TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**KHÓA LUẬN TỐT NGHIỆP**

**HIỆN THỰC VÀ SO SÁNH CÁC MÔ HÌNH TOÁN HỌC CỦA THỊ TRƯỜNG CHỨNG KHOÁN**

**Đây là trang bìa (in giấy cứng khi nộp phản biện / in mạ vàng khi nộp lần cuối)**

*Người hướng dẫn*: **TS NGUYỄN CHÍ THIỆN**

*Người thực hiện*: **HỒ ĐẮC NGHĨA**

**TRẦN QUỐC HUY**

Lớp **: 14050301**

Khoá  **: 18**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2018**

TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**KHÓA LUẬN TỐT NGHIỆP**

**HIỆN THỰC VÀ SO SÁNH CÁC MÔ HÌNH TOÁN HỌC CỦA THỊ TRƯỜNG CHỨNG KHOÁN**

*Người hướng dẫn*: **TS NGUYỄN CHÍ THIỆN**

*Người thực hiện*: **HỒ ĐẮC NGHĨA**

**TRẦN QUỐC HUY**

Lớp **: 14050301**

Khoá  **: 18**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2018**

LỜI CẢM ƠN

*Chúng em chân thành cám ơn thầy Nguyễn Chí Thiện đã tận tình hướng dẫn, chỉ bảo Chúng em và giải đáp những thắc mắc trong suốt quá trình làm việc. Giúp chúng em có thể hoàn thành luận văn này*.

**CÔNG TRÌNH ĐƯỢC HOÀN THÀNH**

**TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi và được sự hướng dẫn khoa học của TS Nguyễn Chí Thiện;. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

Ngoài ra, trong luận văn còn sử dụng một số nhận xét, đánh giá cũng như số liệu của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

**Nếu phát hiện có bất kỳ sự gian lận nào tôi xin hoàn toàn chịu trách nhiệm về nội dung luận văn của mình.** Trường đại học Tôn Đức Thắng không liên quan đến những vi phạm tác quyền, bản quyền do tôi gây ra trong quá trình thực hiện (nếu có).

*TP. Hồ Chí Minh, ngày tháng năm*

*Tác giả*

*(ký tên và ghi rõ họ tên)*

*Hồ Đắc Nghĩa*

*Trần Quốc Huy*

TÓM TẮT

Thị trường chứng khoán là một trong những yếu tố cơ bản của nền kinh tế thị trường hiện đại. Mọi biến động về kinh tế, chính trị, xã hội… sẽ tác động ngay trên thị trường chứng khoán và cứ nhìn vào chỉ số giá chứng khoán người ta có thể thấy rõ mức ảnh hưởng ấy tác động như thế nào. Ở Việt Nam hiện nay, thị trường chứng khoán là thị trường sôi động, tin tức chứng khoán luôn được đưa lên hàng đầu trên các phương tiện truyền thông và đầu tư chứng khoán trở thành hoạt động phát triển mạnh mẽ được nhiều người quan tâm. Những nhà đầu tư chứng khoán cũng chính là đem tiền tham gia vào hoạt động kinh doanh của công ty và hy vọng sau một thời gian sẽ nhận được lợi nhuận. Nhưng liệu việc kinh doanh này có đạt được mục đích như mong muốn không khi mà công ty bạn mua chứng khoán làm ăn thua lỗ, phá sản …và giá cổ phiếu giảm đáng kể. Đầu tư chứng khoán kiếm lời là một hoạt động hấp dẫn thu hút các nhà đầu tư, nhưng đầu tư chứng khoán như thế nào để đạt được lợi nhuận mục tiêu với mức rủi ro thấp nhất. Vì thế chúng tôi nên thiết lập và quản lý danh mục đầu tư chứng khoán.

Quản lý danh mục đầu tư là một trong những vấn đề nóng nhất trong lĩnh vực tài chính. Nó chủ yếu là mối quan tâm đến sự kết hợp tốt nhất các chứng khoán đem lại lợi nhuận cho các nhà đầu tư. Và để đạt được lợi nhuận tốt nhất thì các nhà đầu tư phải tối ưu hoá danh mục đầu tư. Việc tối ưu hóa cụ thể là lựa chọn phân bổ đầu tư của các nhà đầu tư thường dựa trên phân tích phương sai. Trong luận văn này chúng tôi tìm hiểu các mô hình toán học của chứng khoán, cài đặt thuật toán và so sánh các mô hình toán học của thị trường chứng khoán.

MỤC LỤC

[LỜI CẢM ƠN 3](#_Toc520213586)

[TÓM TẮT 5](#_Toc520213587)

[MỤC LỤC 1](#_Toc520213588)

[DANH MỤC KÍ HIỆU VÀ CHỮ VIẾT TẮT 3](#_Toc520213589)

[DANH MỤC CÁC BẢNG BIỂU, HÌNH VẼ, ĐỒ THỊ 4](#_Toc520213590)

[CHƯƠNG 1 – MỞ ĐẦU 5](#_Toc520213591)

[1.1 Đặt vấn đề 5](#_Toc520213592)

[1.2 Mục tiêu và phạm vi nghiên cứu 5](#_Toc520213593)

[1.2.1 Mục tiêu nghiên cứu 5](#_Toc520213594)

[1.2.2 Phạm vi nghiên cứu 6](#_Toc520213595)

[CHƯƠNG 2 – TỔNG QUAN 7](#_Toc520213596)

[2.1 Thu nhập dữ liệu 7](#_Toc520213597)

[2.2 Phương pháp xử lý dữ liệu thu được 9](#_Toc520213598)

[2.2.1 Chiến lược danh mục động sử dụng giải thuật Clustering 9](#_Toc520213599)

[2.2.2 Lây lan rủi ro trên thị trường chứng khoán 10](#_Toc520213600)

[2.2.3 Phân tích loại bỏ xu hướng tương quan chéo phụ thuộc q của thị trường chứng khoán 12](#_Toc520213601)

