

## 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 mapeiam-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* [1]. Redes mapeiam diversas situações reais. Um exemplo consiste nas comunidades virtuais como Facebook, Twitter e Instagram [1]. Nelas, os usuários, representados pelos vértices, estabelecem relações entre si, representadas pelas 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 [2–6]. Em [2] definem-se ciclos no sinal e faz-se a análise da correlação entre eles. Tal modelagem é aplicável a séries cujos ciclos são bem definidos. Em [3] gera-se uma rede digitalizando uma série em certo número de níveis ou *quantis* e construindo a matriz de pesos que representa o número de vezes que dois quantis adjacentes são encontrados. Em [4] a matriz de recorrência obtida da série é utilizada como matriz de adjacências da rede. Em [5] é proposto um modelo em que o espaço de fase do sinal é reconstruído e cada um de seus pontos representa um vértice na rede resultante. A modelagem proposta por [6] 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 do presente artigo é utilizar o algoritmo de mapeamento proposto em [3] para converter duas séries temporais provindas do mercado financeiro em redes complexas. A partir das redes infere-se propriedades das séries originais. A primeira série consiste no valor das cotações diárias de fechamento da ação preferencial da empresa Petrobrás, *PETRA*. A outra série consiste nos valores de retorno diário da mesma ação, que mostra a variação do seu preço em relação ao preço no dia anterior.

Esse artigo está estruturado da seguinte forma: na Seção 2 são conceituadas propriedades úteis à caracterização de redes complexas, além de revistos alguns modelos presentes na literatura. Na Seção 3 o mapeamento de uma série em uma rede por meio de 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 [7]. Vértices e arestas são 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  $l_{ij} \in E$  [1]. Uma rede também é representada por uma matriz de adjacências  $N \times N$ ,  $\mathbf{A}$ , com elementos  $a_{ij}$ . 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 uma sequência de arestas que se inicia em  $i$  e termina em  $j$ . Entre os possíveis caminhos, aqueles que apresentam o menor número de arestas são chamados de *menores caminhos* [1]. Caso não exista caminho entre os vértices  $i$  e  $j$ ,  $L_{ij} = \infty$  [1]. 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 terminam nele [1]. 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*  $C_i$  do vértice  $i$  pertencente a uma rede direcionada é definido como a razão entre o número de arestas  $E_i$  existentes entre os vizinhos desse vértice e o número máximo de arestas possíveis entre eles, dado por

$$C_i = \frac{E_i}{V_i(V_i-1)}, \quad (3)$$

em que  $V_i$  representa o número de vizinhos de  $i$  [1]. O *coeficiente de agrupamento médio*  $\langle C \rangle$  de uma rede desse tipo é dado por

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^N C_i. \quad (4)$$

A *modularidade*  $\mathcal{M}$  mede a possibilidade de dividir uma rede em subredes ou módulos. O conjunto de módulos em que uma rede é dividida denomina-se *partição*. Em geral, procuram-se partições em que vértices de um mesmo módulo tenham muitas conexões e vértices de módulos diferentes tenham poucas conexões entre si. Para uma partição, a modularidade  $\mathcal{M}_P$  é dada por

$$\mathcal{M}_P = \sum_{s=1}^r \left[ \frac{M_s}{M} - \left( \frac{d_s}{2M} \right)^2 \right], \quad (5)$$

sendo  $r$  o número de módulos da partição,  $M$  o total de arestas na rede,  $M_s$  o número de arestas entre os vértices do  $s$ -ésimo módulo e  $d_s$  a soma dos graus de conectividade dos vértices do  $s$ -ésimo módulo [3, 8]. A modularidade da rede é o máximo valor de  $\mathcal{M}_P$  dentre todas as possíveis partições [8].

