

## Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

---

# Mapeamento de Séries Financeiras em Redes Complexas

Amanda Leite de Camargo<sup>1</sup>

Centro de Engenharia, Modelagem e Ciências Sociais, UFABC, Santo André, SP

Marcio Eisencraft<sup>2</sup>

Escola Politécnica, Universidade de São Paulo, SP

**Resumo.** Recentemente, teorias referentes ao estudo de redes complexas estão sendo utilizadas em uma aplicação até então inédita: a análise da dinâmica de séries temporais. Diversos autores propuseram procedimentos distintos para converter uma série temporal em uma rede complexa. Nesse trabalho mapeia-se séries temporais de mercado financeiro em redes complexas por meio de quantis. Algumas propriedades das redes resultantes são analisadas, inferindo-se informações a respeito do comportamento das séries.

**Palavras-chave.** Redes Complexas, Séries Temporais, Séries Financeiras, Dinâmica de Séries Temporais, PETRA.

## 1 Introdução

Redes são representações das relações existentes entre pares de elementos de um conjunto. Nesse contexto, tais elementos são chamados de *vértices* e as relações entre eles, *arestas* ou *ligações* [9]. Redes mapeiam diversas situações reais. Um exemplo consiste nas comunidades virtuais como Facebook, Twitter e Instagram [9]. Nelas, os usuários, representativos dos vértices, estabelecem relações com os demais, constituindo arestas.

Tem-se utilizado a teoria de redes complexas na análise de séries temporais. Essas redes podem ser construídas a partir de séries temporais por meio de diversas técnicas [5, 7, 8, 14, 15].

Em [15], definem-se ciclos no sinal e faz-se a análise da correlação entre seus pontos. Tal modelagem é aplicável a séries cujos ciclos são bem definidos. Em [5] gera-se uma rede dividindo-se uma série em quantis e construindo-se uma matriz de pesos que representa o número de vezes que dois quantis adjacentes são encontrados. Já em [8], a matriz de recorrência obtida da série equivale à matriz de adjacência da rede. Em [14] é proposto um modelo em que o espaço de fase do sinal é reconstruído e cada um de seus pontos representa um vértice da rede resultante. A modelagem proposta por [7] utiliza o *Visibility Algorithm*. Nele, os valores da série são dispostos em um gráfico de barras. Cada barra associa-se a um vértice e as arestas estabelecem-se de forma que cada barra é conectada a todas as outras que podem ser vistas do topo da primeira.

---

<sup>1</sup>amanda.leite@ufabc.edu.br

<sup>2</sup>marcio@lcs.poli.usp.br

O objetivo desse artigo é converter duas séries financeiras em redes complexas pelo mapeamento por quantis, proposto por [5], além de analisar as propriedades dessas redes, inferindo-se o comportamento das séries. A primeira série consiste no valor das cotações diárias de fechamento da ação preferencial da empresa Petrobrás, *PETRA*. A outra, consiste nos valores de retorno diário da mesma ação, que mostra a variação do seu preço em relação ao preço do dia anterior.

Esse artigo está estruturado da seguinte forma: na Seção 2 são conceituadas propriedades úteis à caracterização de redes complexas, além dos modelos presentes na literatura. Na Seção 3, o mapeamento de uma série em uma rede por quantis é explicado detalhadamente. Na Seção 4, aplica-se esse mapeamento às séries financeiras. Na Seção 5 traçam-se as conclusões e trabalhos futuros.

## 2 Redes Complexas e Métricas

Uma rede com  $N$  vértices é representada matematicamente por um grafo  $G$ , composto por um conjunto de  $M$  arestas,  $E(G)$ , e relações que associam a cada uma delas, dois vértices [13]. Vértices e arestas serão denotados, respectivamente por:  $\{i, j, k, \dots, n\}$  e  $\{l_{ij}, l_{jk}, \dots, l_{mn}\}$ . Dois vértices  $i$  e  $j$  são vizinhos se a aresta  $l_{ij} \in E$  [9]. Uma rede também é representada por uma matriz de adjacências  $N \times N$ ,  $\mathbf{A}$ . Nela, se dois vértices  $i$  e  $j$  estiverem conectados,  $A_{ij} = 1$ . Caso contrário,  $A_{ij} = 0$ .

