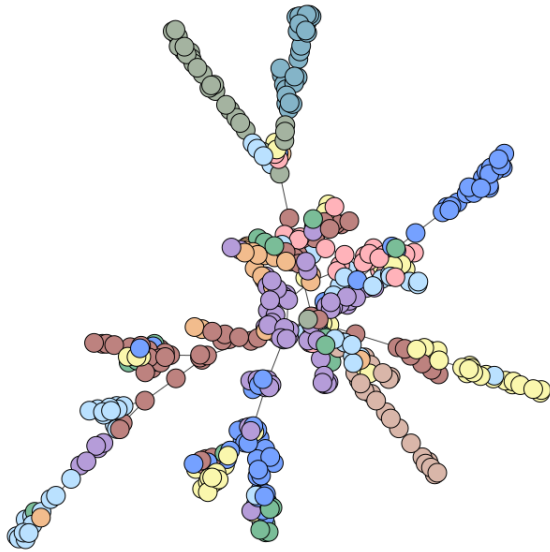


Figura 1 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 0.0$.

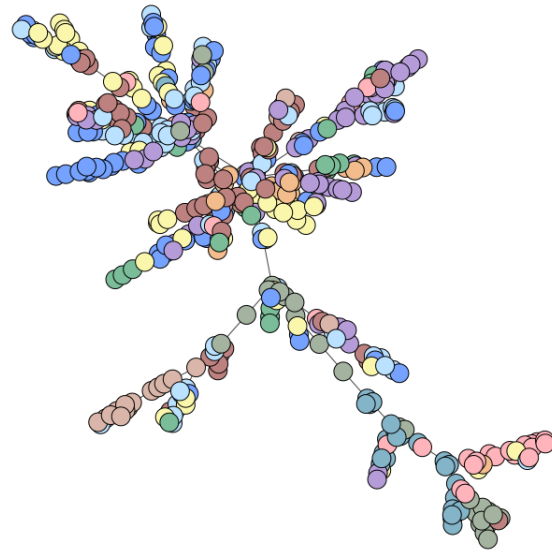


Figura 2 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 1.0$.

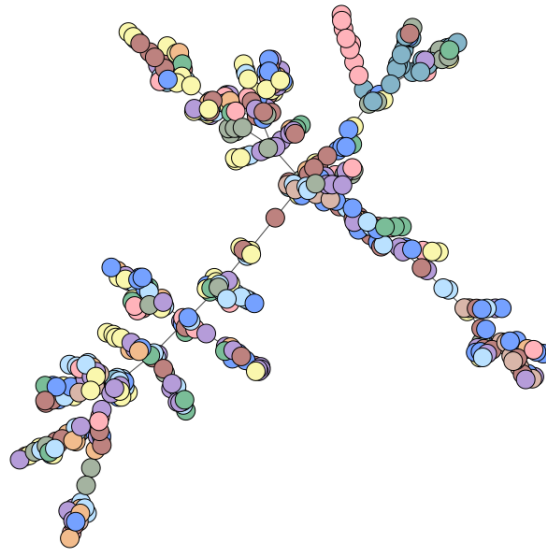


Figura 3 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 2.0$.

Para interpretar a eficiência, uma alternativa é analisar o tamanho médio dos caminhos na Figura 6.

O tamanho médio dos caminhos pode ser interpretado exatamente como a eficiência, mas nesse caso é somado as distâncias entre todas as combinações possíveis de nós e dividido por $N(N - 1)$, enquanto a eficiência é calculada por inversas das distâncias. Não surpreendentemente, a mudança na eficiência se traduz numa mudança no tamanho característico na direção contrária. Dessa forma, podemos ver que nos primeiros casos, a

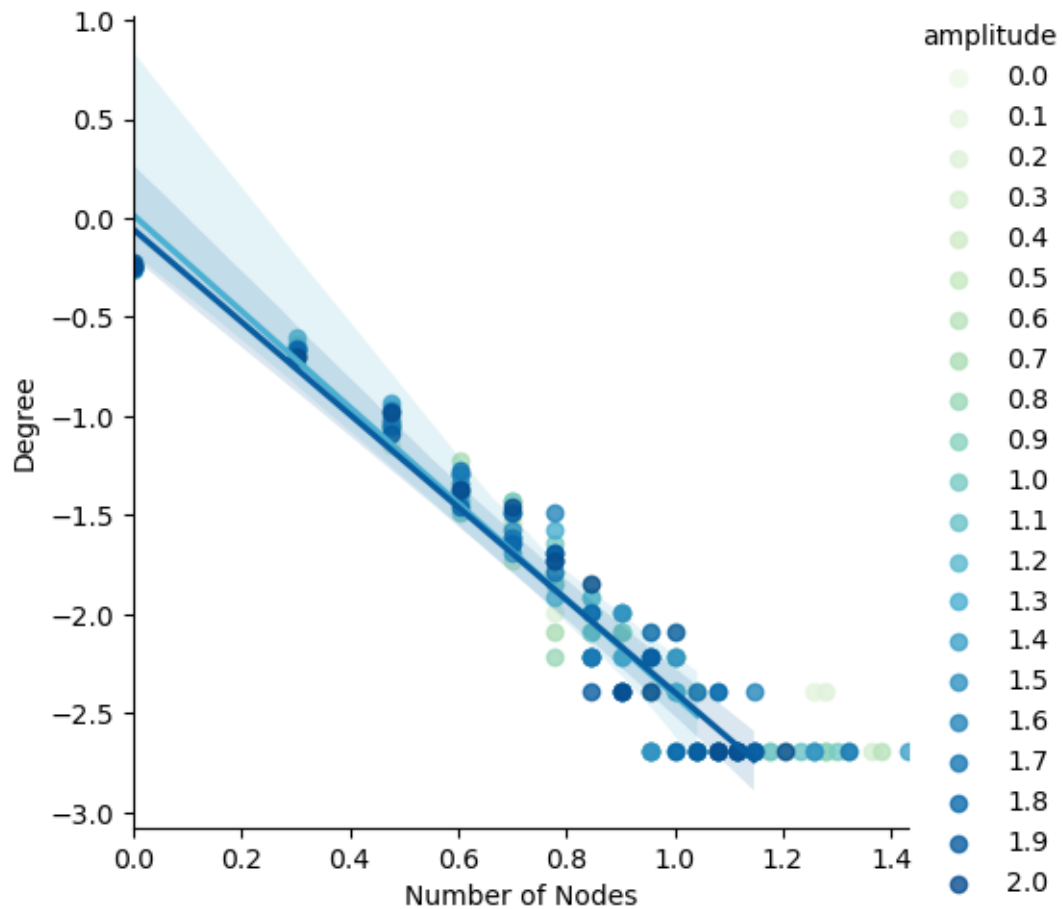


Figura 4 – Lei de potência para o índice S&P 500 para amplitudes diferentes.

média dos caminhos da rede eram cerca de 10 passos, ou seja, um “small-world network”. Analiso o fenômeno que ocorre entre as amplitudes 0.4 e 0.7 na subseção com detalhes no Apêndice A.