[2.3 Trình tự thực hiện nghiên cứu 12](#_Toc520213602)

[CHƯƠNG 3 – NHỮNG NGHIÊN CỨU THỰC NGHIỆM HOẶC LÝ THUYẾT 14](#_Toc520213603)

[3.1 Chiến lược danh mục động sử dụng giải thuật Clustering 14](#_Toc520213604)

[3.1.1 Dữ liệu và phương thức 14](#_Toc520213605)

[3.1.2 Xây dựng mạng dựa trên thuật toán cây khung nhỏ nhất (MST) 15](#_Toc520213606)

[3.1.3 Lựa chọn danh mục dựa trên các tham số tô pô 16](#_Toc520213607)

[3.1.4 Xác định thời gian đầu tư 18](#_Toc520213608)

[3.1.5 Xác định điều kiện thị trường 18](#_Toc520213609)

[3.1.7 Kết luận 20](#_Toc520213610)

[3.2 Lây lan rủi ro trên thị trường chứng khoán 21](#_Toc520213611)

[3.2.1 Giới thiệu 21](#_Toc520213612)

[3.2.2 Phương pháp 25](#_Toc520213613)

[3.3 Phân tích loại bỏ xu hướng tương quan chéo phụ thuộc q của thị trường chứng khoán 27](#_Toc520213614)

[3.3.2 Giới thiệu 27](#_Toc520213615)

[3.3.3 Phương thức 29](#_Toc520213616)

[3.3.5 Ứng dụng 32](#_Toc520213617)

[CHƯƠNG 4 – TRÌNH BÀY, ĐÁNH GIÁ, BÀN LUẬN CÁC KẾT QUẢ 35](#_Toc520213618)

[4.1 Chiến lược danh mục động sử dụng giải thuật Clustering 35](#_Toc520213619)

[4.2 Lây lan rủi ro trên thị trường chứng khoán 40](#_Toc520213620)

[4.3 Phân tích loại bỏ xu hướng tương quan chéo phụ thuộc q của thị trường chứng khoán 43](#_Toc520213621)

[CHƯƠNG 5 – KẾT LUẬN 52](#_Toc520213622)

DANH MỤC KÍ HIỆU VÀ CHỮ VIẾT TẮT

**CÁC KÝ HIỆU**

*E[ ] biểu diễn kỳ vọng toán học của dãy theo khoảng thời gian t*

*Giá đóng cửa của cổ phiếu i vào ngày thứ t*

*ri lợi nhuận của các cổ phiếu i*

*rj lợi nhuận của các cổ phiếu j*

**CÁC CHỮ VIẾT TẮT**

PMFG Planst Maximumally Filtered Graphs

MST Minimum Spanning Tree

DANH MỤC CÁC BẢNG BIỂU, HÌNH VẼ, ĐỒ THỊ

**DANH MỤC HÌNH**

[Hình 1.1 Trang web cophieu68.vn 7](file:///C:\Users\love_\LuanVan\word\51403009-51403280_LuanVan.docx#_Toc520245410)

[Hình 1.2 Trang web finance.yahoo.com 7](file:///C:\Users\love_\LuanVan\word\51403009-51403280_LuanVan.docx#_Toc520245411)

[Hình 3.1 Mạng lưới danh mục đầu tư sàn Hà Nội 16](#_Toc520245447)

[Hình 3.2 Mạng lưới danh mục đầu tư sàn Hồ Chí Minh 16](#_Toc520245448)

**DANH MỤC BẢNG**

[Bảng 3.1 Kết quả đầu tư của 20 thị trường theo điều kiện thị trường 24](#_Toc518817688)

CHƯƠNG 1 – MỞ ĐẦU

1.1 Đặt vấn đề

Thị trường chứng khoán là một trong những yếu tố cơ bản của nền kinh tế thị trường hiện đại. Mọi biến động về kinh tế, chính trị, xã hội sẽ tác động ngay trên thị trường chứng khoán và cứ nhìn vào chỉ số giá chứng khoán người ta có thể thấy rõ mức ảnh hưởng ấy tác động như thế nào. Ở Việt Nam hiện nay, thị trường chứng khoán là thị trường sôi động, tin tức chứng khoán luôn được đưa lên hàng đầu trên các phương tiện truyền thông và đầu tư chứng khoán trở thành hoạt động phát triển mạnh mẽ được nhiều người quan tâm. Những nhà đầu tư chứng khoán cũng chính là những người sử dụng tài sản tham gia vào hoạt động kinh doanh của công ty và hy vọng sau một thời gian sẽ nhận được giá trị cao hơn. Nhưng liệu việc kinh doanh này có đạt được mục đích như mong muốn không khi mà công ty bạn mua chứng khoán làm ăn thua lỗ, phá sản và giá cổ phiếu giảm đáng kể. Đầu tư chứng khoán kiếm lời là một hoạt động hấp dẫn thu hút các nhà đầu tư, nhưng đầu tư chứng khoán như thế nào để đạt được lợi nhuận mục tiêu với mức rủi ro thấp nhất. Vì lý do trên nên đã có rất nhiều mô hình toán học của thị trường chứng khoán được ra đời và được áp dụng cho rất nhiều thị trường khác nhau. Vậy một mô hình toán học đó có đúng trên tất cả các thị trường và đều đem lại cho chúng tôi lợi nhuận tốt nhất. Đó là lý do chúng tôi nghiên cứu đề tài này: “Hiện thực và so sánh các mô hình toán học trên thị trường chứng khoán”.

1.2 Mục tiêu và phạm vi nghiên cứu

1.2.1 Mục tiêu nghiên cứu

Tìm hiểu các mô hình toán học của thị trường chứng khoán.

Cài đặt các môn hình toán học của thị trường chứng khoán và các thuật toán giao dịch trên các mô hình.

Thử nghiệm thực nghiêm các mô hình toán học trên nhiều thị trường để xác minh độ chính xác của thuật toán và mô hình có thể áp dụng trên mọi thị trường được hay không.