Um *caminho*  $L_{ij}$  entre dois vértices  $i$  e  $j$  é um conjunto de arestas que se inicia em  $i$  e termina em  $j$ . Entre os possíveis caminhos, aquele que apresenta menos arestas é chamado de *menor caminho* [9]. Caso não exista caminho entre os nós  $i$  e  $j$ ,  $L_{ij} = \infty$  [9]. Nesse caso, um desses vértices está *desconectado* da rede. O *menor caminho médio*,  $\langle L \rangle$ , para uma rede direcionada é definido como

$$\langle L \rangle = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1}^N L_{ij}. \quad (1)$$

O *grau de conectividade*,  $K_i$  de um vértice  $i$  é dado pelo número de arestas que partem ou chegam nele [9]. O grau de conectividade médio,  $\langle K \rangle$  para uma rede direcionada é

$$\langle K \rangle = \frac{1}{N} \sum_{i=1}^N K_i. \quad (2)$$

A *distribuição de graus de conectividade*,  $P(K)$ , indica a fração de vértices com grau  $K$  existente na rede.

O *coeficiente de agrupamento* do vértice  $i$ ,  $C_i$ , é definido como a razão entre o número de arestas existentes entre os vizinhos desse vértice e o número máximo de arestas possíveis entre eles [9]. O *coeficiente de agrupamento médio*,  $\langle C \rangle$  de uma rede direcionada é dado por

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^N C_i. \quad (3)$$

A *modularidade* quantifica a divisão de uma rede em partições de elevado coeficiente de agrupamento [6]. Para uma partição de vértices em uma rede, a modularidade,  $\mathcal{M}_p$ , é dada por

$$\mathcal{M}_p = \sum_{s=1}^r \left[ \frac{M_s}{M} - \left( \frac{d_s}{2M} \right)^2 \right], \quad (4)$$

sendo  $r$  o número de partições na rede,  $M$  o total de arestas na rede,  $M_s$ , o número de arestas entre os vértices da partição  $s$  e  $d_s$ , a soma dos graus de conectividade dos vértices da partição  $s$ , [5, 6]. A modularidade da rede,  $\mathcal{M}$ , é o máximo valor de  $\mathcal{M}_p$  dentre todas as possíveis partições da rede [6].

Na literatura, encontram-se modelos de redes que apresentam propriedades topológicas características. Nesse trabalho, são de relevância as redes aleatórias, *small world* e livres de escala. Numa rede aleatória de  $N$  vértices, a probabilidade de dois vértices quaisquer  $i$  e  $j$  estarem conectados é igual a  $p$ . Essas redes possuem  $\langle L \rangle$  pequeno em relação a  $N$ ,  $P(K)$  binomial e  $\langle K \rangle$  é a moda da distribuição [5, 9]. Quando  $N \rightarrow \infty$  e  $p \rightarrow 0$ , essa distribuição binomial pode ser aproximada por uma distribuição do tipo Poisson [9].

Denomina-se rede regular, a rede de forma circular cujas arestas estabelecem-se apenas entre os  $n$  vizinhos mais próximos de seus vértices. Essas redes são úteis à geração de redes *small world*. Para tal, realoca-se algumas arestas da rede regular conectando vértices distantes. Assim, as redes *small world* incorporam características de redes regulares e aleatórias [12]. Tem-se  $\langle L \rangle_{aleatória} \ll \langle L \rangle_{smallworld} \ll \langle L \rangle_{regular}$ , em que  $\langle L \rangle_{smallworld} \simeq \langle L \rangle_{aleatória} \ll N$ , ou seja,  $\langle L \rangle_{smallworld}$  aproxima-se mais de  $\langle L \rangle_{aleatória}$  do que de  $\langle L \rangle_{regular}$ . Logo, nessas redes, a distância média entre os vértices é pequena em comparação com o número de vértices, caracterizando o *fenômeno mundo pequeno* [9]. Em relação ao  $\langle C \rangle$  de uma rede *small world*, espera-se que  $\langle C \rangle_{aleatória} \ll \langle C \rangle_{smallworld}$  e  $P(K)$  do tipo Poisson [9].

