
PHÂN TÍCH CẤU TRÚC PHỤ THUỘC CỦA CÁC CỔ PHIẾU ĐƯỢC NIÊM YẾT TRÊN SỞ GIAO DỊCH CHỨNG KHOÁN TP. HỒ CHÍ MINH - TIẾP CẬN ĐỒ THỊ LỌC PHẪNG CỰC ĐẠI

Nguyễn Thị Thảo

Trường Đại học Kinh tế Quốc dân

Email: thaonguyen@neu.edu.vn

Hoàng Đức Mạnh

Trường Đại học Kinh tế Quốc dân

Email: manhhd@neu.edu.vn

Nguyễn Thị Hà Giang

Học viên cao học I.S.F.A - Lyon 1 University

Email: hagiangnguyen96@gmail.com

Mã bài: JED - 691

Ngày nhận bài: 30/05/2022

Ngày nhận bài sửa: 11/08/2022

Ngày duyệt đăng: 15/08/2022

Tóm tắt

Trong bài báo này, chúng tôi dùng đồ thị lọc phẳng cực đại PMFG (Planar Maximally Filtered Graphs) để mô tả sự phụ thuộc lẫn nhau giữa các mã cổ phiếu được niêm yết trên Sở Giao dịch Chứng khoán Thành phố Hồ Chí Minh (HOSE). Từ đồ thị PMFG, danh mục đầu tư sẽ được xây dựng lần lượt trên nhóm các mã cổ phiếu trung tâm và trên nhóm các mã cổ phiếu ngoại vi của đồ thị. Việc so sánh hiệu quả của các danh mục nhận được sẽ cho biết vai trò của mối quan hệ giữa các mã cổ phiếu trong việc thiết lập danh mục đầu tư để từ đó đưa ra các chiến lược đầu tư phù hợp. Với thị trường chứng khoán Việt Nam, danh mục đầu tư trên nhóm cổ phiếu ngoại vi cũng mang lại hiệu quả vượt trội, ngoại trừ giai đoạn 2019-2020, khi nền kinh tế nói chung và thị trường chứng khoán nói riêng bị ảnh hưởng nặng nề bởi dịch COVID-19.

Từ khóa: danh mục đầu tư, đồ thị mạng PMFG, hiệu quả danh mục.

Mã JEL: C6, C61, D4, D53, G11, G12

Analyzing the network structure of listing stocks on Ho Chi Minh City Stock Exchange - a Planar Maximally Filtered Graphs approach

Abstract:

In this paper, we use the Planar Maximally Filtered Graphs (PMFG) to describe the interdependence between stocks listed on the Ho Chi Minh Stock Exchange (HOSE). From the PMFG graph, portfolios will be built on the central group of stocks and on the peripheral group of stocks, respectively. Portfolios performances will be compared to evaluate the role of the relationship between stocks in establishing portfolios so that appropriate investment strategies can be made. For Vietnamese stock market, the portfolio on peripheral stocks brings outstanding performance, except for the period of 2019-2020, when the economy in general, and in particular, the stock market are severely affected by the COVID-19 pandemic.

Keywords: Planar Maximally Filtered Graph, Portfolio, Portfolio performance.

JEL codes: C6, C61, D4, D53, G11, G12

1. Giới thiệu

Phân tích và quản lý danh mục đầu tư là một bộ phận quan trọng của lý thuyết tài chính hiện đại. Cho đến nay, chủ đề này đang được phát triển theo nhiều hướng khác nhau và vẫn nhận được sự quan tâm của nhiều nhà nghiên cứu trên thế giới.

Việc nắm được mối quan hệ giữa các cổ phiếu được niêm yết cùng một thị trường, mức độ biến động cùng nhau giữa chúng trên sẽ giúp các nhà đầu tư đưa ra những quyết định đúng đắn khi lựa chọn các mã cổ phiếu đưa vào danh mục. Mantegna (1999), Tumminello & cộng sự (2005) và Peralta & Zareei (2016) đã nghiên cứu xây dựng các đồ thị mạng mô tả mối quan hệ giữa các tài sản trên thị trường, trên đồ thị này sẽ có một nhóm các tài sản nằm ở phía trung tâm của đồ thị, có mối tương quan cao với các tài sản khác và có mức biến động cùng nhau cao, một nhóm tài sản nằm ở khu vực ngoại vi của đồ thị, ít chịu tác động từ sự biến động giá của các tài sản khác. Đồ thị được xây dựng bởi Mantegna có tên gọi là cây bao trùm cực tiểu (Minimal Spanning Tree – MST) và đồ thị lọc phẳng cực đại (Planar Maximally Filtered Graph - PMFG) được xây dựng theo nghiên cứu của Tumminello & cộng sự (2005).

Từ cấu trúc cơ bản của các mối quan hệ trên thị trường được thể hiện trên đồ thị mạng, Pozzi & cộng sự (2013) nâng cao quy trình lựa chọn danh mục đầu tư khi chỉ ra danh mục đảm bảo được tính đa dạng hóa sẽ gồm các tài sản nằm trên các nút ngoại vi của đồ thị.

Trong bài báo này, nhóm tác giả sẽ đi theo hướng nghiên cứu xây dựng đồ thị PMFG để trực quan hóa cấu trúc phụ thuộc của các tài sản trên thị trường chứng khoán Việt Nam sau đó, thực hiện xây dựng danh mục đầu tư gồm các tài sản thuộc nhóm trung tâm, nhóm ngoại vi của đồ thị và so sánh hiệu quả của các danh mục này với danh mục gồm các tài sản được chọn một cách ngẫu nhiên trên thị trường.

Bài báo gồm 5 phần, ngoài phần giới thiệu, phần 2 sẽ trình bày về tổng quan các công trình nghiên cứu liên quan, phương pháp nghiên cứu sẽ được trình bày ở phần 3, phần 4 trình bày về dữ liệu nghiên cứu và kết quả thực nghiệm, phần 5 trình bày các kết luận của nghiên cứu.