4.4 Robustez topológica

Para analisar a robustez topológica, calculamos o tamanho relativo da rede à medida que removemos uma quantidade de nós. A Figura 7 é o resultado do procedimento quando retiramos o nó mais conectado, enquanto a Figura 8 é o caso da remoção de nós aleatoriamente.

Veja que, exatamente como propõe a literatura, a rede é extremamente frágil contra ataques direcionados, mas resiliente para erros aleatórios (BARABÁSI; BONABEAU, 2003; ALBERT; JEONG; BARABÁSI, 2000; YAN; XIE; WANG, 2014), em alguns casos, para a remoção de 5% das arestas. A característica permanece independente da amplitude usada para gerar a rede.

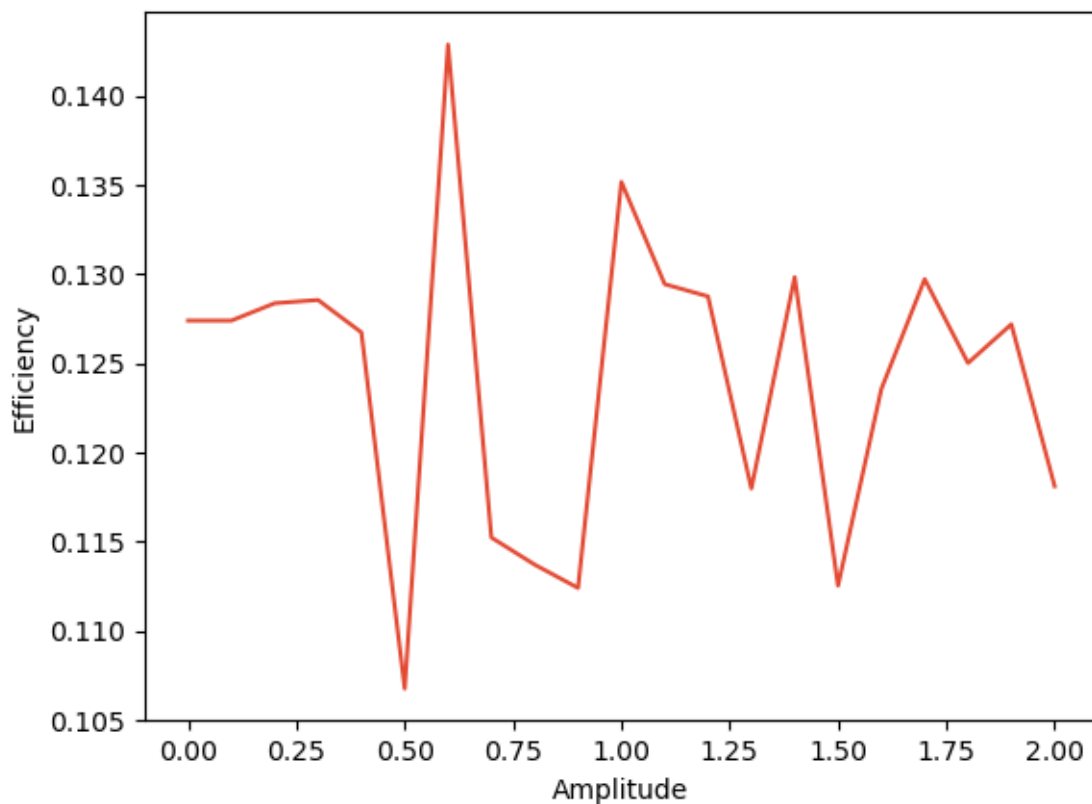


Figura 5 – Eficiência para amplitudes diferentes

4.5 Discussão dos Resultados

Os resultados são que, apesar do embaralhamento dos índices das ações, a estrutura da rede de ações é ainda caracterizada pelos fenômenos “small-world” e livres de escala. Além disso, não há motivos para crer que a robustez topológica é diferente para qualquer amplitude, tanto para a remoção aleatória e a remoção direcionada.

O embaralhamento preserva tendências de longo prazo e periodicidades. A preservação da eficiência, da robustez, da lei de potência mostra que, mesmo no longo prazo, os fundamentos pelos quais algumas ações são mais correlacionadas com outras e a presença de poucos nós de grande importância permanecem e não são frutos apenas de flutuações de curto prazo ou ruído.

Dessa forma, a única alteração significativa parece ser os setores se tornarem menos aglutinados à medida que a amplitude aumenta. Isso pode indicar que o setor não é um bom preditor para tendências de longo prazo e outros aspectos se tornam mais importantes. Agora, quais seriam esses aspectos é uma incógnita e uma pergunta interessante para ser explorada.

A ocorrência de uma rede pouco eficiente e pouco conectada na amplitude $A = 0.5$