O modelo livre de escala foi proposto por [1]. Para construir uma rede desse tipo inicia-se com um pequeno número de vértices de mesmo grau de conectividade. Posteriormente, adicionam-se novos vértices conectando-os, com maior probabilidade, aos vértices de maior grau de conectividade. Como os vértices mais conectados tendem a receber mais arestas, essas redes apresentam *hubs*, vértices altamente conectados que ocorrem em menor número na rede [5]. Nessas redes,  $P(K)$  tem a forma de uma lei de potência [9].

### 3 Mapeamento de Séries Temporais em Redes Complexas

No mapeamento proposto por [5], uma rede complexa com pesos e arestas direcionadas é obtida a partir de uma série temporal qualquer, não sendo restrito a nenhum tipo específico de sinal. Nele, a série é discretizada com base nos *quantis* dos seus valores. Define-se *quantil* de ordem  $100m$ , indicado por  $q(m)$ , onde  $m$  é uma proporção qualquer,  $0 < m < 1$ , tal que  $100m\%$  das observações sejam menores do que  $q(m)$  [10].

A obtenção de uma rede por meio desse mapeamento é descrita a seguir: i. dada uma série temporal, divide-se ela em  $Q$  quantis com  $q_i = (T \cdot i)/Q$ ; ii. cada quantil é associado a um vértice  $i$  pertencente à rede; iii. os vértices da rede são conectados dois a dois por

meio de arestas  $l_{ij}$ , cujos pesos são determinados pelo número de vezes que um ponto  $x_n$  pertencente a um quantil  $q_i$  é seguido por outro ponto  $x_{n+1}$ , pertencente ao quantil  $q_j$ .

Uma vez que as redes geradas são direcionadas e com peso, elas devem ser limiarizadas para possibilitar os cálculos das propriedades da Seção 2. Na limiarização, retiram-se as arestas cujo peso é menor que um limiar pré-definido e associam-se pesos unitários às arestas remanescentes [5].

## 4 Aplicação a Séries de Mercado Financeiro

Nessa seção aplica-se o mapeamento em questão a duas séries financeiras. A primeira,  $x(n)$ , consiste no valor em reais das cotações diárias de fechamento da ação preferencial da Petrobrás, *PETRA* [2]. Essa ação faz parte do índice Bovespa, que mede o desempenho dos ativos de maior representatividade do mercado brasileiro [3]. Tomou-se o intervalo entre o dia 26 de outubro de 2010 e o dia 09 de abril de 2015 [2], considerando-se assim, 1000 amostras. Na Figura 1a, mostra-se a evolução temporal de  $x(n)$ . Nota-se que entre 2010 e 2011, o preço da ação oscila entre 23 e 27 reais. Do final de 2011 até aproximadamente a metade de 2014, ele sofre queda, oscilando entre 12 e 23 reais. Após alta no final de 2014, observa-se uma queda acentuada, verificada até abril de 2015.

A segunda série,  $r(n)$ , compreende os valores porcentuais dos retornos diários da mesma ação, calculados durante o mesmo período. O retorno diário mostra a variação do preço da ação em relação ao preço do dia anterior. Obtém-se os valores dos retornos por  $r(n) = x(n)/x(n-1) - 1$  [11].  $r(n)$  também possui 1000 amostras, sendo visualizada na Figura 1b. Seu comportamento reflete o de  $x(n)$ , visto que no período em que ela sofre queda, os retornos são predominantemente negativos e nos momentos de alta nos preços, os retornos são em sua maioria positivos. Ambas as séries foram divididas em 30 quantis com o mesmo número de amostras e posteriormente convertidas em redes com 30 vértices. Nelas, cada vértice associa-se a um quantil de forma que a ordem dos quantis corresponde à ordem dos vértices. Logo, o primeiro quantil é representado pelo vértice 01 nas duas redes e assim por diante.