2. Tổng quan nghiên cứu

Việc ứng dụng lý thuyết mạng vào phân tích mối quan hệ giữa các cổ phiếu trên thị trường tài chính là một chủ đề hấp dẫn trong những năm gần đây. Ngoài việc cho thấy sự phụ thuộc giữa các cổ phiếu, cấu trúc mạng còn cho biết vai trò của từng cổ phiếu trên thị trường tài chính. Từ đó, có thể xem mạng như một màng lọc thông tin thị trường, chỉ giữ lại những thông tin có vai trò quan trọng và loại bỏ đi các thông tin gây nhiễu. Trong nghiên cứu của Mantegna (1999), ông đề xuất xây dựng một cấu trúc mạng có tên là cây bao trùm cực tiểu (Minimum Spanning Tree – MST), ở đó các đỉnh chính là các mã cổ phiếu của các công ty, độ đo khoảng cách được dùng để xây dựng các cạnh nối các đỉnh được tính dựa trên hệ số tương quan giữa các chuỗi lợi suất theo ngày của các mã cổ phiếu. Trong MST sẽ không có sự xuất hiện của các chu trình (cycle), các đỉnh được kết nối với nhau thông qua các cạnh với nguyên tắc làm cực tiểu tổng khoảng cách (tổng độ dài các cạnh), với đỉnh thì sẽ có cạnh trong một MST.

Tuy nhiên, nghiên cứu của Pozzi & cộng sự (2008) đã chỉ ra rằng khi mỗi liên hệ giữa các cổ phiếu chỉ được thể hiện trên MST sẽ mất nhiều thông tin về sự phụ thuộc giữa các cổ phiếu trên thị trường. Hơn nữa, trong dài hạn, MST cũng cho thấy tính ổn định (stability) và tính vững (robustness) thấp. Vào năm 2005, Tumminello & cộng sự (2005) đã xây dựng mạng lọc thông tin có tên là đồ thị lọc phẳng cực đại (Planar Maximally Filtered Graph – PMFG), đồ thị này chứa nhiều thông tin về sự tương tác giữa các cổ phiếu trên thị trường hơn MST, hay chính xác hơn, MST là một bộ phận trong PMFG, tất cả các cạnh có trong MST đều xuất hiện trong PMFG. Đồ thị PMFG được xây dựng dựa trên nguyên tắc đảm bảo số *genus* của đồ thị luôn bằng 0, các đỉnh được kết nối với nhau hoặc có dạng tam giác (chu trình gồm ba đỉnh) hoặc có dạng đồ thị con gồm bốn đỉnh với đầy đủ các kết nối giữa chúng. Nghiên cứu cũng đã chỉ ra đồ thị này đặc biệt phù hợp khi khoảng cách giữa các đỉnh được xây dựng dựa trên hệ số tương quan. Kết quả nghiên cứu thực nghiệm trên 100 cổ phiếu trên thị trường Hoa Kỳ cho thấy các chu trình dạng tam giác và nhóm đầy đủ gồm bốn đỉnh phản ánh khá đầy đủ các mối quan hệ quan trọng, có ý nghĩa với cấu trúc và tính chất của thị trường.

Từ cấu trúc mạng dạng MST của thị trường tài chính, Onnela & cộng sự (2003) đã đề xuất khai thác cấu trúc cơ bản của mạng như là một công cụ hữu hiệu trong việc nâng cao quy trình chọn danh mục đầu tư vì ở khía cạnh đa dạng hóa, nghiên cứu cho thấy danh mục đầu tư tạo bởi mô hình M-V thường gồm các tài sản

nằm trên các lá ngoài của cây MST. Tuy nhiên, nhóm nghiên cứu chưa cung cấp tiêu chí để dựa trên mạng, tìm ra trọng số tối ưu của danh mục. Peralta & Zareei (2016) đã thiết lập một cầu nối giữa mô hình M-V và lý thuyết mạng, cho thấy mối quan hệ nghịch chiều giữa độ lớn của trọng số các tài sản trong danh mục được xây dựng bởi mô hình M-V và độ trung tâm của tài sản trong đồ thị MST.

Pozzi & cộng sự (2013) đã ứng dụng cấu trúc liên kết dạng đồ thị PMFG của thị trường tài chính để thiết lập danh mục đầu tư, họ xem xét việc phân bổ đều vốn hoặc sử dụng mô hình M-V để tìm trọng số phân bổ vốn vào các tài sản ứng với các đỉnh nằm ở vùng trung tâm và ngoại vi nhất của đồ thị PMFG. Việc phân định một đỉnh nằm ở ngoại vi hay trung tâm của biểu đồ được dựa trên chỉ số tổng hợp được tạo nên từ các chỉ số Degree, Betweenness, Eccentricity, Closeness và Eigenvector Centrality của lý thuyết đồ thị. Kết quả nghiên cứu cho thấy dù vốn được phân bổ đều hay trọng số được xác định qua phương pháp M-V thì danh mục được xây dựng từ các tài sản ngoại vi cũng mang lại hiệu quả vượt trội so với danh mục được xây dựng trên nhóm các tài sản nằm ở nhóm trung tâm của đồ thị.

Ngoài phương pháp phân nhóm cổ phiếu dựa trên đồ thị mạng, một số các phương pháp phân cụm khác cũng được các nhà nghiên cứu sử dụng và danh mục đầu tư được thiết lập từ các cụm cổ phiếu này cũng mang lại hiệu quả cao hơn so với chỉ số thị trường (Tola & cộng sự, 2008; Pai & Michel, 2009; Nanda & cộng sự, 2010).

Thực tế cho thấy, đồ thị mạng được sử dụng có hiệu quả trong việc nghiên cứu sự thay đổi của cấu trúc thị trường khi xảy ra khủng hoảng (Tu, 2014; Zhao & cộng sự, 2016; Aslam & cộng sự, 2020; Millington & Niranjana, 2021). Việc quan sát được sự thay đổi của cấu trúc ở các giai đoạn đặc biệt này sẽ giúp các nhà điều hành chính sách và các nhà đầu tư đưa ra những quyết định đúng đắn.

Với thị trường chứng khoán Việt Nam, cấu trúc phụ thuộc của các mã cổ phiếu niêm yết trên thị trường cũng thu hút sự quan tâm của nhiều nhà nghiên cứu (Hoàng Đức Mạnh, 2013; Trần Trọng Nguyên & Nguyễn Thu Thủy, 2017; Moslehpour & cộng sự, 2022). Nghiên cứu của Nguyen & cộng sự (2019) với dữ liệu chứng khoán trong giai đoạn từ 09/01/2008 đến 31/12/2017, dùng sự thay đổi cấu trúc của MST để mô tả tác động của cuộc khủng hoảng tài chính 2011-2012 lên thị trường chứng khoán. Theo tìm hiểu của chúng tôi, tại Việt Nam, các nghiên cứu ứng dụng lý thuyết mạng vào thị trường chứng khoán còn khá ít, đặc biệt là cách tiếp cận PMFG hầu như chưa được áp dụng. Vì vậy, nghiên cứu này sẽ tập trung vào cách tiếp cận PMFG để phân tích cấu trúc phụ thuộc của các cổ phiếu niêm yết trên sàn HOSE, một sàn giao dịch lớn của Thị trường Chứng khoán Việt Nam, và ứng dụng nó trong xây dựng danh mục đầu tư.