So sánh các mô hình toán học của thị trường chứng khoán và phát triển các mô hình này để đạt được lợi nhuận cao nhất có thể áp dụng cho nhiều thị trường.

1.2.2 Phạm vi nghiên cứu

Thị trường:

* Các thuật toán sẽ được áp dụng trển nhiều thị trường chứng khoán trên toàn thế giới như: Việt Nam, Mỹ, Na Uy, Hàn Quốc, Nhật Bản, Úc…

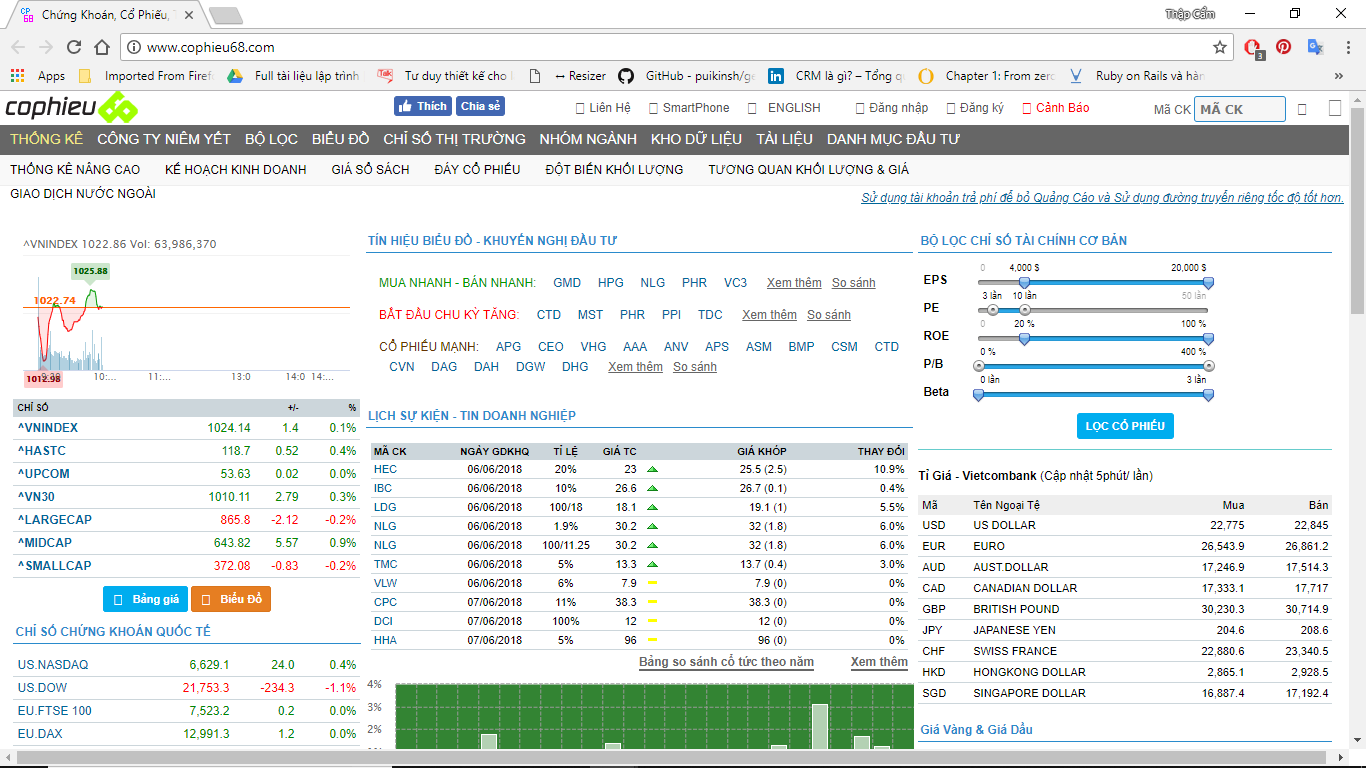
Tập đầu tư:

* Các tập đầu tư gồm các cổ phiếu của các công ty đã được niêm yết.
* Dữ liệu nghiên cứu trong khoảng thời gian: từ tháng 4 năm 2015 đến tháng 4 năm 2018.