A seguir revisitam-se os principais modelos de redes complexas presentes na literatura: 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 fixa e igual a  $p$ . Essas redes possuem  $\langle L \rangle$  pequeno em relação a  $N$  e  $P(K)$  binomial com  $\langle K \rangle$  sendo a moda da distribuição [1, 3]. Quando  $N \rightarrow \infty$  e  $p \rightarrow 0$ , essa distribuição binomial pode ser aproximada por uma distribuição do tipo Poisson [1].

Denomina-se rede regular à rede de forma circular cujas arestas estabelecem-se apenas entre os  $n$  vizinhos mais próximos de seus vértices. A partir delas pode-se gerar as redes *small world* [9]. Para tal, realocam-se algumas arestas da rede regular conectando vértices distantes. Assim, as redes *small world* incorporam características de redes regulares e aleatórias [9]. Tem-se  $\langle L \rangle_{\text{aleatória}} < \langle L \rangle_{\text{smallworld}} < \langle L \rangle_{\text{regular}}$ , em que  $\langle L \rangle_{\text{smallworld}} \simeq \langle L \rangle_{\text{aleatória}} \ll N$ , ou seja,  $\langle L \rangle_{\text{smallworld}}$  aproxima-se mais de  $\langle L \rangle_{\text{aleatória}}$  do que de  $\langle L \rangle_{\text{regular}}$ . Assim sendo, nessas redes a distância média entre os vértices é pequena em comparação ao seu número de vértices, caracterizando o *fenômeno mundo pequeno* [1]. Em relação ao  $\langle C \rangle$  de uma rede *small world*, espera-se que  $\langle C \rangle_{\text{aleatória}} \ll \langle C \rangle_{\text{smallworld}}$  e que  $P(K)$  seja do tipo Poisson [1].

O modelo livre de escala foi proposto por [10]. 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*, que são vértices altamente conectados que ocorrem em menor número [3]. Nessas redes,  $P(K)$  têm a forma de uma lei de potência [1].

### 3 Mapeamento de Séries Temporais em Redes Complexas

No mapeamento proposto por [3], uma rede complexa com pesos e arestas direcionadas é obtida a partir de uma série temporal qualquer. Define-se *quantil* de ordem  $m$  de uma série temporal, indicado por  $q(m)$ , em que  $0 < m < 1$ , à amplitude da amostra tal que  $100m\%$  das amostras da série sejam menores do que  $q(m)$  [3].

A obtenção de uma rede por meio desse mapeamento é descrita a seguir: i. dada uma série temporal  $x(n)$  com  $N_p$  amostras, obtém-se seu histograma; ii. a partir do histograma, divide-se  $x(n)$  em  $Q$  quantis  $\{q_1, q_2, \dots, q_Q\}$  tais que  $q_i = q(i/Q)$ . O número de amostras cujos valores estão entre  $q_i$  e  $q_{i+1}$  é dado por  $N_p/Q$ ; iii. cria-se uma rede com  $Q$  vértices  $\{v_1, v_2, \dots, v_Q\}$  que representam esses quantis; iv. os vértices da rede são conectados dois a dois por meio de arestas  $l_{ij}$ , cujos pesos são determinados pelo número de vezes que uma amostra  $x_n$  pertencente a um quantil  $q_i$  é seguida por outra amostra  $x_{n+1}$ , pertencente ao quantil  $q_j$ .

Uma vez que as redes geradas são direcionadas e com peso, deve ser utilizado um limiar de forma a definir se dois vértices estão ou não conectados, possibilitando o cálculo das propriedades da Seção 2. Na aplicação do limiar, retiram-se as arestas cujo peso é menor que o valor pré-definido e associam-se pesos unitários às arestas remanescentes [3]. A escolha do limiar é feita de forma a excluir exceções do mapeamento. Esse valor deve ser calibrado de acordo com a série em análise. Um limiar excessivamente elevado causa a perda de informações relevantes enquanto um limiar muito baixo considera as exceções.

## 4 Aplicação a Séries de Mercado Financeiro

Nessa seção aplica-se o mapeamento discutido a duas séries financeiras.