3. Phương pháp nghiên cứu và dữ liệu

3.1. Mô hình trung bình – phương sai (M-V)

Theo Markowitz, mục tiêu của mô hình M-V là xác định các tỷ lệ vốn X_i được đầu tư vào tài sản thứ i trong một giỏ các tài sản đã chọn sao cho rủi ro của danh mục đo bởi phương sai là nhỏ nhất và kỳ vọng của danh mục không thấp hơn một mức \bar{r} đã được ấn định trước. Bài toán được phát biểu như sau:

Tìm véc tơ trọng số $x = (x_1, x_2, \dots, x_n)$ sao cho:

$$\sigma_p^2 = \sum_{i=1}^n \sum_{j=1}^n x_i x_j \text{Cov}(r_i, r_j) \rightarrow \min \quad (1)$$

với các ràng buộc:

$$E_p = \sum_{i=1}^n x_i E(r_i) \geq \bar{r}, \quad \sum_{i=1}^n x_i = 1, \quad 0 \leq x_i \leq 1, \quad i = 1, 2, \dots, n$$

Trong đó, n là số tài sản trong danh mục, r_i và r_j là chuỗi lợi suất theo ngày của tài sản thứ i và thứ j tương ứng ($i, j = 1, 2, \dots, n, i \neq j$), $x = (x_1, \dots, x_n)$ là véc tơ trọng số của các tài sản trong danh mục, điều kiện $0 \leq x_i \leq 1, i = 1, \dots, n$ để đảm bảo không có hiện tượng bán khống, \bar{r} là mức lợi suất cho trước.

3.2. Tương quan có trọng số của các chuỗi lợi suất

Theo các chuyên gia phân tích dữ liệu, thông tin từ các sự kiện gần hiện tại có giá trị cao hơn so với các sự

kiện từ xa cho cả mục đích mô tả và dự báo. Trong nghiên cứu của mình, Pozzi & cộng sự (2012) đã dùng trọng số dạng mũ khi tính các đặc trưng của chuỗi lợi suất trong khoảng thời gian T, cụ thể:

$$w_t = w_0 e^{\alpha(t-\Delta t)} \quad \forall t \in \{1, 2, \dots, T\} \quad (2)$$

Với $\alpha \geq 0$ là tham số cho biết mức độ tăng theo thời gian của trọng số w_t và

$$w_0(\alpha) = \frac{1 - e^{-\alpha T}}{1 - e^{-\alpha \Delta t}} \quad (3)$$

Các giá trị đặc trưng có trọng số của chuỗi lợi suất như trung bình \bar{r}_i^w , độ lệch chuẩn σ_i^w , hiệp phương sai $(\sigma_{ij}^w)^2$ và hệ số tương quan giữa các chuỗi lợi suất ρ_{ij} sẽ được tính theo công thức sau:

$$\begin{aligned} \bar{r}_i^w &= \sum_{t=1}^T w_t r_{it} \\ \sigma_i^w &= \sqrt{\sum_{t=1}^T w_t (r_{it} - \bar{r}_i^w)^2} \\ \sigma_{ij}^w &= \sqrt{\sum_{t=1}^T w_t (r_{it} - \bar{r}_i^w)(r_{jt} - \bar{r}_j^w)} \\ \rho_{ij} &= \frac{\sigma_{ij}^w}{\sigma_i^w \sigma_j^w} \end{aligned}$$

Ma trận hệ số tương quan có trọng số của n tài sản sẽ là ma trận cấp $n \times n$ với các phần tử của ma trận là ρ_{ij} , ($i, j = 1, \dots, n$). Ma trận này sẽ được dùng làm đầu vào để xây dựng đồ thị mạng PMFG.

3.3. Đồ thị mạng PMFG

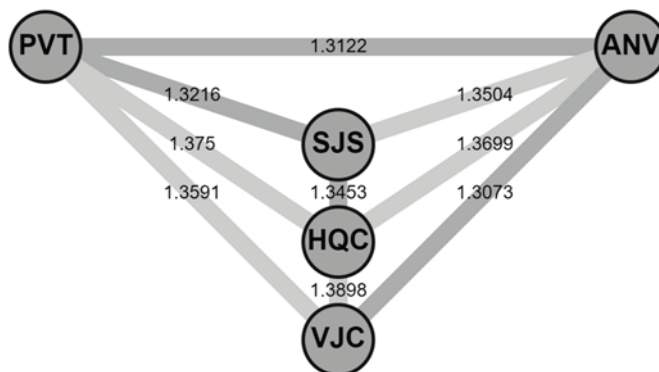
Để xây dựng đồ thị mạng, trước hết ta sẽ định nghĩa *khoảng cách* giữa hai cổ phiếu i và j trên thị trường thông qua hệ số tương quan có trọng số ρ_{ij} :

$$d_{ij} = \sqrt{2(1 - \rho_{ij})} \quad (5)$$

Giá trị của d_{ij} sẽ nằm trong khoảng $[0; 2]$ và thỏa mãn cả ba tiên đề về khoảng cách. Khi giá trị của hệ số tương quan ρ_{ij} càng nhỏ (càng lớn) thì khoảng cách giữa hai cổ phiếu được định nghĩa qua d_{ij} càng lớn (càng nhỏ).

Theo Tumminello & cộng sự (2005), đồ thị PMFG được xây dựng dựa trên ma trận khoảng cách D gồm các phần tử d_{ij} ($i = 1, \dots, n$; $j = 1, \dots, n$). Với thị trường chứng khoán, đồ thị PMFG sẽ kết nối n đỉnh (chính là n mã cổ phiếu trên thị trường) bằng $3(n-2)$ cạnh theo nguyên tắc: nếu vẽ tất cả các đỉnh và các cạnh nối chúng trên một mặt phẳng thì luôn tồn tại cách vẽ để không có hai cạnh nào của đồ thị cắt nhau và tổng độ dài các cạnh trên đồ thị là nhỏ nhất. Trong PMFG, các đồ thị con liên thông (các đỉnh phân biệt của đồ thị con đều được nối với nhau) chỉ gồm 3 hoặc 4 đỉnh (3-cliques or 4-cliques). Các mã cổ phiếu trong một đồ thị con liên thông sẽ có giá biến động cùng chiều hay nói cách khác giá của một cổ phiếu trong đồ thị con liên thông biến động sẽ có thể là nguyên nhân để giá của các cổ phiếu khác trong đồ thị con biến động (theo Boginski & cộng sự, 2005).