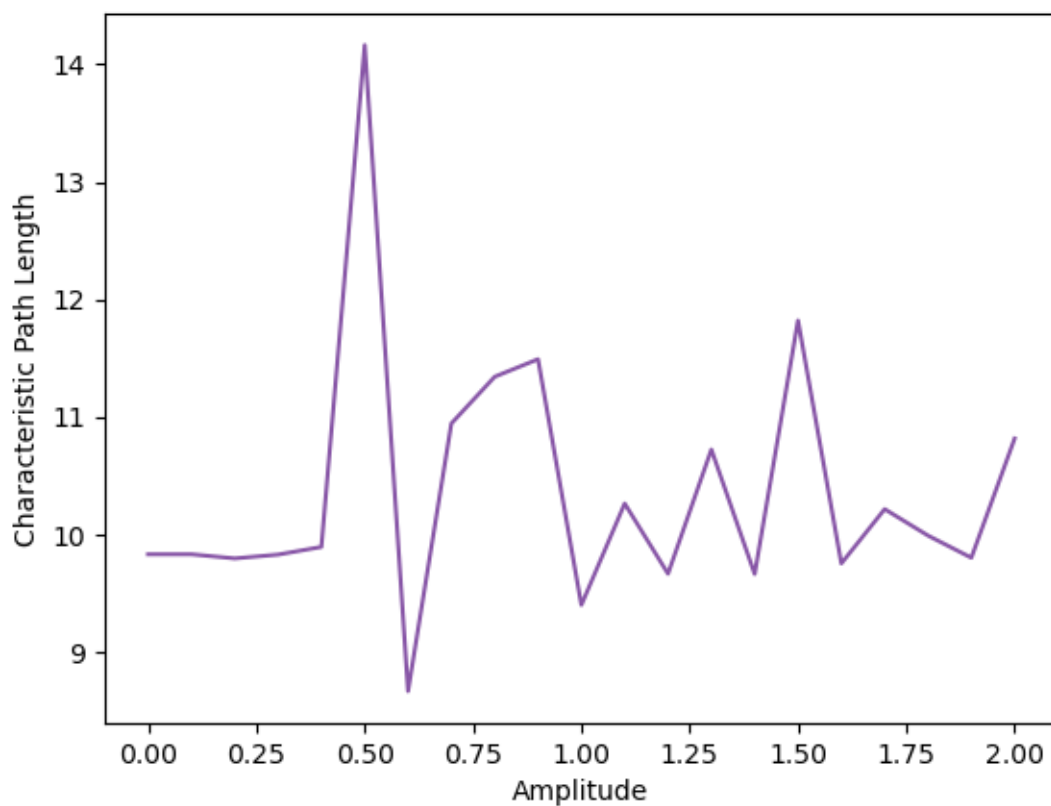


Figura 6 – Tamanho médio dos caminhos

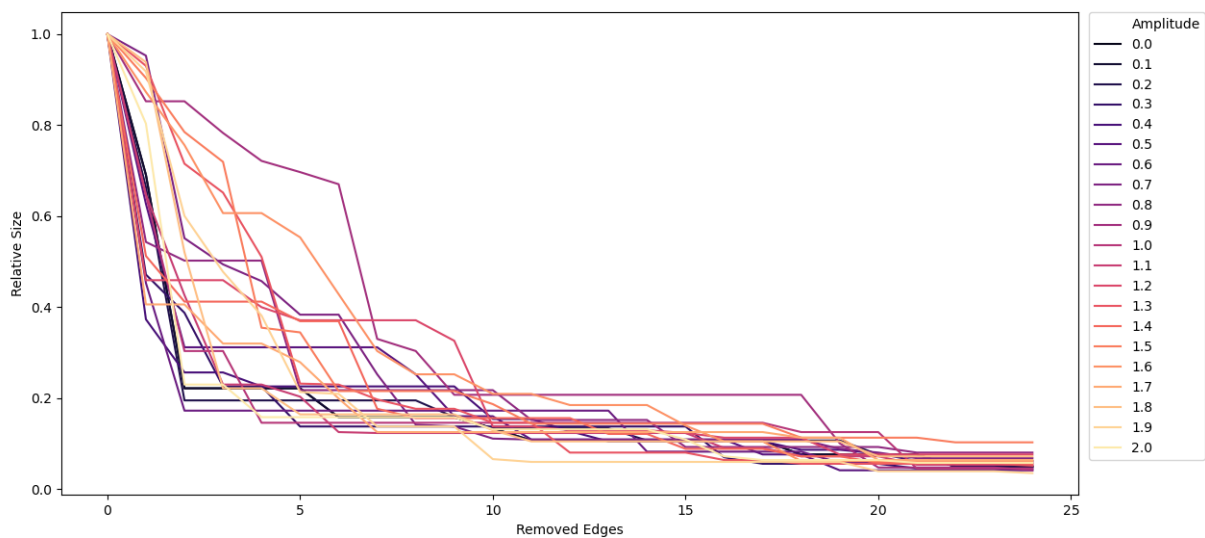


Figura 7 – Remoção das arestas mais conectadas

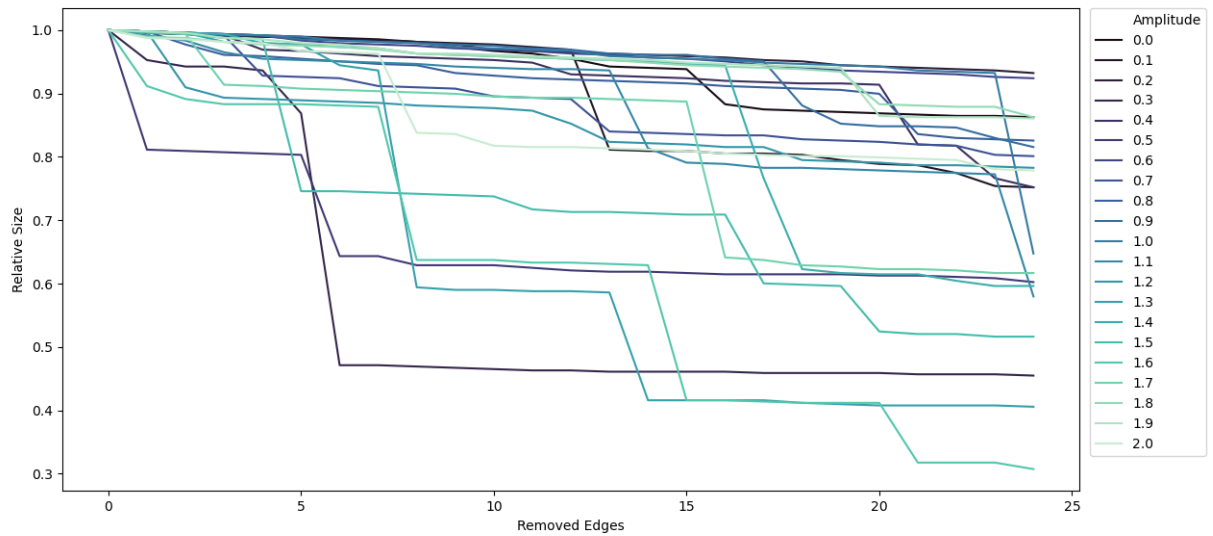


Figura 8 – Remoção aleatória das arestas

é um achado interessante e seria necessário averiguar se foi uma coincidência aleatória ou se existe algum mecanismo para a formação de redes daquela forma.

Podemos concluir que, independente de qual seja o processo gerador dos dados, ele permanece o mesmo para diferentes amplitudes e que há, de fato, uma dinâmica determinística que dirige o movimento das ações e que ela é responsável por manter a unidade do movimento de preços de ações pertencentes ao mesmo setor.

5 Conclusões e Trabalhos Futuros

Concluimos então que apesar do SSS embaralhar os dados de forma a quebrar estruturas locais de curto-prazo, e preservar tendências de longo-prazo, os aspectos topológicos se mantêm.

Outro fato interessante é que os setores são mais agrupados quando SSS tem amplitude menor, ou seja, a interconectividade dos setores é muito grande quando incluído flutuações de curto prazo. Esse efeito vai esvaecendo à medida que a amplitude aumenta, mostrando que o movimento conjunto das ações é menos interligado no longo prazo e indica a presença de dinâmicas determinísticas de curto prazo para o movimento de ações do mesmo setor. Além disso, quando a amplitude chega a níveis maiores, tais como $A = 2.0$, já não há uma clara relação entre o comportamento das ações.