CHƯƠNG 2 – TỔNG QUAN

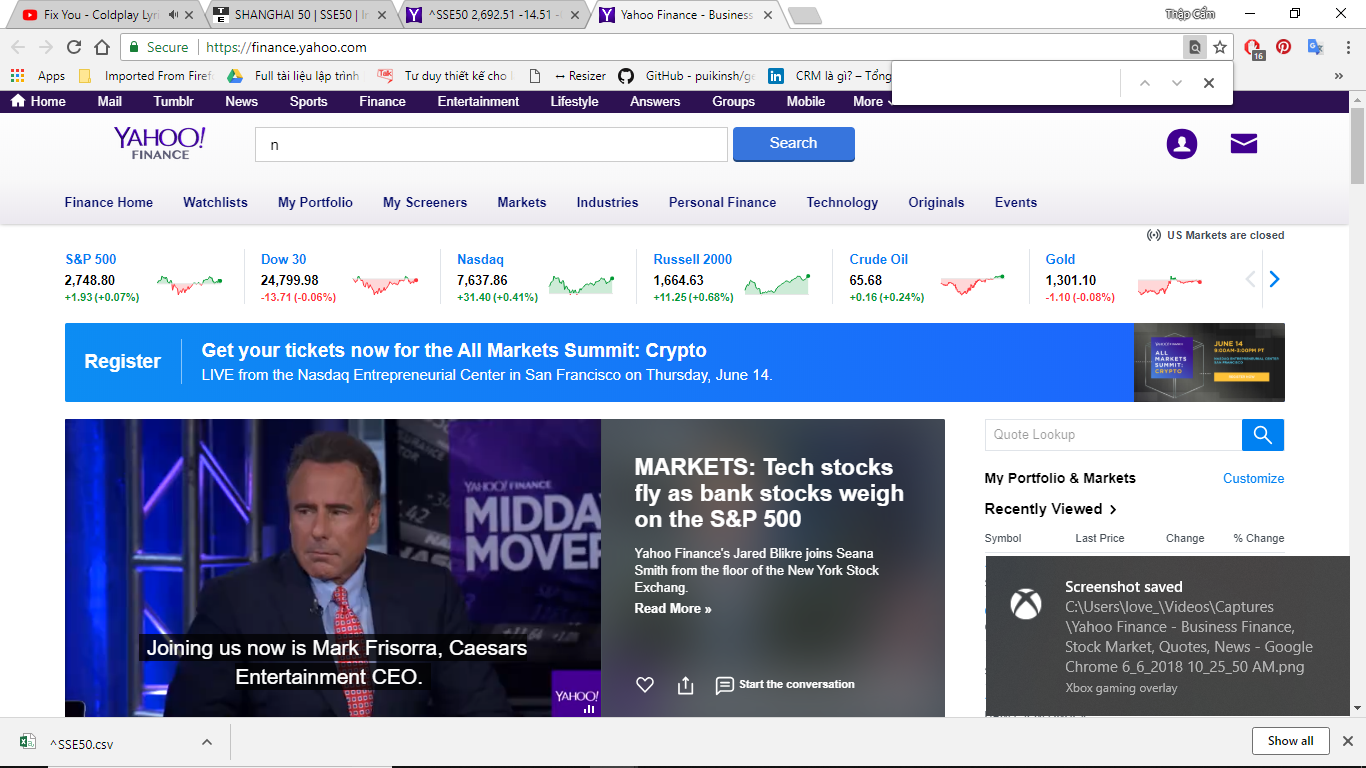
2.1 Thu nhập dữ liệu

Thu thập dữ liệu chứng khoán từ các công ty được niên yết theo các sàn.

* + Các cổ phiếu của các sàn ở Việt Nam chúng tôi thu thập từ trang web: <http://www.cophieu68.com>

Hình 1. Trang web cophieu68.vn

* + Các cổ phiếu quốc tể chúng tôi thu nhập từ trang web:

<http://www.finance.yahoo.com>

Hình 1. Trang web finance.yahoo.com

2.2 Phương pháp xử lý dữ liệu thu được

2.2.1 Chiến lược danh mục động sử dụng giải thuật Clustering

Phân cụm theo thứ bậc là một trong những phương pháp hiệu quả nhất để chọn một tập cổ phiếu cho danh mục đầu tư tối ưu. Trên thực tế, việc lựa chọn một bộ cổ phiếu là một điều kiện tiên quyết cho lý thuyết tối ưu danh mục đầu tư.

Bằng cách sử dụng phương pháp phân cụm theo thứ bậc, mối quan hệ giữa các cổ phiếu được mạng lưới xây dựng cổ phiếu biểu diễn bởi cấu trúc tô pô[1] và nó áp dụng cho tối ưu hóa danh mục đầu tư. Mô tả về mối tương quan giữa các cổ phiếu gồm những cổ phiếu có độ rủi ro ít nhất trong danh mục đầu tư có xu hướng nằm ngoài của mạng lưới giao dịch bằng cách sử dụng cây khung nhỏ nhất MST (MST). Bằng cách khai thác cấu trúc phụ thuộc của các cổ phiếu tài chính bao gồm cả thuật toán MST. Người ta thấy rằng danh mục đầu tư từ các cổ phiếu ngoại vi có rủi ro thấp hơn và lợi nhuận cao hơn danh mục đầu tư từ các cổ phiếu trung tâm. Tính trung tâm / ngoại vi được đo bằng các chỉ số như bậc,độ trung tâm trung gian, độ lệch tâm, độ chặt chẽ và độ đặc trưng trung tâm. Thuật toán phân cụm dữ K-means và thuật toán phân cụm dữ liệu dùng để phân loại chứng khoán. Các cổ phiếu được lựa chọn từ các nhóm phân loại này được sử dụng để xây dựng danh mục đầu tư. Phương pháp này tương tự như thuật toán phân cụm cho sự giống nhau trong các lựa chọn cổ phiếu từ các cụm hoặc cộng đồng được phân chia theo các cách tiếp cận cụ thể.

Động cơ chính của nghiên cứu này là đề xuất một chiến lược đầu tư mới năng động dựa trên cơ cấu thời gian của các mạng lưới lọc tài chính trên các thị trường chứng khoán. Một khung di chuyển với kích thước δt được sử dụng để nghiên cứu sự khác biệt của mạng lưới chứng khoán theo khoảng thời gian t. Phương pháp MST để lọc ra đồ thị mạng trong mỗi khung vì tính chính xác và đơn giản của nó, được tạo ra bằng cách kết nối các nút với các tương quan quan trọng nhất. Việc lựa chọn danh mục đầu tư được xác định bởi cấu trúc mạng trong khung trước đó (vùng lựa chọn), được chọn từ một số cổ phiếu ngoại vi, đa dạng nhất tương ứng với danh mục Markowitz với độ lệch tối thiểu và các cổ phiếu trung tâm, tương quan cao và đồng bộ trong diễn biến giá cả. Danh mục đầu tư được lựa chọn sau đó được sử dụng để đầu tư vào khoảng thời gian đầu tư.

Các điều kiện thị trường cơ bản được xem xét thêm trong chiến lược danh mục động, bao gồm chiến lược đầu tư cùng với việc lựa chọn danh mục đầu tư. Trong nghiên cứu ở đây, Họ cho rằng tối ưu danh mục đầu tư có thể thay đổi theo các điều kiện thị trường khác nhau và đầu tư theo danh mục đầu tư được thực hiện dựa trên sự thay đổi giá cả và giá dự đoán trong tương lai. Để đơn giản, ba điều kiện thị trường: xu hướng giảm, tăng và ổn định của giá hàng ngày sẽ được sử dụng trong việc lựa chọn các khoảng thời gian đầu tư. Một loạt các danh mục được lựa chọn sẽ được so sánh dưới sự kết hợp khác nhau của điều kiện thị trường trong hai phạm vi, và danh mục đầu tư tối ưu có lợi nhuận lớn nhất sẽ được xác định theo từng điều kiện thị trường. Để chứng minh hiệu quả chiến lược danh mục động của họ, chúng tôi sẽ tiến hành quá trình huấn luyện sử dụng nửa đầu của dữ liệu mẫu và sử dụng danh mục đầu tư tối ưu từ huấn luyện để đầu tư sử dụng nửa còn lại của dữ liệu mẫu.