Hình 1: Đồ thị mạng PMFG gồm 5 đỉnh



Nguồn: Tác giả vẽ dựa trên chuỗi giá đóng cửa của một số mã cổ phiếu niêm yết trên sàn HOSE

Ma trận đỉnh kề không có trọng số cấp $n \times n$ (the unweighted adjacency matrix) A_F của đồ thị PMFG là ma trận gồm các phần tử a_{ij} xác định như sau

$$a_{ij} = \begin{cases} 1, & \text{nếu trong PMFG có cạnh nối đỉnh } i \text{ và đỉnh } j \\ 0, & \text{nếu trong PMFG không có cạnh nối đỉnh } i \text{ và đỉnh } j. \end{cases}$$

Ma trận đỉnh kề có trọng số $\mathbf{A}_F^W = \mathbf{A}_F \circ \mathbf{D}$, trong đó \mathbf{D} là ma trận khoảng cách được nêu ở trên, ký hiệu “ \circ ” chính là tích Hadamard của hai ma trận \mathbf{A}_F và \mathbf{D} , mỗi phần tử của \mathbf{A}_F^W được xác định bằng tích của từng phần tử của ma trận \mathbf{A}_F với từng phần tử của ma trận \mathbf{D} .

Ma trận \mathbf{H}^W cấp $n \times n$ với các phần tử $h_{ij} = 1 + \rho_{ij}$, được gọi là ma trận đo độ mạnh của các kết nối giữa các đỉnh trong đồ thị, h_{ij} nhận giá trị trong khoảng $[0; 2]$. Với đồ thị PMFG, ma trận đo độ mạnh của các kết nối trong đồ thị là $\mathbf{H}_F^W = \mathbf{H}^W \circ \mathbf{A}_F$.

3.4. Độ đo mức độ trung tâm và ngoại vi của các đỉnh trong đồ thị

Trong phần này ta sẽ xem xét vị trí của các cổ phiếu, chính là các đỉnh trong đồ thị PMFG, và cách thức kết nối các đỉnh này với nhau, cho biết cấu trúc phụ thuộc của các mã cổ phiếu trên thị trường chứng khoán. Vị trí của các cổ phiếu sẽ được xác định thông qua việc so sánh các chỉ số: *Degree (DC)*, *Betweenness (BC)*, *Eccentricity (EC)*, *Closeness (CC)* và *Eigenvector Centrality (EV)* của từng đỉnh trên đồ thị. Trong khi *Degree*, *Betweenness* và *Eigenvector Centrality* đo độ trung tâm của một đỉnh, khi chúng nhận giá trị càng lớn thì đỉnh càng ở vị trí trung tâm của đồ thị thì *Eccentricity* và *Closeness* lại là đo độ ngoại vi của các đỉnh, khi các độ đo này càng lớn thì đỉnh càng ở vị trí xa trung tâm của đồ thị. Để thống nhất cách tiếp cận các độ đo này từ cùng một góc nhìn hoặc là độ đo trung tâm hoặc là độ đo ngoại vi, ở đây ta lựa chọn góc nhìn trung tâm, ta sẽ sắp xếp các độ đo *Degree*, *Betweenness* và *Eigenvector* theo thứ tự giảm dần, còn độ đo *Eccentricity* và *Closeness* theo thứ tự tăng dần. Như vậy, sau khi sắp xếp lại, những cổ phiếu nào có thứ hạng cao đều là các cổ phiếu được xếp ở vị trí trung tâm của đồ thị. Với mỗi chỉ số này ta đều tính cho cả trường hợp có trọng số và không có trọng số, cụ thể, ký hiệu giá trị có trọng số và không có trọng số của *Degree* là DC^w và DC^u của *Betweenness* ký hiệu là BC^w và BC^u , của *Eigenvector* là EV^w và EV^u , của *Eccentricity* là EC^w và EC^u , của *Closeness* là CC^w và CC^u .

Để so sánh mức độ trung tâm của các đỉnh, ta cần tạo ra một chỉ số tổng hợp. Bằng cách phân tích nhân tố, ta nhận được hai nhân tố. Nhân tố thứ nhất ký hiệu là PCI_1 gồm các chỉ số DC^w , DC^u , BC^w và BC^u ; nhân tố thứ hai, ký hiệu là PCI_2 gồm các chỉ số EV^w , EV^u , CC^w , CC^u , EC^w và EC^u . Nhân tố PCI_1 cho biết mức độ kết nối của nút đang xét, PCI_2 cho biết mức độ quan trọng của các nút mà nút đang xét kết nối đến trên đồ thị. Theo nghiên cứu của Pozzi & cộng sự (2013), các chỉ số PCI_1 và PCI_2 được tính theo công thức sau:

$$PCI_1 = \frac{(DC^w + DC^u + BC^w + BC^u - 4)}{4(n - 1)} \quad (6)$$

$$PCI_2 = \frac{(EV^w + EV^u + CC^w + CC^u + EC^w + EC^u - 6)}{6(n - 1)} \quad (7)$$

Để tổng hợp thông tin nhận được từ cả PCI_1 và PCI_2 , ta tạo nên chỉ số tổng hợp

$$PC = PCI_1 + PCI_2 \quad (8)$$

cho biết đầy đủ thông tin về mức độ trung tâm và kết nối của các nút trên đồ thị. Các giá trị của chỉ số tổng hợp PC được sắp xếp theo thứ tự tăng dần, những nút có thứ hạng theo PC cao (giá trị của PC nhỏ) là những nút nằm ở vị trí trung tâm nhất, còn những nút có thứ hạng theo PC thấp (giá trị của PC lớn) thì nằm ở vùng ngoại vi của đồ thị.

3.5. So sánh hiệu quả của danh mục

Sau khi tính được chỉ số tổng hợp, ta có thể xác định được mức độ trung tâm và ngoại vi của các đỉnh trên đồ thị hay chính là các mã cổ phiếu trên thị trường. Danh mục được xây dựng dựa trên nhóm cổ phiếu ngoại vi được dự báo là sẽ mang lại hiệu quả tốt hơn với đồ thị MST (Onnela & cộng sự (2003)), điều này liệu có đúng với đồ thị lọc PMFG? Để trả lời câu hỏi này, ta sẽ sử dụng phương pháp M-V hoặc trọng số bằng nhau nhau để xây dựng các loại danh mục trên các nhóm cổ phiếu: gồm các cổ phiếu ngoại vi; gồm các cổ phiếu trung tâm và gồm các cổ phiếu được chọn ngẫu nhiên sau đó so sánh hiệu quả của chúng.