Futuramente, pretendo testar algumas questões que não puderam ser devidamente respondidas pelo documento. A primeira é se a abordagem feita no documento é útil para seleção de portfólio para minimização de riscos no longo prazo, uma vez que o SSS deve refletir tendências e ignorar movimentos de curto prazo. A segunda, seria criar uma métrica que permite auferir quantitativamente a alteração das aglutinações intersetoriais das ações. A terceira é testar as propriedades com timestamps menores e ver que tipo de dinâmica pode estar presente em períodos de tempo menores. A quarta seria avaliar as propriedades mas levando em conta períodos de crise, pré-crise e períodos estáveis.

Referências

- ALBERT, R.; JEONG, H.; BARABÁSI, A.-L. Error and attack tolerance of complex networks. *nature*, Nature Publishing Group, v. 406, n. 6794, p. 378–382, 2000. Citado 4 vezes nas páginas 2, 11, 15 e 33.
- ALSTOTT, J.; BULLMORE, E.; PLENZ, D. powerlaw: a python package for analysis of heavy-tailed distributions. *PloS one*, Public Library of Science, v. 9, n. 1, p. e85777, 2014. Citado 2 vezes nas páginas 33 e 34.
- BARABÁSI, A.-L.; BONABEAU, E. Scale-free networks. *Scientific american*, JSTOR, v. 288, n. 5, p. 60–69, 2003. Citado 3 vezes nas páginas 2, 10 e 15.
- BOCCALETTI, S. et al. Complex networks: Structure and dynamics. *Physics reports*, Elsevier, v. 424, n. 4-5, p. 175–308, 2006. Citado na página 1.
- BONANNO, G. 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. Citado na página 9.
- BONANNO, G. et al. Networks of equities in financial markets. *The European Physical Journal B*, Springer, v. 38, n. 2, p. 363–371, 2004. Citado na página 9.
- CALDARELLI, G. et al. Emergence of complexity in financial networks. In: *Complex Networks*. [S.l.]: Springer, 2004. p. 399–423. Citado na página 1.
- CHAKRABORTI, A. et al. Econophysics. physicists’ approaches to a few economic problems. *J. Econ. Inter. Coord*, 2013. Citado na página 1.
- CHI, K. T.; LIU, J.; LAU, F. C. A network perspective of the stock market. *Journal of Empirical Finance*, Elsevier, v. 17, n. 4, p. 659–667, 2010. Citado na página 9.
- CLAUSET, A.; SHALIZI, C. R.; NEWMAN, M. E. Power-law distributions in empirical data. *SIAM review*, SIAM, v. 51, n. 4, p. 661–703, 2009. Citado na página 33.
- DONG, G. et al. Robustness of network of networks under targeted attack. *Physical Review E*, APS, v. 87, n. 5, p. 052804, 2013. Citado na página 2.
- GAO, Z.-K.; SMALL, M.; KURTHS, J. Complex network analysis of time series. *EPL (Europhysics Letters)*, IOP Publishing, v. 116, n. 5, p. 50001, 2017. Citado 2 vezes nas páginas 1 e 2.
- GEORG, C.-P. *The effect of the interbank network structure on contagion and financial stability*. [S.l.], 2010. Citado na página 10.
- HEIBERGER, R. H. Stock network stability in times of crisis. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 393, p. 376–381, 2014. Citado na página 1.
- JEFFERYS, W. H.; BERGER, J. O. Ockham’s razor and bayesian analysis. *American Scientist*, JSTOR, v. 80, n. 1, p. 64–72, 1992. Citado na página 1.
- KIM, H. et al. Scale-free network in stock markets. *Journal-Korean Physical Society*, Citeseer, v. 40, p. 1105–1108, 2002. Citado na página 9.

- KIM, H.-J. et al. Weighted scale-free network in financial correlations. *Journal of the Physical Society of Japan*, The Physical Society of Japan, v. 71, n. 9, p. 2133–2136, 2002. Citado na página 9.
- LATORA, V.; MARCHIORI, M. Efficient behavior of small-world networks. *Physical review letters*, APS, v. 87, n. 19, p. 198701, 2001. Citado na página 10.
- LATORA, V.; MARCHIORI, M. Economic small-world behavior in weighted networks. *The European Physical Journal B-Condensed Matter and Complex Systems*, Springer, v. 32, n. 2, p. 249–263, 2003. Citado 2 vezes nas páginas 1 e 10.
- LILLO, F.; MANTEGNA, R. N. Power-law relaxation in a complex system: Omori law after a financial market crash. *Physical Review E*, APS, v. 68, n. 1, p. 016119, 2003. Citado 3 vezes nas páginas 9, 10 e 33.
- LIN, M.; LI, N. Scale-free network provides an optimal pattern for knowledge transfer. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 389, n. 3, p. 473–480, 2010. Citado na página 9.
- LÓPEZ-PINTADO, D. Contagion and coordination in random networks. *International Journal of Game Theory*, Springer, v. 34, n. 3, p. 371–381, 2006. Citado na página 10.
- MANTEGNA, R. 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. Citado 2 vezes nas páginas 1 e 8.
- MILGRAM, S. Six degrees of separation. *Psychology Today*, v. 2, p. 60–64, 1967. Citado 2 vezes nas páginas 2 e 10.
- NAKAMURA, T.; HIRATA, Y.; SMALL, M. Testing for correlation structures in short-term variabilities with long-term trends of multivariate time series. *Physical Review E*, APS, v. 74, n. 4, p. 041114, 2006. Citado 2 vezes nas páginas 5 e 6.
- NAKAMURA, T.; SMALL, M. Small-shuffle surrogate data: Testing for dynamics in fluctuating data with trends. *Physical Review E*, APS, v. 72, n. 5, p. 056216, 2005. Citado 4 vezes nas páginas 1, 5, 6 e 7.
- NAKAMURA, T.; SMALL, M. Small-shuffle surrogate data: Testing for dynamics in fluctuating data with trends. *Physical Review E*, APS, v. 72, n. 5, p. 056216, 2005. Citado 2 vezes nas páginas 5 e 13.
- NAKAMURA, T.; SMALL, M. Testing for dynamics in the irregular fluctuations of financial data. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 366, p. 377–386, 2006. Citado 3 vezes nas páginas 5, 6 e 7.
- NEWMAN, M. E. Power laws, pareto distributions and zipf's law. *Contemporary physics*, Taylor & Francis, v. 46, n. 5, p. 323–351, 2005. Citado na página 13.
- NOBI, A. et al. Effects of global financial crisis on network structure in a local stock market. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 407, p. 135–143, 2014. Citado na página 1.
- ONNELA, J.-P. et al. Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E*, APS, v. 68, n. 5, p. 056110, 2003. Citado na página 9.