2.2.2 Lây lan rủi ro trên thị trường chứng khoán

Rủi ro không được lan truyền đồng bộ trên thị trường chứng khoán và với thực tế này chúng tôi có thể khai thác để giảm rủi ro đầu tư góp phần cải thiện lợi nhuận. Bằng cách trích xuất cấu trúc phụ thuộc của cổ phiếu tài chính có hiệu quả làm giảm rủi ro đầu tư. Họ cho rằng các khoản đầu tư vào các khu vực ngoại vi, các cổ phiếu kém kết nối trong mạng lọc chứng khoán sẽ cải thiện lợi nhuận. Ngược lại, đầu tư vào tập hợp con của các cổ phiếu trung tâm,các cổ phiếu có nhiều kết nối cao được đặc trưng bởi rủi ro lớn hơn và hình thành nhiều rủi ro hơn. Cụ thể là sử dụng cây bao trùm nhỏ nhất (Spanning Trees) và đồ thị phẳng (Planar Maximally Filtered Graphs), phương pháp thành công nhất trong việc đa dạng hóa, cải thiện tỷ lệ giữa độ lệch trung bình và độ lệch chuẩn, giảm khả năng tiêu cực. Phương pháp này có lợi thế hơn trong việc hình dung danh mục đầu tư.

Xây dựng một mạng lọc tài chính, cụ thể là Spanning Trees (MST) và Planar Maximally Filtered Graphs (PMFG) có thể được sử dụng để mô tả sự lây lan không đồng nhất của rủi ro trên thị trường tài chính. giảm rủi ro đầu tư bằng cách xây dựng danh mục đầu tư đa dạng. MST là một cây bao trùm (một mạng được kết nối không có vòng lặp hoặc chu kỳ) tối đa hóa tổng các mối tương quan trên các kết nối trong cây. Tương tự, PMFG là biểu đồ phẳng chứa MST dưới dạng đồ thị con và giữ lại mối tương quan lớn nhất trên các cạnh. Các cổ phiếu được định vị khác nhau trong mạng lọc tài chính thể hiện các kiểu hành vi khác nhau nên việc lựa chọn cổ phiếu từ đa số các vùng thay thế của mạng có thể được sử dụng để thiết lập danh mục đa dạng hiệu quả.

Chúng tôi tính tương quan qua khung thời gian sáu tháng, giảm ảnh hưởng quá mức của khủng hoảng kinh tế từ xa về tương quan hiện tại bằng cách sử dụng phương pháp tiên tiến lũy thừa 8. Sau đó, chúng tôi đã cải thiện bộ ước lượng bằng cách tính toán ma trận tương quan trung bình với độ rút gọn trong khoảng thời gian sáu tháng thu được theo cách này một ước tính mạnh mẽ về các mối tương quan so với năm trước ngày đầu tư t. Sau đó chúng tôi sử dụng các mối tương quan có trọng số trung bình này với độ co rút để xây dựng các mạng lọc tài chính MST và PMFG. Chúng tôi thiết lập danh mục đầu tư bằng cách chọn cổ phiếu từ khu vực ngoại vi của mạng được lọc tài chính và so sánh hiệu suất của các danh mục này với hiệu suất của danh mục được thiết lập bằng cách chọn cổ phiếu trung tâm hoặc cổ phiếu ngẫu nhiên hoặc bằng các phương pháp truyền thống khác. Với mục đích này, trước tiên chúng tôi phân biệt giữa các cổ phiếu nằm trong vùng trung tâm của mạng và những cổ phiếu nằm ở vùng ngoại vi. Bằng cách sử dụng phương pháp phân loại cổ phiếu theo thứ bậc, mối quan hệ giữa các cổ phiếu được mạng lưới xây dựng cổ phiếu biểu diễn bởi cấu trúc tô pô và nó áp dụng cho việc đầu tư theo cổ phiếu ngoại vi.

2.2.3 Phân tích loại bỏ xu hướng tương quan chéo phụ thuộc q của thị trường chứng khoán

Các tính chất của ma trận tương quan chéo q-dependent của thị trường chứng khoán đã phân tích được bằng cách sử dụng lý thuyết ma trận ngẫu nhiên và mạng lưới phức tạp cho thấy cấu trúc tương quan của các dao động ở các độ lớn khác nhau có tính chất duy nhất. Các tương quan chéo giữa các biến động nhỏ tác động lớn hơn nhiều so với các biến động lớn. Các biến động lớn và nhỏ bị chi phối bởi các nhóm cổ phiếu khác nhau. Chúng tôi sử dụng biểu diễn mạng phức tạp để nghiên cứu các ma trận q-dependent này và khám phá một số nhận dạng mới. Bằng cách sử dụng các mạng tương quan q-dependent đó, chúng tôi có thể xây dựng một số danh mục đầu tư bởi những cổ phiếu độc lập nhất luôn hoạt động tốt nhất. Thứ tự đa chiều tối ưu cho tối ưu hóa danh mục đầu tư là khoảng q = 2.

Xử lý dữ liệu sau khi thu được theo các phương pháp của từng mô hình toán học của thị trường chứng khoán. Ở đây chúng tôi thực hiện 2 phương pháp:

* + Chiến lược danh mục đầu tư động sử dụng phương pháp Clustering
  + Giải quyết các mô hình lựa chọn danh mục đầu tư với lợi nhuận không chắc chắn bằng cách sử dụng một sơ đồ mạng nơron nhân tạo

2.3 Trình tự thực hiện nghiên cứu

Nghiên cứu lý thuyết về đầu tư chứng khoán, các từ ngữ chuyên ngành.