A primeira,  $x(n)$ , consiste no valor em reais das cotações diárias de fechamento da ação preferencial da Petrobrás, *PETRA* [12]. Essa ação faz parte do índice Bovespa, que mede o desempenho dos ativos de maior representatividade do mercado brasileiro [13]. Tomou-se o intervalo entre o dia 26 de outubro de 2010 e o dia 09 de abril de 2015 [12], 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 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 percentuais dos retornos diários da mesma ação, calculados durante o mesmo período. O retorno diário mostra a variação no preço da ação em relação ao preço no dia anterior. Obtém-se os valores dos retornos por  $r(n) =$

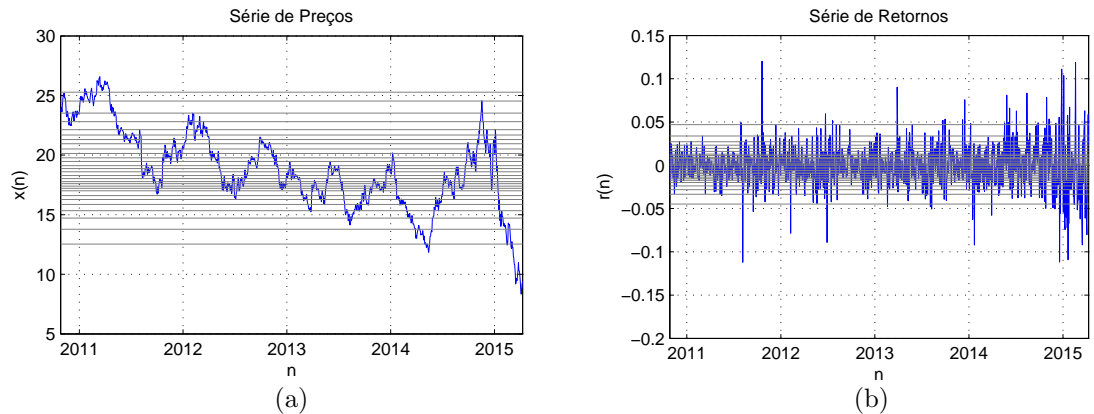


Figura 1: (a) Série de preços e (b) retornos diários da ação *PETRA* de 26/10/2010 a 09/04/2015.



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

$x(n)/x(n-1) - 1$  [14]. Essa série 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 em sua maior parte negativos e nos momentos de alta nos preços, os retornos são em sua maioria positivos. Nas Figuras 1a e 1b também são mostrados os limites entre os quantis empregados.

Ambas as séries foram divididas em 30 quantis e convertidas em redes com 30 vértices. Nelas, cada vértice associa-se a um quantil sendo que o primeiro quantil corresponde ao Vértice 01 e assim por diante. Aplicou-se para ambos os casos um limiar igual a 3 de forma que não houvesse vértices desconexos. A rede resultante para  $x(n)$  é mostrada na Figura 2a e a rede resultante para  $r(n)$ , na Figura 2b. Ambas as redes também foram particionadas utilizando-se o algoritmo *Spectral Partitioning* [15]. Para essas redes, a partição que maximiza a modularidade é composta por dois módulos. Nelas, os vértices de cor azul pertencem ao Módulo 01 e os vértices de cor rosa, ao Módulo 02. 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  $\langle L \rangle$  elevado em relação a  $N$ . O  $\langle K \rangle$  da mesma rede não aproxima-se da moda da sua  $P(K)$ , conforme observa-se na Figura 3a. O  $\langle C \rangle$  obtido para ela não é elevado, porém não aproxima-se de zero. Essa 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 na 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