- PLEROU, V. et al. Universal and nonuniversal properties of cross correlations in financial time series. *Physical review letters*, APS, v. 83, n. 7, p. 1471, 1999. Citado na página 5.
- PLEROU, V. et al. Random matrix approach to cross correlations in financial data. *Physical Review E*, APS, v. 65, n. 6, p. 066126, 2002. Citado na página 5.
- PRIESTER, C.; SCHMITT, S.; PEIXOTO, T. P. Limits and trade-offs of topological network robustness. *PloS one*, Public Library of Science, v. 9, n. 9, p. e108215, 2014. Citado na página 10.
- SENSOY, A.; TABAK, B. M. Dynamic spanning trees in stock market networks: The case of asia-pacific. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 414, p. 387–402, 2014. Citado na página 1.
- TABAK, B. M.; SERRA, T. R.; CAJUEIRO, D. O. 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. Citado 2 vezes nas páginas 1 e 7.
- TABAK, B. M. et al. Directed clustering coefficient as a measure of systemic risk in complex banking networks. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 394, p. 211–216, 2014. Citado 2 vezes nas páginas 1 e 9.
- THEILER, J. et al. Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena*, Elsevier, v. 58, n. 1-4, p. 77–94, 1992. Citado na página 5.
- TUMMINELLO, M. et al. A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 102, n. 30, p. 10421–10426, 2005. Citado na página 1.
- VANDEWALLE, N.; BRISBOIS, F.; TORDOIR, X. Non-random topology of stock markets. *Quantitative finance*, Taylor & Francis, v. 1, p. 372–374, 2001. Citado na página 9.
- WATTS, D. J.; STROGATZ, S. H. Collective dynamics of ‘small-world’ networks. *nature*, Nature Publishing Group, v. 393, n. 6684, p. 440–442, 1998. Citado na página 10.
- YAN, X.-G.; XIE, C.; WANG, G.-J. The stability of financial market networks. *EPL (Europhysics Letters)*, IOP Publishing, v. 107, n. 4, p. 48002, 2014. Citado 4 vezes nas páginas 1, 9, 11 e 15.
- ZHANG, W.; ZHUANG, X. The stability of Chinese stock network and its mechanism. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V., v. 515, p. 748–761, 2019. ISSN 03784371. Disponível em: <<https://doi.org/10.1016/j.physa.2018.09.140>>. Citado na página 1.

Apêndices

APÊNDICE A – Imagens dos MST de acordo com a amplitude

Abaixo apresento todas as redes geradas pelo SSS. Vamos analisar com cuidado a vizinhança da amplitude $A = 0.5$, uma vez que é o ponto onde a eficiência tomou seu menor valor.

As redes geradas pelo SSS entre as amplitudes 0.4 e 0.7 estão representados nas figuras 13, 14, 15, 15. Note que as redes construídas costumam ter dois eixos grandes que criam caminhos bem largos e esses eixos tem vários pequenos galhos. Uma configuração nesse formato, de fato, garante que a maior parte dos nós sejam bastante conectados.

A figura 14 é um outlier interessante, pois nesse ponto um desses eixos apresentou quase nenhum galho, então as ações deixaram de estar densamente conectados. Esse é um fenômeno interessante e seu fator gerador é incerto.

Veja que à medida que a amplitude aumenta, menos conectados estão as ações do mesmo setor, mas não há mudanças explícitas nos formatos das redes que permitem dizer que um mecanismo gerador diferente foi o responsável por criar as redes com maior embaralhamento.

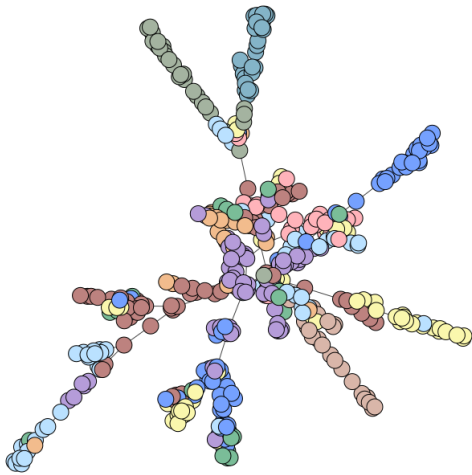


Figura 9 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 0.0$.

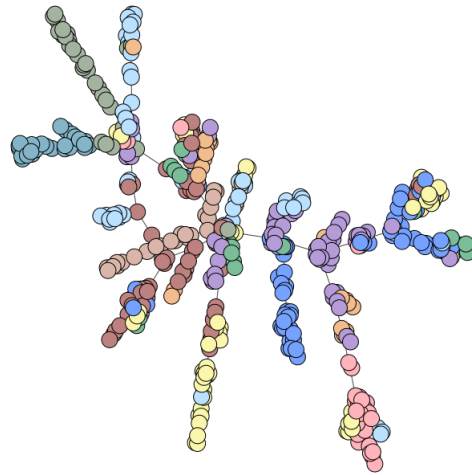


Figura 10 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 0.1$.

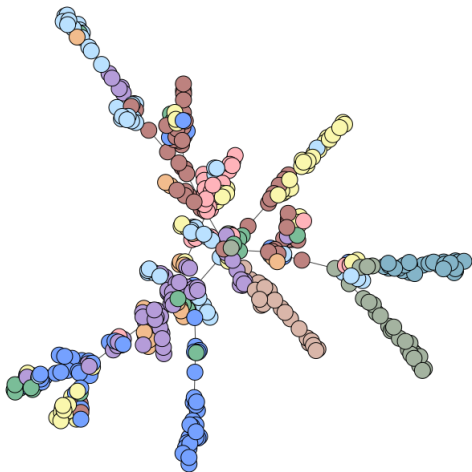


Figura 11 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 0.2$.

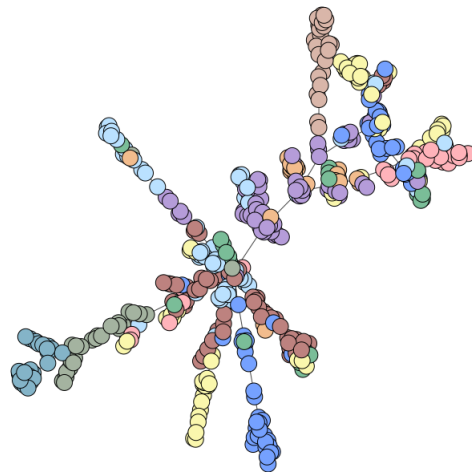


Figura 12 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 0.3$.