Hiệu quả của danh mục được đo dựa trên tỷ số của lợi suất trung bình của danh mục và độ lệch chuẩn của lợi suất danh mục, ký hiệu:

$$IR = \frac{\bar{r}_P}{s_P}$$

tỷ số này được biết đến với tên gọi ‘signal-to-noise ratio’ hoặc ‘information ratio’. Một danh mục đầu tư hiệu quả sẽ là danh mục mang lại lợi suất cao và rủi ro danh mục thấp, hay giá trị ‘signal-to-noise ratio’ cao, và ngược lại sẽ là danh mục có hiệu quả thấp.

3.6. Dữ liệu

Dữ liệu được sử dụng trong nghiên cứu là giá đóng cửa theo ngày đã điều chỉnh của tất cả 404 mã cổ phiếu đang được niêm yết trên Sở Giao dịch Chứng khoán Thành phố Hồ Chí Minh (HOSE) trong khoảng thời gian từ 27/09/2018 đến 28/12/2021.

Trong nghiên cứu này, để so sánh hiệu quả của việc xây dựng danh mục đầu tư từ các nhóm cổ phiếu trung tâm và ngoại vi của đồ thị PMFG qua các thời điểm khác nhau của thị trường, nhóm tác giả chia dữ liệu thành ba giai đoạn năm 2019, năm 2020 và năm 2021. Ở mỗi giai đoạn các mã cổ phiếu, việc xử lý các quan sát khuyết thiếu được thực hiện theo nguyên tắc nếu bị mất 5 quan sát liên tiếp thì mã cổ phiếu đó sẽ bị loại khỏi bộ dữ liệu, nếu mất ít hơn 5 quan sát thì các quan sát bị mất sẽ được lấp đầy bởi giá trị quan sát liền trước nó. Vì thị trường chứng khoán Việt Nam nằm trong nhóm các thị trường cận biên, quy mô thị trường thay đổi liên tục, luôn có các công ty mới được niêm yết trên thị trường trong các giai đoạn nghiên cứu nên để đảm bảo tính ổn định, chúng tôi chỉ lựa chọn những cổ phiếu có khối lượng giao dịch từ 1000 cổ phiếu mỗi ngày (Nguyen & cộng sự, 2019). Sau khi làm sạch dữ liệu và xử lý các quan sát khuyết thiếu ở mỗi giai đoạn, chúng tôi sẽ tiến hành xây dựng đồ thị PMFG, xác định nhóm cổ phiếu trung tâm và ngoại vi, sau đó xây dựng danh mục và tính toán hiệu quả danh mục. Chẳng hạn, với số liệu lợi suất của chuỗi giá đóng cửa của các cổ phiếu trong 250 ngày từ 27/9/2018 đến 27/9/2019, nghiên cứu này sẽ xây dựng đồ thị PMFG của các cổ phiếu và tìm ra các danh mục đầu tư tối ưu rồi so sánh hiệu quả của các danh mục này trong khoảng thời gian 20 ngày tiếp theo (1 tháng). Các đồ thị và kết quả tính toán trong bài báo được thực hiện với sự hỗ trợ của phần mềm Python.

4. Kết quả nghiên cứu và thảo luận

4.1. Xây dựng đồ thị PMFG

Ký hiệu giá của cổ phiếu thứ i trong ngày giao dịch t của mỗi giai đoạn là P_{it} , $t = 0, \dots, 250$ lợi suất của cổ phiếu được tính theo công thức

$$r_{it} = \frac{P_{it} - P_{i,t-1}}{P_{i,t-1}}, t = 1, \dots, 250$$

Với các cổ phiếu i và cổ phiếu j , khoảng cách được dùng để xây dựng đồ thị mô tả mối quan hệ giữa chúng được tính dựa trên hệ số tương quan có trọng số có hiệu chỉnh ρ_{ij} , cụ thể:

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

Đồ thị PMFG được xây dựng dựa trên các bộ dữ liệu chia theo giai đoạn: giai đoạn 1 từ 27/9/2018 đến 30/9/2019, giai đoạn 2 từ ngày 02/10/2019 đến 30/9/2020, giai đoạn 3 từ ngày 30/9/2020 đến 30/9/2021. Trong mỗi giai đoạn sẽ có dữ liệu về lợi suất của tất cả các cổ phiếu trong 250 ngày giao dịch. Sau khi xây dựng đồ thị PMFG, chỉ số PC được tính toán, đây là chỉ tiêu để cho biết một cổ phiếu trên đồ thị ở vị trí trung tâm hay ngoại vi. Kết quả về nhóm cổ phiếu trung tâm và ngoại vi của mỗi giai đoạn được tổng hợp trong Bảng 1 và Bảng 2.

Trước thời điểm dịch covid bắt đầu, 09/2018-09/2018, nhóm trung tâm chủ yếu gồm các công ty thuộc ngành nguyên liệu và ngành tài chính. Trong giai đoạn 09/2019-09/2020, ngoài các cổ phiếu của ngành nguyên liệu nằm ở nhóm trung tâm, còn có thêm cổ phiếu của các công ty trong ngành công nghiệp. Sau thời gian dịch COVID-19 bùng phát mạnh, với số liệu từ 09/2020 đến 09/2021, nhóm cổ phiếu trung tâm của thị trường xuất hiện 3 công ty thuộc nhóm ngành chăm sóc sức khỏe ngoài các công ty thuộc ngành công nghiệp. Những cổ phiếu xuất hiện ở nhóm trung tâm là những cổ phiếu có nhiều kết nối hay có tương quan mạnh với nhiều các cổ phiếu khác trên thị trường.

Bảng 1: Bảng nhóm cổ phiếu trung tâm trong từng giai đoạn

Thứ hạng	09/2018 – 9/2019	09/2019 – 9/2020	09/2020 – 9/2021
1	MIG	HPX	VPS
2	SHA	DTL	SPM
3	DTA	TCD	OPC
4	DHC	SFI	GAB
5	SGT	TTE	YEG
6	NAV	VID	VPD
7	KOS	SBV	VNL
8	MCG	TCO	CTF
9	TLD	SMA	TCO
10	SVI	PTL	SVC

Nguồn: Tính toán của tác giả dựa trên giá đóng cửa của các mã cổ phiếu niêm yết trên sàn HOSE

Bảng 2: Nhóm cổ phiếu ngoại vi trong từng giai đoạn

Thứ hạng	09/2018 – 9/2019	09/2019 – 9/2020	09/2020 – 9/2021
1	PVT	TRC	SRC
2	HQC	PHC	JVC
3	VJC	HAP	TDG
4	SJS	KMR	TLG
5	ANV	BTP	HAX
6	IDI	KDC	ADS
7	VCI	NKG	SGR
8	ACL	GEG	HNG
9	NHA	DCL	KHP
10	PMG	SFG	TCB