Essas redes foram limiarizadas de forma que não houvesse vértices desconexos. Para

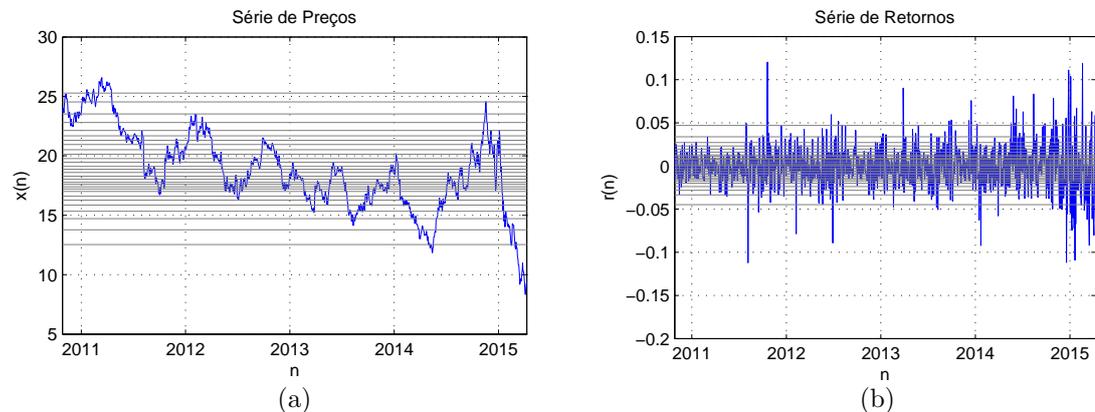


Figura 1: (a) Série de preços e (b) retornos diários da ação *PETRA* de 26/10/2010 a 09/04/2015.



Figura 2: Redes resultantes do mapeamento das séries (a)  $x(n)$  e (b)  $r(n)$

tal, definiu-se 3 como limiar para ambos os casos e por isso, arestas de peso inferior a 3 foram desconsideradas. A rede resultante de  $x(n)$  é mostrada na Figura 2a e a rede resultante de  $r(n)$ , na Figura 2b.

Ambas as redes também foram particionadas utilizando-se o algoritmo Spectral Partitioning, do MIT Strategic Engineering Research Group [4]. Para as duas séries, a partição que maximiza a modularidade é composta por dois agrupamentos. Nas redes da Figura 2, todos os vértices de cor azul pertencem à partição 1 e os de cor rosa, à partição 2. Após a limiarização, as métricas:  $\langle L \rangle$ ,  $\langle K \rangle$ ,  $\langle C \rangle$  e  $\mathcal{M}$  foram calculadas conforme definições da Seção 2 para ambas as redes. Esses resultados são visualizados na Tabela 1. Adicionalmente, obtiveram-se os histogramas representativos da  $P(K)$  dessas redes, visualizados nas Figuras 3a para  $x(n)$  e 3b para  $r(n)$ .

Para  $x(n)$ , a rede obtida dessa série apresentou um  $\langle L \rangle$  elevado em relação a  $N$ . O  $\langle K \rangle$  da mesma rede não aproxima-se da moda em sua  $P(K)$ , conforme observa-se na Figura 3a. O  $\langle C \rangle$  obtido para essa rede não é elevado, porém não aproxima-se de zero. A mesma rede possui  $\mathcal{M}$  negativa, indicando que quando ela é dividida, não apresenta partições de modularidade elevada. Nela, observa-se um elevado número de conexões entre vértices próximos. Tal comportamento indica a existência de um grande número de amostras consecutivas na série que encontram-se no mesmo quantil ou em quantis adjacentes. Por esse motivo, concluiu-se que  $x(n)$  varia lentamente ao longo do tempo. Tendências de alta e baixa em  $x(n)$  são estimadas pela direção das arestas da rede resultante. Conexões que partem de determinado vértice e chegam a outro vértice de ordem superior representam um comportamento crescente na série (aumento do preço). Por outro lado, arestas que partem de um vértice qualquer e chegam a outro de ordem inferior indicam comportamento decrescente da série, sugerindo a diminuição do preço. Assim, é possível afirmar que essa rede exibe características próprias, possuindo algumas propriedades de redes regulares, como o tipo de conexões estabelecidas e o valor de  $\langle C \rangle$ .