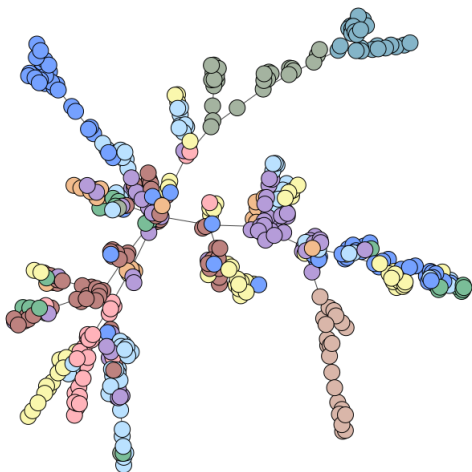


Figura 13 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 0.4$.

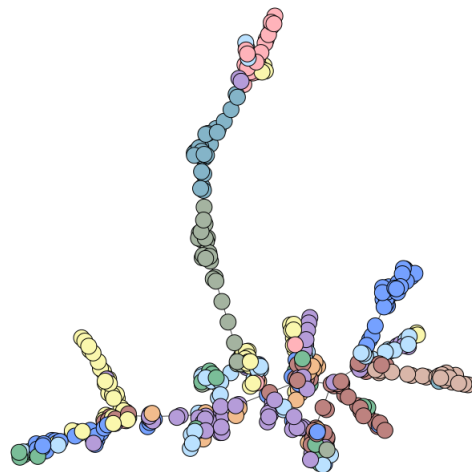


Figura 14 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 0.5$.

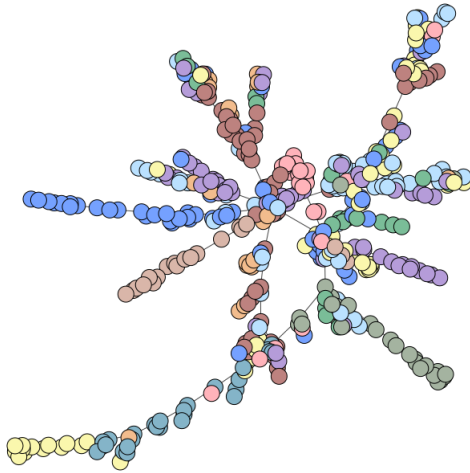


Figura 15 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 0.6$.

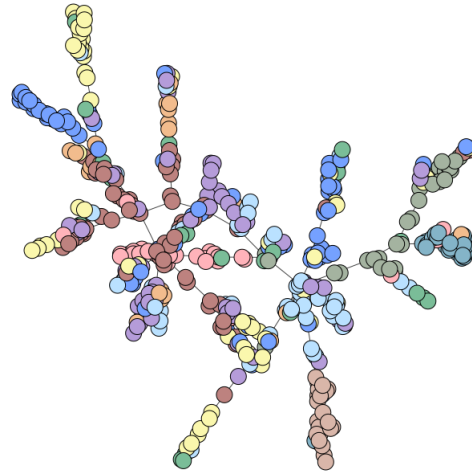


Figura 16 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 0.7$.

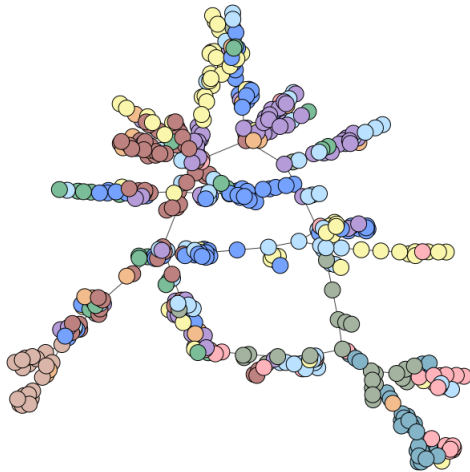


Figura 17 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 0.8$.

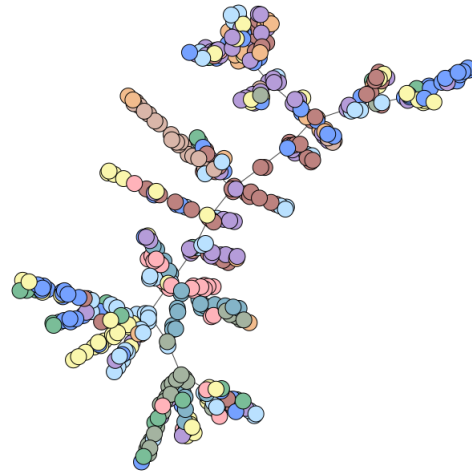


Figura 18 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 0.9$.

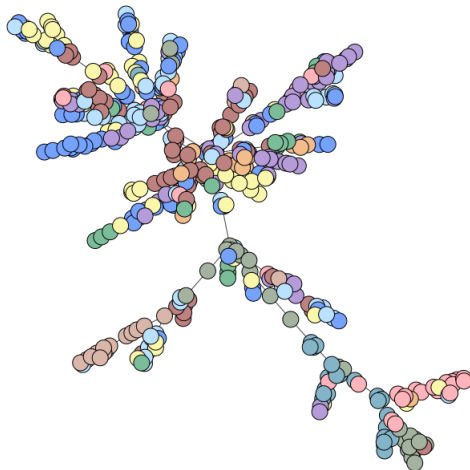


Figura 19 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 1.0$.

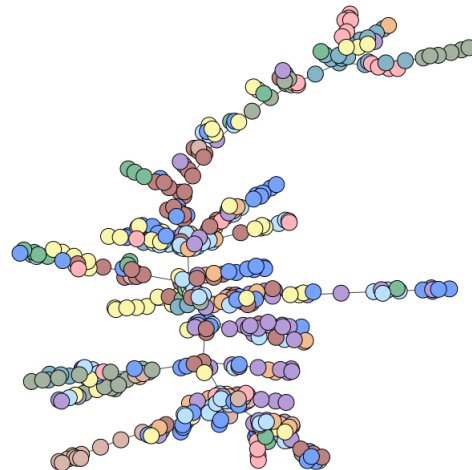


Figura 20 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 1.1$.

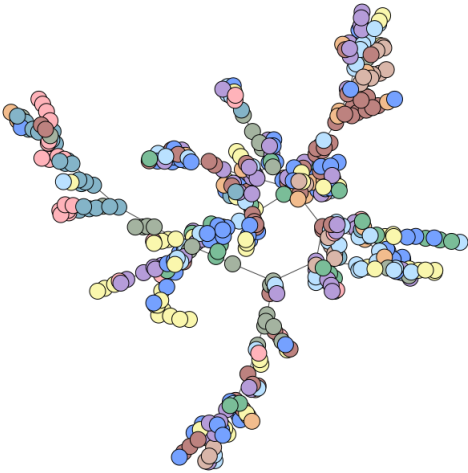


Figura 21 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 1.2$.