Nguồn: Tính toán của tác giả dựa trên giá đóng cửa của các mã cổ phiếu niêm yết trên sàn HOSE

Với nhóm cổ phiếu ngoại vi, kết quả thực nghiệm cho thấy ở trước giai đoạn dịch COVID-19 bắt đầu, nhóm này chủ yếu bao gồm cổ phiếu của các công ty sản xuất mặt hàng tiêu dùng, thực phẩm và dịch vụ (ANV, ACL, IDI, VJC). Dịch COVID-19 bùng phát mạnh đã ảnh hưởng lên mọi mặt của đời sống xã hội, việc giãn cách, hạn chế giao thương giữa các quốc gia, và có thêm các rào cản trong việc xuất nhập khẩu đã tác động mạnh lên nền kinh tế cũng như mối quan hệ giữa các cổ phiếu trên của thị trường, điều này cũng thể hiện qua sự thay đổi của đồ thị PMFG. Với số liệu 09/2019-09/2020, nhóm ngoại vi nhất của PMFG gồm cổ phiếu của các công ty trong các ngành năng lượng (BTP, GEG). Các kết quả này vẫn thống nhất với kết quả nhận được từ nghiên cứu của Pozzi & cộng sự (2008)

4.2. Thiết lập danh mục đầu tư

Với số liệu của mỗi giai đoạn, sau khi xây dựng đồ thị PMFG, nhóm cổ phiếu trung tâm và ngoại vi đã được xác định dựa trên chỉ số tổng hợp, theo như đề xuất trong nghiên cứu của Pozzi & cộng sự (2013), bài báo tiến hành xây dựng danh mục đầu tư với từng nhóm cổ phiếu: nhóm gồm 5 cổ phiếu trung tâm nhất, nhóm gồm 5 cổ phiếu ngoại vi nhất và nhóm gồm 5 cổ phiếu được lựa chọn ngẫu nhiên từ tất cả các cổ phiếu đang niêm yết trên sàn HOSE. Với mỗi nhóm cổ phiếu trọng số của danh mục tối ưu được xác định theo hai cách, cách thứ nhất là lấy trọng số của tất cả các tài sản trong danh mục bằng nhau và cách thứ hai là xác định dựa trên mô hình M-V. Khi sử dụng mô hình M-V, trọng số tối ưu của danh mục được chọn sao cho giá trị của tỷ số Sharpe của danh mục, SR lớn nhất

$$SR = \frac{r_P - r_f}{\sigma_P}$$

trong đó r_p , r_f và σ_p lần lượt là lãi suất phi rủi ro, kỳ vọng và độ lệch chuẩn của lợi suất danh mục được tính dựa trên số liệu trong khoảng thời gian 125 ngày (từ 04/01/2019 đến 27/09/2019 với giai đoạn 1; từ 31/03/2020 đến 25/09/2020 với giai đoạn 2, từ 31/03/2021 đến 28/09/2021 với giai đoạn 3). Trong nghiên cứu này nhóm tác giả sẽ sử dụng lãi suất của trái phiếu chính phủ kỳ hạn 10 năm làm lãi suất phi rủi ro. Bảng

3 sẽ liệt kê trọng số của các tài sản có được bằng mô hình M-V.

Bảng 3: Trọng số của các danh mục được xây dựng ở các nhóm tài sản khác nhau

	Nhóm trung tâm	Trọng số	Nhóm ngoại vi	Trọng số	Nhóm chọn ngẫu nhiên	Trọng số
Năm 2019	MIG	0	PVT	0.178228996	PLX	0
	SHA	0	HQC	0	YBM	10^{-16}
	DHC	0.550014748	VJC	0.821771004	FPT	0.900349208
	DTA	0.115404056	SJS	0	HHS	10^{-16}
	SGT	0.334581196	ANV	0	HPX	0.099650792
Năm 2020	HPX	0.58708858	TRC	0.290391335	SHB	0
	DTL	0	PHC	0.171313592	PPC	0
	TCD	0.30433487	BTP	0.246337151	JVC	0.43400308
	SFI	0.10857655	HAP	0.291957923	BTP	0.284046104
	TTE	0	KMR	0	ACL	0.281950815
Năm 2021	VPS	0.013180538	SRC	0	NKG	0.286586553
	SPM	0.22688324	JVC	0.101544805	TMT	0.018743151
	OPC	0.000000649	ADS	0.898455195	VNL	0.242716863
	GAB	0.759927832	TLG	0	FMC	0.245943559
	YEG	0.00000774	HAX	0	SGT	0.206009874

Nguồn: Tính toán của tác giả dựa trên giá đóng cửa của các mã cổ phiếu niêm yết trên sàn HOSE

4.3. So sánh hiệu quả các danh mục vừa được thiết lập

Trong nghiên cứu này chúng tôi dùng chỉ số signal-to-noise IR làm tiêu chí để so sánh hiệu quả của các danh mục khác nhau. Với mỗi danh mục nhận được, lợi suất của danh mục với chu kỳ τ ngày được tính theo công thức:

$$r_t(\tau) = \frac{P_{t+\tau} - P_t}{P_t}$$

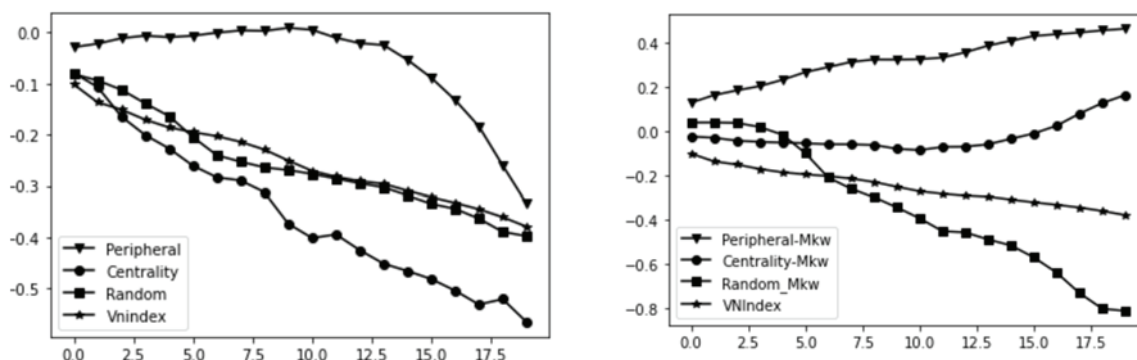
với $\tau = 1, 2, \dots, 20$, P_t là giá trị của danh mục ở thời điểm t , $t = 252, \dots, T_0$. Ký hiệu $\bar{r}(\tau)$ là mức thay đổi trung bình của giá sau τ ngày, $s(\tau)$ là độ lệch chuẩn của mức thay đổi giá sau τ ngày $\tau = 1, 2, \dots, 20$. Tỷ số