Thu thập dữ liệu chứng khoán của các thị trường.

Cài đặt các môn hình toán học của thị trường chứng khoán và các thuật toán giao dịch trên đó.

Cải tiến các mô hình để tăng lợi nhuận.

So sánh các mô hình toán học của thị trường chứng khoán để lựa chọn ra mô hình thích hợp.

CHƯƠNG 3 – NHỮNG NGHIÊN CỨU THỰC NGHIỆM HOẶC LÝ THUYẾT

3.1 Chiến lược danh mục động sử dụng giải thuật Clustering

3.1.1 Dữ liệu và phương thức

Dữ liệu hàng ngày của chúng tôi lấy trong khoảng thời gian 3 năm từ ngày 3 tháng 1 năm 2015 đến ngày 3 tháng 1 năm 2018. Để đảm bảo sự liên tục và toàn vẹn của dữ liệu, các cổ phiếu được lựa chọn trong nghiên cứu là cổ phiếu có cổ phần tích cực nhất trong giao dịch suốt giai đoạn lấy mẫu. Với mục đích này, chúng tôi loại bỏ những cổ phiếu đã bị đình chỉ thị trường trong hơn 30 ngày giao dịch và có giao dịch suốt 8 năm không ít hơn. Dòng trả về của một cổ phiếu nhất định i được tính bằng:

Trong đó:

* : là giá đóng cửa của cổ phiếu i vào ngày thứ t.

Giá của các cổ phiếu được tính toán và các ảnh hưởng của các hành động công ty được loại bỏ, ví dụ như cổ tức bằng tiền mặt, phần thưởng và vấn đề quyền.

Vào một ngày nhất định, một ma trận tương quan được tính bằng cách sử dụng bộ ước lượng hệ số tương quan Pearson trên chuỗi trả về trong khung {t - δt + 1, ..., t} và mạng chứng khoán được xây dựng bằng phương pháp MST. Phân loại cổ phiếu trả về thành 10% loại danh mục đầu tư được lựa chọn tương ứng từ 10% của hầu hết các chứng khoán trung tâm và ngoại vi trong đồ thị MST, trung tâm / ngoại vi của nó được đo bằng năm tham số ghi lại topo mạng: bậc, độ trung tâm trung gian, khoảng cách về bậc, khoảng cách về tương quan và khoảng cách về khoảng cách

Các danh mục được lựa chọn được sử dụng để đầu tư vào khoảng thời gian sau (t + 1, ..., t + Δt}, với trọng số bằng nhau cho mỗi cổ phiếu đã được chọn trong đó chiến lược danh mục đầu tư 1 / N được chứng minh là hiệu quả hơn mô hình phương sai trung bình. Thu nhập đầu tư của danh mục đầu tư được lựa chọn được tính toán dưới chín kết hợp của điều kiện thị trường trong việc lựa chọn và đầu tư. Các điều kiện thị trường bao gồm tăng, giảm và ổn định được xác định theo các tiêu chuẩn ngày giao dịch.

Sau đó di chuyển đến t + φ. Một chiến lược danh mục giống nhau được thông qua bằng cách chọn danh mục đầu tư trong khung (đường thời gian) {t + φ - δt + 1, ..., t + φ} và sau đó sử dụng danh mục đầu tư đã chọn để đầu tư vào đường thời gian {t + φ + 1, ..., t + φ + Δt}. Lợi nhuận đầu tư của 10% loại danh mục đầu tư được lựa chọn được tính toán dưới kết hợp của điều kiện thị trường trong hai phạm vi, và tối ưu danh mục đầu tư bằng cách đánh giá hiệu suất trung bình cho các khung di chuyển khác nhau.

Sự lựa chọn thích hợp của δt và φ thực sự sẽ giúp mạng nắm bắt được thông tin của dữ liệu ban đầu càng nhiều càng tốt. Δt lớn hơn và nhỏ hơn φ thì cấu trúc mạng ổn định hơn, và càng có nhiều thông tin thị trường được lọc ra. Ngược lại, cấu trúc mạng dễ bay hơi hơn và không đúng , mặc dù sự dao động trong khoảng thời gian có thể dễ dàng nhận thấy. Chúng tôi chọn δt = 10 tháng (≈ 200 ngày) và φ = 1 tháng (≈ 20 ngày) bằng cách quan sát cẩn thận và tính toán chính xác.

3.1.2 Xây dựng mạng dựa trên thuật toán cây khung nhỏ nhất (MST)

Sử dụng hệ số tương quan Pearson giữa các chuỗi trả về của chúng bằng công thức:

Trong đó:

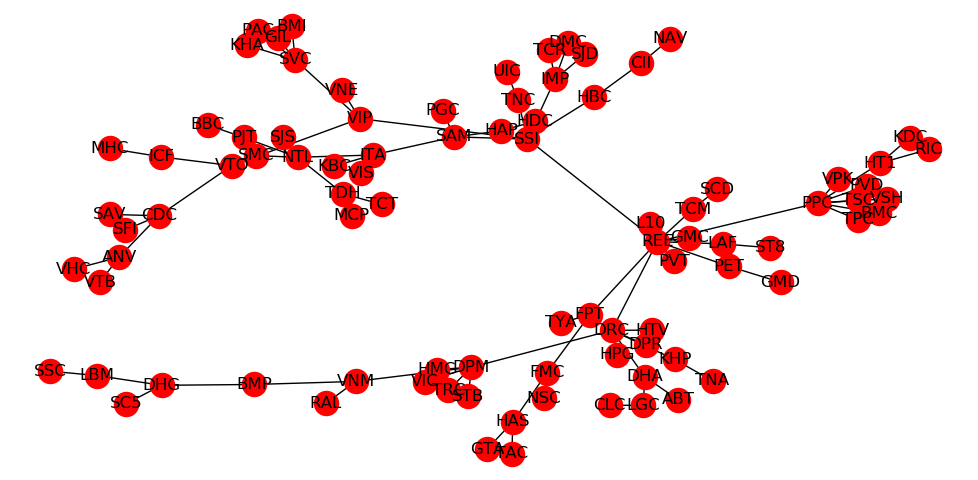
* + E[ ]: biểu diễn kỳ vọng toán học của dãy theo khoảng thời gian t.
  + ri: lợi nhuận của các cổ phiếu i.
  + rj: lợi nhuận của các cổ phiếu j