Por outro lado, a rede resultante do mapeamento de  $r(n)$  possui  $\langle L \rangle$  baixo em relação a  $N$ . Essa tendência é observada em redes *small world* e aleatórias [5,12]. O valor

Tabela 1: Métricas e desvios padrão obtidos para as redes resultantes das séries  $x(n)$  e  $r(n)$ .

Série	$N$	$\langle L \rangle$	$\langle K \rangle$	$\langle C \rangle$	$\mathcal{M}$
$x(n)$	30	$(6.71 \pm 4.44)$	$(5.66 \pm 1.98)$	$(0.32 \pm 0.25)$	-0.09
$r(n)$	30	$(2.95 \pm 1.09)$	$(5.73 \pm 1.77)$	$(0.10 \pm 0.06)$	0.02

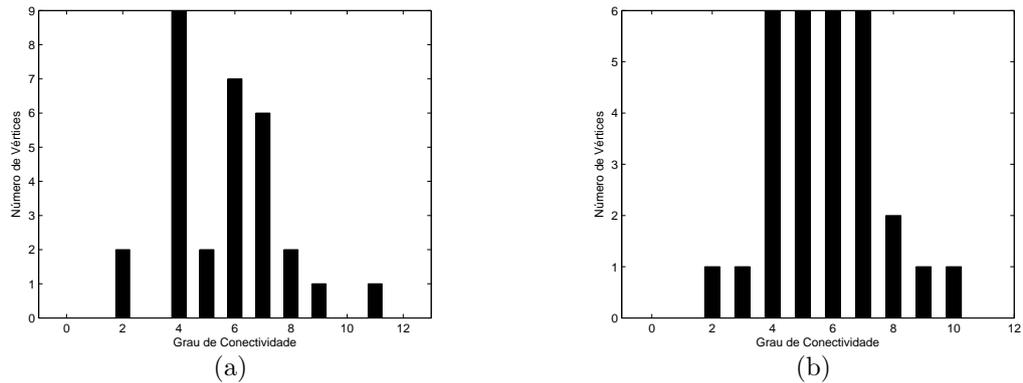


Figura 3: Histogramas representativos da  $P(K)$  das redes obtidas das séries (a)  $x(n)$  e (b)  $r(n)$ .

de  $\langle K \rangle = 5.73$ , obtido para a mesma rede, aproxima-se de  $K = 5$  e  $K = 6$  que são dois dos valores que mais ocorrem na  $P(K)$  dessa rede, conforme é visualizado na Figura 3b. O  $\langle C \rangle$  para a rede obtida de  $r(n)$  não é elevado, caracterizando conexões que não estabelecem-se necessariamente entre vértices adjacentes. A  $\mathcal{M}$  dessa rede é positiva, porém não é próxima de um, inferindo-se que quando a rede é dividida, as partições resultantes não apresentam modularidade elevada. Nessa rede, as conexões distribuem-se entre vértices correspondentes a quantis não necessariamente adjacentes. Assim, concluiu-se que  $r(n)$  varia de forma semelhante a uma série de ruído. Tendências de alta e baixa em  $r(n)$  também podem ser detectadas pela direção das arestas da rede associada a  $r(n)$ . Dessa forma, a rede associada a  $r(n)$  pode ser classificada como aleatória devido aos valores obtidos para  $\langle L \rangle$  e  $\langle K \rangle$ , e ao tipo de conexão estabelecidas entre seus vértices.