$$IR = \frac{\bar{r}(\tau)}{s(\tau)}, \tau = 1, 2, \dots, 20.$$

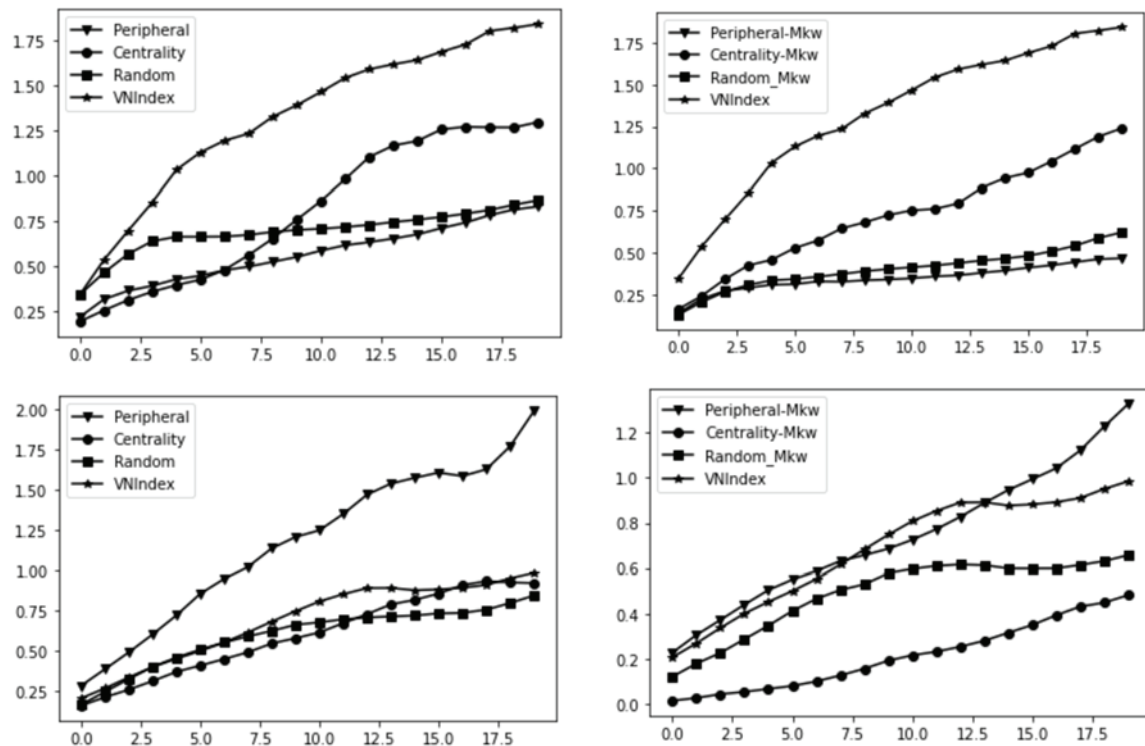
được gọi là tỷ số thông tin (Information Ratio) hay signal-to-noise của danh mục. Khi τ thay đổi $\tau = 1, 2, \dots, 20$ với các nhóm tài sản khác nhau, sẽ vẽ nên các đường cong khác nhau cho biết mức độ hiệu quả của danh mục nhận được, danh mục nào có đường IR nằm cao hơn thì sẽ càng hiệu quả.

Ở giai đoạn trước khi có dịch COVID-19, ta nhận thấy hiệu quả vượt trội của danh mục được xây dựng trên nhóm tài sản ngoại vi dù với trọng số bằng nhau hay trọng số được xác định bằng mô hình M-V, còn cao hơn lợi suất của cả thị trường (đường nối các ngôi sao). Tới giai đoạn đầu năm 2020, khi dịch COVID-19 bắt

Hình 2: Tỷ số signal-to-noise của các danh mục được lập từ các nhóm tài sản



Hình 2 (tiếp)



Chú ý: Hai hình trên cùng, hai hình ở giữa và hai hình dưới cùng lần lượt là tỷ số IR của danh mục dựa trên số liệu năm 2019, 2020 và 2021 với hai cách tính trọng số: trọng số bằng nhau (bên trái) và trọng số xác định theo mô hình M-V (bên phải).

Nguồn: Tính toán của tác giả dựa trên giá đóng cửa của các mã cổ phiếu niêm yết trên sàn HOSE.

đầu lây lan vào Việt Nam, hàng loạt các biện pháp cách ly, giãn cách được đưa ra đã ảnh hưởng đến chuỗi cung ứng từ đó tác động mạnh mẽ đến tất cả các khu vực của thị trường, hiệu quả của danh mục được xây dựng trên nhóm ngoại vi ở giai đoạn này thấp hơn danh mục được lập bởi các tài sản trong nhóm trung tâm và thấp hơn đường hiệu quả chung của cả thị trường. Vào năm 2021, khi vaccine đã bắt đầu được tiêm cho người dân, Chính phủ đã chủ động hơn trong việc điều hành và ứng phó dịch bệnh, những thay đổi này đã có tác động lên thị trường chứng khoán, cấu trúc mối quan hệ giữa các tài sản được thể hiện trên đồ thị PMFG đã dần quay trở về đúng quỹ đạo. Lúc này, danh mục đầu tư được xây dựng trên nhóm cổ phiếu ngoại vi lại thể hiện tính hiệu quả tốt hơn danh mục được xây dựng từ các nhóm tài sản khác và tốt hơn đường lợi suất chung của cả thị trường.