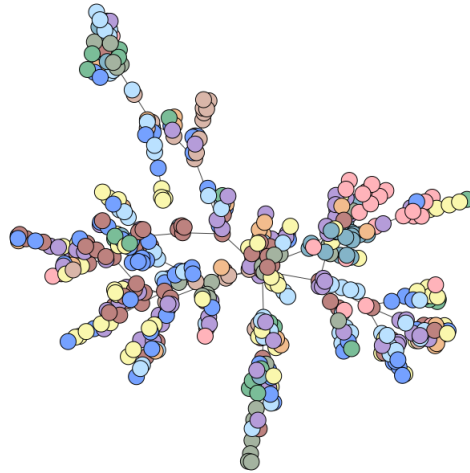


Figura 22 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 1.3$.

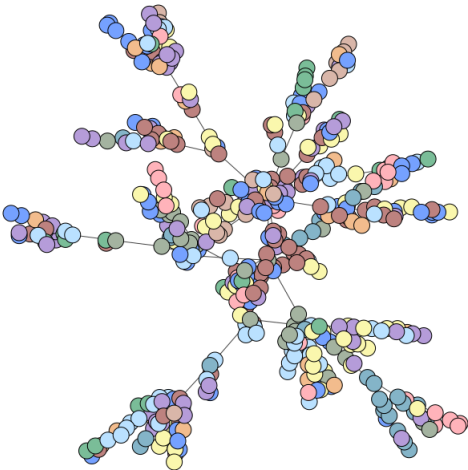


Figura 23 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 1.4$.

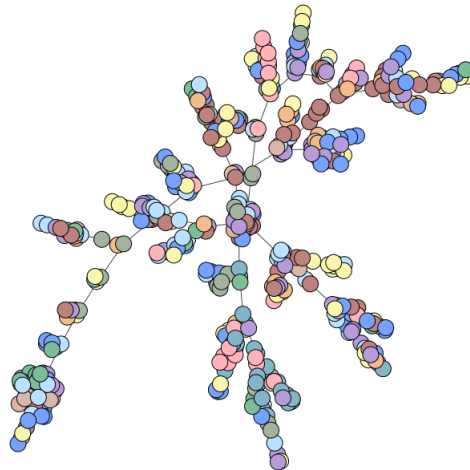


Figura 24 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 1.5$.

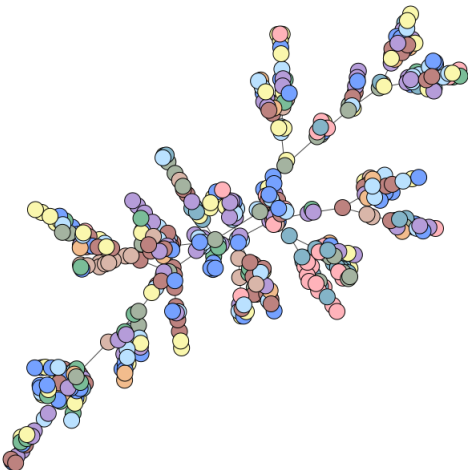


Figura 25 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 1.6$.

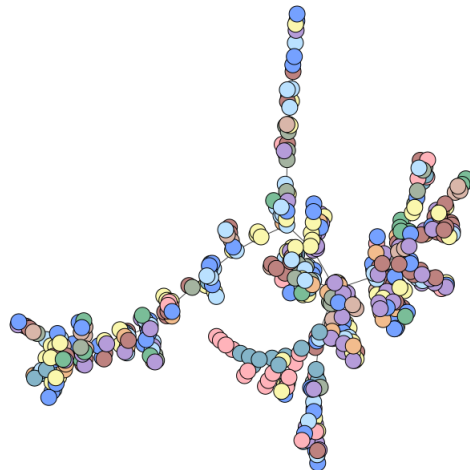


Figura 26 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 1.7$.

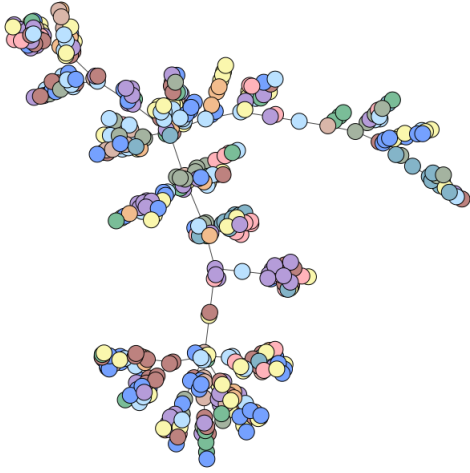


Figura 27 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 1.8$.

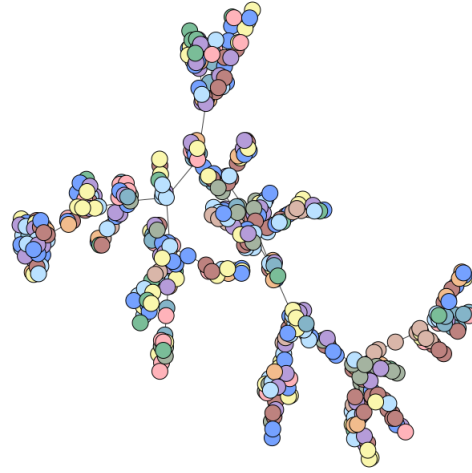


Figura 28 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 1.9$.

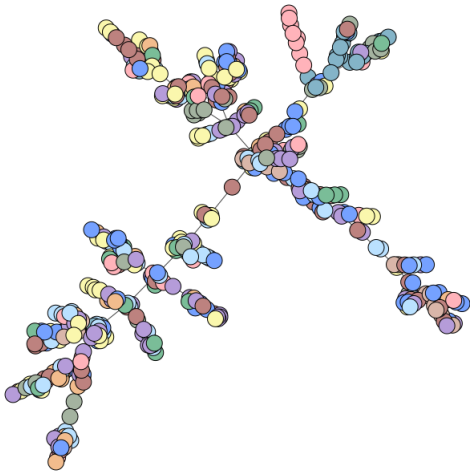


Figura 29 – MST das ações do S&P 500 colorida por setor. O algoritmo SSS aplicado tem amplitude $A = 2.0$.

APÊNDICE B – Lei de Potência

A qualidade do ajuste da lei de potência é analisada nessa seção. A Tabela 1 apresenta estatísticas relacionadas à esse ajuste, tais como o $\hat{\alpha}$ encontrado pelo seu estimador de máxima verossimilhança e o seu desvio-padrão $\hat{\sigma}$, a distância de Kolmogorov-Smirnov D entre os dados e o modelo ajustado e, por fim, a log-razão de verossimilhança entre a lei de potência e um modelo alternativo R e o p-valor em favor da lei de potência p .