Trước khi xây dựng biểu đồ MST, hệ số tương quan được chuyển đổi thành khoảng cách giữa các cổ phiếu i và j theo phương trình sau:

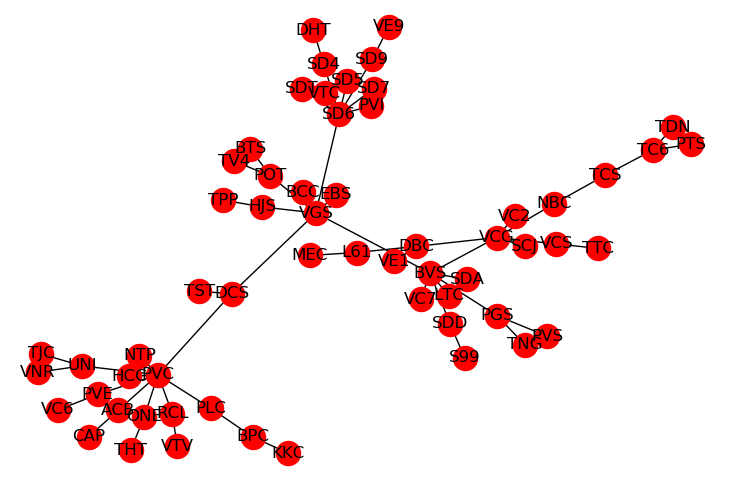
Khoảng cách d (i, j) dao động từ 0 đến 2, và một khoảng cách nhỏ tương ứng với hệ số tương quan lớn. Đối với kho dữ liệu mẫu tại từng thị trường chứng khoán, mỗi một thị trường thu được ma trận khoảng cách. Việc ước lượng ma trận tương quan đã không thể tránh khỏi với sự không chắc chắn về mặt thống kê do độ dài hữu hạn của dòng hồi tiếp cũng như các thông tin dư thừa.

Chúng tôi chọn phương pháp MST để lọc ra các đồ thị mạng trong mỗi khung để loại bỏ thông tin dư thừa nhưng vẫn giữ các liên kết quan trọng. Bằng cách xây dựng cây khung tối thiểu giúp giảm không gian thông tin từ n (n - 1) / 2 cạnh thành n - 1 cạnh. Nói cách khác, lượng thông tin đang được nén đáng kể. Thủ tục xây dựng mạng MST có thể được thực hiện như sau:

* Thứ nhất, bố trí khoảng cách giữa tất cả các cặp cổ phiếu theo thứ tự tăng dần.
* Thứ hai, bắt đầu bằng cách kết hợp các nút gần nhất. Tiếp tục nối các nút theo danh sách đã sắp xếp khi và chỉ khi đồ thị thu được sau khi kết hợp vẫn còn là một cây. Các cạnh tối đa hóa tổng các mối tương quan giữa các kết nối trong cây có nhiều khả năng không thay đổi bằng phương pháp này. Các nghiên cứu như tài liệu tham khảo đã sử dụng mô hình MST để lọc mạng. Với dữ liệu được sử dụng trong nghiên cứu của chúng tôi, chúng tôi chọn thuật toán Prim để xây dựng mạng của chúng tôi.



Hình 3. Mạng lưới danh mục đầu tư sàn Hà Nội



Hình 3. Mạng lưới danh mục đầu tư sàn Hồ Chí Minh

3.1.3 Lựa chọn danh mục dựa trên các tham số tô pô

Năm tham số được sử dụng để đo trung tâm và ngoại vi của các nút trong danh mục đầu tư lựa chọn:

* Bậc K (Degree K): số nút hàng xóm kết nối với một nút. K càng lớn thì càng có nhiều cạnh kết hợp với nút này.
* Độ trung tâm trung gian C (Betweenness centrality C), phản ánh sự đóng góp của một nút vào kết nối mạng. Gọi V là tập hợp các nút trong mạng. Đối với các nút i và j, C của một nút k có thể được tính như

trong đó là số các đường đi ngắn nhất từ nút i đến nút j, (V) là một tiểu ngành của mà các tuyến đường đi qua nút này k.

Khoảng cách dùng để chỉ độ dài ngắn nhất từ nút đến nút trung tâm của mạng. Ở đây, ba loại định nghĩa của một nút trung tâm được giới thiệu để giảm lỗi gây ra bởi một phương pháp duy nhất. Vì vậy, ba loại khoảng cách được mô tả ở đây.

* Khoảng cách dựa trên bậc (): Nút trung tâm là nút có bậc lớn nhất.
* Khoảng cách dựa trên sự tương quan (): Sự tương quan, một nút trung tâm là nút có giá trị cao nhất của tổng các hệ số tương quan với các cạnh kề của nó;
* Khoảng cách dựa trên khoảng cách (): Nút trung tâm là nút cho giá trị nhỏ nhất cho khoảng cách trung bình.

Chúng tôi sử dụng các thông số được xác định ở trên để chọn danh mục đầu tư. 10% các nút có bậc K cao nhất hoặc trung tâm trung gian C cao nhất được chọn để nằm trong danh mục trung tâm, và các nút có bậc K =1 hoặc trung tâm trung gian C = 0 được chọn để nằm trong danh mục ngoại vi.

Tương tự như vậy, chúng tôi xác định 10% các nút trong đầu trang theo khoảng cách là cổ phiếu của các danh mục ngoại vi, và 10% cuối trang là cổ phiếu của danh mục trung tâm.