5. Kết luận

Thông qua việc xây dựng đồ thị PMFG, bài báo đã góp phần cung cấp một cái nhìn trực quan về cấu trúc phụ thuộc của các cổ phiếu được niêm yết trên thị trường chứng khoán Việt Nam và tác động của cuộc khủng hoảng do dịch bệnh lên thị trường. Trước khi COVID-19 xuất hiện, nhóm trung tâm sẽ gồm cổ phiếu của các công ty tài chính và các công ty thuộc ngành nguyên vật liệu, nhóm ngoại vi gồm cổ phiếu của các công ty trong ngành năng lượng và ngành dịch vụ tiện ích, điều này thống nhất với cấu trúc thị trường được mô tả trong nghiên cứu của Pozzi & cộng sự (2008). Trong giai đoạn dịch COVID-19 xuất hiện, cấu trúc thị trường cũng có nhiều thay đổi. Nhóm cổ phiếu trung tâm khi đó không chỉ gồm các cổ phiếu của các công ty thuộc ngành nguyên vật liệu, mà còn có các công ty hoạt động trong ngành công nghiệp vận tải, nhóm ngoại vi gồm cổ phiếu thuộc các ngành năng lượng và hàng tiêu dùng. Ở giai đoạn năm 2020-2021, nhóm trung tâm lúc này bao gồm cả cổ phiếu của các công ty trong ngành chăm sóc sức khỏe, nhóm ngoại vi gồm các công ty thuộc ngành tiêu dùng và năng lượng. Sự thay đổi của cấu trúc thị trường trong giai đoạn khủng hoảng cũng đã được tìm thấy trong nghiên cứu của Uechi & cộng sự (2015); Li & cộng sự (2019) và Millington & Niranjan (2021).

Hạn chế của nghiên cứu là việc xây dựng đồ thị mạng và danh mục đầu tư được thực hiện một cách rời rạc trên dữ liệu của ba giai đoạn nên chưa cập nhật kịp thời được sự chuyển biến của thị trường. Điều này sẽ được khắc phục ở bài báo tiếp theo, khi nhóm tác giả làm với bộ dữ liệu lớn hơn, đồ thị mạng được xây dựng lại sau mỗi ngày giao dịch, sự thay đổi của cấu trúc thị trường vì thế được cập nhật nhanh hơn. Từ đó tìm ra nhóm cổ phiếu mà danh mục được tạo từ nhóm này mang lại hiệu quả cao một cách ổn định.

Tài liệu tham khảo

- Aslam, F., Mohmand, Y. T., Ferreira, P., Memon, B. A., Khan, M. & Khan, M. (2020), 'Network analysis of global stock markets at the beginning of the coronavirus disease (Covid-19) outbreak', *Borsa Istanbul Review*, 20, S49-S61.
- Boginski, V., Butenko, S. & Pardalos, P. M. (2005), 'Statistical analysis of financial networks', *Computational statistics & data analysis*, 48(2), 431-443.
- Hoàng Đức Mạnh (2013), 'Phân tích sự phụ thuộc của các chuỗi lợi suất tài sản-Tiếp cận bằng mô hình hồi quy phân vị và phương pháp Copula', *Kỷ yếu Hội thảo Đào tạo và Ứng dụng Toán học trong Kinh tế - Xã hội*, Đại học Kinh tế Quốc dân, Hà Nội, 311-321.
- Li, B., Sun, Y., Aw, G. & Teo, K. L. (2019), 'Uncertain portfolio optimization problem under a minimax risk measure', *Applied Mathematical Modelling*, 76, 274-281.
- Mantegna, R. N. (1999), 'Hierarchical structure in financial markets', *The European Physical Journal B-Condensed Matter and Complex Systems*, 11(1), 193-197.
- Millington, T. & Niranjana, M. (2021), 'Stability and similarity in financial networks—How do they change in times of turbulence?', *Physica A: Statistical Mechanics and its Applications*, 574, 126016, DOI:10.1016/j.physa.2021.126016.
- Moslehpour, M., Al-Fadly, A., Ehsanullah, S., Chong, K. W., Xuyen, N. T. M. & Tan, L. P. (2022), 'Assessing financial risk spillover and panic impact of COVID-19 on European and Vietnam stock market', *Environmental Science and Pollution Research*, 29(19), 28226-28240.
- Nanda, S., Mahanty, B. & Tiwari, M. (2010), 'Clustering Indian stock market data for portfolio management', *Expert Systems with Applications*, 37(12), 8793-8798.
- Nguyen, Q., Nguyen, N. & Nguyen, L. (2019), 'Dynamic topology and allometric scaling behavior on the Vietnamese stock market', *Physica A: Statistical Mechanics and its Applications*, 514, 235-243.
- Onnela, J.-P., Chakraborti, A., Kaski, K., Kertesz, J. & Kanto, A. (2003), 'Dynamics of market correlations: Taxonomy and portfolio analysis', *Physical Review E*, 68(5), 056110, DOI: <https://doi.org/10.1103/PhysRevE.68.056110>.
- Pai, G. V. & Michel, T. (2009), 'Evolutionary optimization of constrained k -means clustered assets for diversification in small portfolios', *IEEE Transactions on Evolutionary Computation*, 13(5), 1030-1053.
- Peralta, G. & Zareei, A. (2016), 'A network approach to portfolio selection', *Journal of Empirical Finance*, 38, 157-180.
- Pozzi, F., Di Matteo, T. & Aste, T. (2008), 'Centrality and peripherality in filtered graphs from dynamical financial correlations', *Advances in Complex Systems*, 11(06), 927-950.
- Pozzi, F., Di Matteo, T. & Aste, T. (2012), 'Exponential smoothing weighted correlations', *The European Physical Journal B*, 85(6), 1-21.
- Pozzi, F., Di Matteo, T. & Aste, T. (2013), 'Spread of risk across financial markets: better to invest in the peripheries', *Scientific Reports*, 3(1), 1-7.
- Tola, V., Lillo, F., Gallegati, M. & Mantegna, R. N. (2008), 'Cluster analysis for portfolio optimization', *Journal of Economic Dynamics and Control*, 32(1), 235-258.
- Trần Trọng Nguyên & Nguyễn Thu Thủy (2017), 'Ứng dụng nghiên cứu cấu trúc phụ thuộc giữa thị trường ngoại hối và thị trường chứng khoán Việt Nam trong đo lường rủi ro - Tiếp cận bằng phương pháp Copula', *Kinh tế và Phát triển*, 238(II), 31-40.

-
- Tu, C. (2014), 'Cointegration-based financial networks study in Chinese stock market', *Physica A: Statistical Mechanics and its Applications*, 402, 245-254.
- Tumminello, M., Aste, T., Di Matteo, T. & Mantegna, R. N. (2005), 'A tool for filtering information in complex systems', *Proceedings of the National Academy of Sciences*, 102(30), 10421-10426.
- Uechi, L., Akutsu, T., Stanley, H. E., Marcus, A. J. & Kenett, D. Y. (2015), 'Sector dominance ratio analysis of financial markets', *Physica A: Statistical Mechanics and its Applications*, 421, 488-509.
- Zhao, L., Li, W. & Cai, X. (2016), 'Structure and dynamics of stock market in times of crisis', *Physics Letters A*, 380(5-6), 654-666.