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

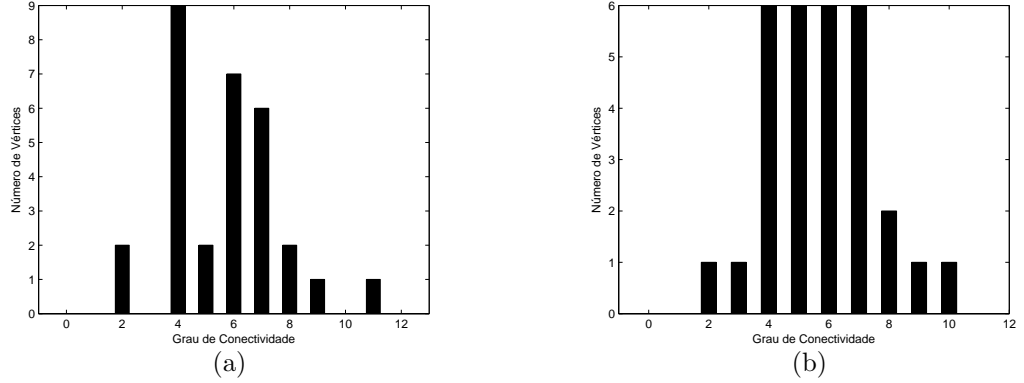


Figura 3: Histogramas representativos da  $P(K)$  das redes obtidas das séries (a)  $x(n)$  e (b)  $r(n)$ .

de 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 [3, 9]. O  $\langle K \rangle = 5.73$  obtido aproxima-se de  $K = 5$  e  $K = 6$  que são dois dos valores que mais ocorrem na sua  $P(K)$ , conforme visualizado na Figura 3b. O  $\langle C \rangle$  da mesma rede não é elevado, caracterizando conexões que não se estabelecem necessariamente entre vértices adjacentes. A  $\mathcal{M}$  dessa rede é positiva, porém não é próxima de um, inferindo-se que quando ela 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 essa série. Dessa forma, a rede obtida de  $r(n)$  pode ser classificada como aleatória devido aos valores de  $\langle L \rangle$  e  $\langle K \rangle$  e ao tipo de conexões 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 dessas séries. Eles associam-se às chances dos preços e retornos realizarem as transições indicadas pela direção das arestas. Pretende-se apresentar em um trabalho futuro essa análise, que pode ser utilizada para a detecção de padrões no comportamento dessas séries.

## 5 Conclusões

Nesse artigo, mostrou-se a aplicação do mapeamento por meio de quantis a séries financeiras. Foi possível detectar tendências no seu comportamento e classificar as redes 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] L. H. A. Monteiro, *Sistemas dinâmicos*, Editora Livraria da Física, Cap.2, (2014).
- [2] J. Zhang and M. Small, Complex network from pseudoperiodic time series: topology versus dynamics, *Physical Review Letters*, vol. 96 (2006)
- [3] A. S. L. O. Campanharo, M. I. Siner, R. D. Malmgrem, F. M. Ramos and L. A. N. Amaral, Duality between time series and networks, *PLoS ONE*, Public Library of Science, vol. 6, (2011).
- [4] 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)
- [5] 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),
- [6] 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)
- [7] D. B. West and others, *Introduction to graph theory*, Prentice Hall Upper Saddle River, vol.2, (2001).
- [8] 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).
- [9] D. J. Watts and S. H. Strogatz, Collective dynamics of 'small-world' networks, *Nature*, vol.393, 440–442, (1998).
- [10] 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.
- [11] P. A. Morettin e W. O. Bussab, *Estatística básica*, Saraiva, (2004).
- [12] BM&F Bovespa, Cotações históricas - PETR4. Disponível em: <http://www.bmfbovespa.com.br/pt-br/cotacoes-historicas/FormSeriesHistoricas.asp>. Acessado em: 24/04/2015.
- [13] 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.
- [14] R. S. Tsay, *Analysis of financial time series*, John Wiley & Sons, Cap.1 vol. 543, (2005).
- [15] 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. Disponível em: <http://link.aps.org/doi/10.1103/PhysRevE.85.016117>. Acessado em: 02/02/2015.