A log-razão de verossimilhança entre a lei de potência e o modelo alternativo apresenta número positivo se a distribuição de lei de potência for mais verossímil que a alternativa e negativo caso contrário. O p-valor p indica a significância dessa hipótese. Na tabela, os p-valores significativos em 10% são indicados por asteriscos simples (*) e em 5% são indicados por asteriscos duplos (**).

De acordo com (CLAUSET; SHALIZI; NEWMAN, 2009), se observa um descuido estatístico na literatura que é a conclusão da presença de uma lei de potência apenas pelo gráfico log-log. Na verdade, é difícil poder afirmar que dados se originaram de uma distribuição de lei de potência e o máximo que se pode afirmar é que os dados são consistentes com essa distribuição. Mostramos nesse apêndice que esse é, de fato, o caso dos dados aqui representados.

A análise da distribuição de lei de potência deve ser comparada com o ajuste com outras curvas que podem explicar o modelo. O primeiro modelo alternativo a ser analisado é o exponencial. Uma distribuição de cauda pesada é, por definição, uma distribuição que não é exponencialmente limitada e, dessa forma, para que seja factível dizer que os dados seguem uma lei de potência, é necessário garantir que os dados não seguem uma distribuição exponencial (ALSTOTT; BULLMORE; PLENZ, 2014).

Como as distribuições alternativas tem 2 parâmetros enquanto a lei de potência tem só 1, não frequentemente, os ajustes são maiores para esses casos, por uma questão de overfitting (ALSTOTT; BULLMORE; PLENZ, 2014), então não necessariamente a lei de potência será a distribuição com maior log-razão de verossimilhança. Importa apenas ser consistente com a possibilidade de haver uma distribuição que é resiliente a choques, embora sendo frágil a ataques (ALBERT; JEONG; BARABÁSI, 2000) e a presença de eventos extremos (LILLO; MANTEGNA, 2003).

Os dados na Tabela 1 indicam que a lei de potência é mais verossímil que a exponencial (exp) em todas as amplitudes e é significativa para algumas. A stretched exponential - Weibull (stretch) parece mais verossímil para os dados que a lei de potência, embora em nenhuma amplitude o resultado seja significativo.

A comparação com o modelo de lei de potência e a exponentially truncated power law (trunc) é interessante, pois os dados se ajustam melhor à versão truncada que a lei de potência original. Isso pode fazer sentido pois a escolha de apenas 488 ações da S&P 500 é arbitrária e não há um motivo para que tenha um corte nesse número. Uma forma mais “natural” de escolher as ações poderia ser uma alternativa para esse problema.

A distribuição log-normal pode ser gerada com uma variação do Teorema do Limite Central, mas para variáveis aleatórias positivas que são multiplicadas, e não somadas. Quando um mecanismo do tipo é plausível, a distribuição log-normal tende, de fato, a ter um ajuste melhor que a lei de potência (ALSTOTT; BULLMORE; PLENZ, 2014). Os dados mostram que esse é o caso dos dados. No entanto, essa diferença não é significativa para quase nenhuma amplitude.

<i>Amplitude</i>	α	σ	D	R_{exp}	p_{exp}	$R_{stretch}$	$p_{stretch}$	R_{trunc}	p_{trunc}	R_{logn}	p_{logn}
0	1.492	0.148	0.098	1.579	0.114	-0.643	0.521	-0.927	0.252	-0.617	0.537
1	1.492	0.148	0.098	1.579	0.114	-0.643	0.521	-0.927	0.252	-0.617	0.537
2	1.627	0.209	0.144	1.476	0.140	-0.170	0.865	-0.552	0.465	-0.221	0.825
3	1.397	0.106	0.070	2.065	0.039**	-1.025	0.305	-1.382	0.166	-0.986	0.324
4	1.417	0.108	0.108	2.234	0.025**	-1.158	0.247	-1.374	0.212	-1.266	0.205
5	1.393	0.105	0.081	2.025	0.043**	-1.104	0.270	-1.429	0.153	-1.029	0.303
6	1.640	0.242	0.141	0.725	0.468	-0.485	0.627	-0.690	0.345	-0.460	0.645
7	1.447	0.141	0.163	1.002	0.316	-0.780	0.436	-1.027	0.153	-0.756	0.450
8	1.377	0.105	0.108	1.801	0.072*	-1.047	0.295	-1.371	0.127	-1.018	0.309
9	1.604	0.214	0.124	0.988	0.323	-0.432	0.666	-0.688	0.341	-0.415	0.678
10	1.356	0.103	0.144	1.377	0.168	-1.110	0.267	-1.373	0.091*	-1.050	0.294
11	1.416	0.111	0.102	2.331	0.020**	-1.392	0.164	-1.280	0.242	-1.980	0.048**
12	1.569	0.180	0.126	1.527	0.127	-0.332	0.740	-0.676	0.366	-0.338	0.735
13	1.805	0.329	0.158	0.644	0.520	-0.247	0.805	-0.473	0.494	-0.237	0.812
14	1.358	0.103	0.118	1.454	0.146	-1.125	0.260	-1.378	0.096*	-1.061	0.289
15	1.374	0.104	0.115	1.636	0.102	-1.092	0.275	-1.354	0.117	-1.040	0.298
16	1.715	0.292	0.150	0.443	0.658	-0.406	0.685	-0.601	0.378	-0.383	0.702
17	1.398	0.106	0.067	2.098	0.036**	-0.998	0.318	-1.365	0.171	-0.979	0.328
18	1.370	0.103	0.106	1.649	0.099*	-1.165	0.244	-1.446	0.104	-1.086	0.277
19	1.393	0.105	0.095	1.927	0.054*	-1.024	0.306	-1.369	0.152	-0.991	0.322
20	1.354	0.102	0.159	1.290	0.197	-1.079	0.281	-1.347	0.086*	-1.028	0.304
Mean	1.476	0.150	0.117	1.51	0.180	-0.831	0.433	-1.096	0.226	-0.836	0.439
Std. Dev	0.132	0.0683	0.0285	0.525	0.177	0.361	0.215	0.348	0.125	0.417	0.218

Tabela 1 – Parâmetros encontrados para a lei de potência com amplitudes diferentes

Analisando tanto a figura quanto as comparações, é razoável dizer que os dados podem ter sido gerados por uma lei de potência, ou talvez mais precisamente, por uma lei de potência exponencialmente truncada. É interessante notar que, independente do verdadeiro processo gerador dos dados, o processo parece ser consistente independente da amplitude e, pela rejeição da distribuição exponencial em grande parte dos casos, a distribuição deve ser de cauda pesada.