De forma geral, concluiu-se que  $x(n)$  varia mais lentamente em relação a  $r(n)$ . Os pesos das arestas também são representativos do comportamento de  $x(n)$  e  $r(n)$ . Eles associam-se às chances dos preços e retornos realizarem as transições indicadas pela direção das arestas. Essa análise, que pretende-se fazer em um trabalho futuro, pode ser utilizada para a detecção de padrões no comportamento futuro dessas séries.

## 5 Conclusões

Nesse artigo, mostrou-se a aplicação do mapeamento por quantis a séries financeiras. Foi possível detectar tendências no seu comportamento e classificar as rede associadas a elas por meio de suas propriedades.

Tratando-se de trabalhos futuros, pretende-se aplicar o mesmo modelo a outras séries financeiras, estudar as características das redes associadas, além de comparar os resultados obtidos aos já observados. Pretende-se também analisar a influência da métrica  $\mathcal{M}$  na distinção dessas séries.

## Agradecimentos

ALC agradece ao suporte financeiro do CNPq e da UFABC. ME agradece ao suporte financeiro do CNPq, processos 479901/2013 – 9 e 311575/2013 – 7 e da FAPESP, processo 2014/04864 – 2.

## Referências

- [1] A. L. Barabási and R. Albert, Emergence of scaling in random networks, *Science*, vol.286, 509–512, (1999), DOI: 10.1126/science.286.5439.509.
- [2] BM&F Bovespa, Cotações históricas - PETR4. Disponível em: <http://www.bmfbovespa.com.br/pt-br/cotacoes-historicas/FormSeriesHistoricas.asp>. Acessado em: 24/04/2015.
- [3] BM&F Bovespa, Índice bovespa. Disponível em: <http://www.bmfbovespa.com.br/indices/ResumoIndice.aspx?Indice=Ibovespa&Idioma=pt-br>. Acessado em 24/04/2015.
- [4] G. Bounova and O. Weck, Overview of metrics and their correlation patterns for multiple-metric topology analysis on heterogeneous graph ensembles, *Phys. Rev. E*, vol. 85, (2012). DOI:10.1103/PhysRevE.85.016117.
- [5] A. S. L. O. Campanharo, M. I. Sirer, R. D. Malmgren, F. M. Ramos and L. A. N. Amaral, Duality between time series and networks, *PLoS ONE*, Public Library of Science, vol. 6, (2011).
- [6] R. Guimerá, M. Sales-Pardo and L. A. N. Amaral, Classes of complex networks defined by role-to-role connectivity profiles, *Nature Physics*, vol. 3, 63–69, (2007).
- [7] L. Lacasa, L. Bartolo, F. Ballesteros, J. Luque and J. C. Nuño, From time series to complex networks: the visibility graph, *Proceedings of the National Academy of Sciences*, vol. 105, 4972–4975, (2008).
- [8] N. Marwan, J. F. Donges, Z. Yong, R. V. Donner and J. Kurths, Complex network approach for recurrence analysis of time series, *Physics Letters A*, vol. 373, 4246–4254, (2009).
- [9] L. H. A. Monteiro, *Sistemas dinâmicos*, Editora Livraria da Física, Cap. 2, (2014).
- [10] P. A. Morettin e W. O. Bussab, *Estatística básica*, Saraiva, (2004).
- [11] R. S. Tsay, *Analysis of financial time series*, John Wiley & Sons, Cap. 1, vol. 543, (2005).
- [12] D. J. Watts and S. H. Strogatz, Collective dynamics of ‘small-world’ networks, *Nature*, vol. 393, 440–442, (1998).
- [13] D. B. West, *Introduction to graph theory*, Prentice Hall Upper Saddle River, vol. 2, (2001).
- [14] X. Xu, J. Zhang and M. Small, Superfamily phenomena and motifs of networks induced from time series, *Proceedings of the National Academy of Sciences*, vol. 105, 19601–19605,(2008).
- [15] J. Zhang and M. Small, Complex network from pseudoperiodic time series: topology versus dynamics, *Physical Review Letters*, vol. 96, (2006).