Danh mục đầu tư trung tâm và các danh mục ngoại vi đại diện cho hai mặt đối diện của mối tương quan và kết tụ. Nói chung, cổ phiếu trung tâm đóng một vai trò quan trọng trên thị trường và gây áp lực mạnh lên các cổ phiếu khác, trong khi mối tương quan giữa các cổ phiếu ngoại vi yếu và có nhiều thông tin dư thừa hơn các cổ phiếu trung tâm.

3.1.4 Xác định thời gian đầu tư

Trong tiểu mục này, chúng tôi sẽ thảo luận về sự lựa chọn tối ưu của độ dài của các khoảng thời gian đầu tư Δt. Nói chung, độ dài của các khoảng thời gian đầu tư không thể quá dài, nếu không các thuộc tính tô pô của mạng sẽ thay đổi và các danh mục đầu tư trung tâm hoặc ngoại vi được lựa chọn sẽ bị thay đổi theo. Mặt khác, độ dài của các khoảng thời gian đầu tư không thể quá ngắn hoặc lợi nhuận sẽ bị ảnh hưởng nhiều bởi các thông tin dư thừa của thị trường hoặc các phát sinh. Ở đây, chúng tôi so sánh lợi nhuận đạt được trong các khoảng thời là 10 tháng vì khoảng thời gian này đủ để quan sát thị trường của tất cả các thị trường cổ phiếu.

3.1.5 Xác định điều kiện thị trường

Các điều kiện thị trường, mô tả xu hướng chung của chỉ số thị trường trong một đường biên cụ thể, được đo bằng hai tiêu chí:

1. Tiêu chí ngày giao dịch: Tỷ lệ số ngày có chỉ số tăng trên tổng số ngày giao dịch trong một khoảng thời gian cụ thể được lấy bởi số ngày giá đóng cửa lớn hơn ngày của ngày hôm trước và là tổng số ngày giao dịch trong khung thời gian thứ i. Tỷ số dao động từ 0 đến 1. Một giá trị lớn của đại diện cho điều kiện tăng và giá trị nhỏ của rd đại diện cho điều kiện giảm. Với các ngưỡng θ+ và θ-, chúng tôi xác định thị trường tăng nếu rd > θ +, thị trường giảm nếu rd <θ-, và thị trường ổn định nếu θ- ≤ rd ≤ θ +.
2. Tiêu chí biên độ: Tỷ lệ rf của tổng biên độ các ngày giao dịch với tổng các biên độ tăng của tổng số ngày giao dịch trong một khung thời gian cụ thể được chọn bởi giá đóng cửa lớn hơn ngày của ngày hôm trước, Ti là tập hợp của tất cả các ngày giao dịch trong khung thời gian thứ i, và P(t) là giá đóng cửa trên t- thứ ngày. Tương tự, với các ngưỡng θ+ và θ-, chúng tôi xác định thị trường tăng nếu > θ+, thị trường giảm nếu < θ-, và một ổn định nếu θ- ≤ ≤ θ+.

Các tỷ số và theo khoảng thời gian cho thấy một số khác biệt gây thay đổi nhỏ trong việc xác định các điều kiện thị trường. Trong phương pháp này, chúng tôi chọn θ+ = 0.55, θ- = 0.45 làm ngưỡng. Đối với giá trị lớn hơn của θ+ và giá trị nhỏ hơn của θ-, số lượng các mẫu tăng hoặc giảm không phải là thống kê đầy đủ cho việc thiếu dữ liệu.

Đối với mỗi khung thời gian chúng tôi có được ba điều kiện thị trường. Do đó, chúng tôi nhận được chín sự kết hợp của các điều kiện thị trường trong khoảng thời gian nghiên cứu và khoảng thời gian đầu tư sau này:

* đường thời gian nghiên cứu tăng và đường thời gian đầu tư tăng Up-Up(UU),
* đường thời gian nghiên cứu tăng và đường thời gian đầu tư ổn định Up-Stable(US),
* đường thời gian nghiên cứu tăng và đường thời gian đầu tư giảm Up-Down(UD)
* đường thời gian nghiên cứu ổn định và đường thời gian đầu tư tăng Stable-Up(SU),
* đường thời gian nghiên cứu và đường thời gian đầu tư đều ổn định Stable-Stable(SS)
* đường thời gian nghiên cứu ổn định và đường thời gian đầu tư giảm Stable-Down(SD)
* đường thời gian nghiên cứu giảm và đường thời gian đầu tư tăng Down-Up(DU)
* đường thời gian nghiên cứu giảm và đường thời gian đầu tư ổn định Down-Stable(DS)
* đường thời gian nghiên cứu và thời gian đầu tư giảm Down-Down(DD).

3.2 Lây lan rủi ro trên thị trường chứng khoán

3.2.1 Giới thiệu

Rủi ro không được lan truyền đồng bộ trên thị trường chứng khoán và với thực tế này chúng tôi có thể khai thác để giảm rủi ro đầu tư góp phần cải thiện lợi nhuận. Bằng cách trích xuất cấu trúc phụ thuộc của cổ phiếu tài chính, một phương pháp tiếp cận mạng có thể được sử dụng để xây dựng một danh mục đầu tư đa dạng, có hiệu quả làm giảm rủi ro đầu tư. Chúng tôi thấy rằng các khoản đầu tư vào các khu vực ngoại vi, các cổ phiếu kém kết nối trong mạng chọn lọc chứng khoán, cụ thể là cây bao trùm nhỏ nhất (Spanning Trees) và đồ thị phẳng (Planar Maximally Filtered Graphs), thành công nhất trong việc đa dạng hóa, cải thiện tỷ lệ giữa độ lệch trung bình và độ lệch chuẩn, giảm khả năng tiêu cực. Ngược lại, đầu tư vào tập hợp con của các cổ phiếu trung tâm,các cổ phiếu có nhiều kết nối cao được đặc trưng bởi rủi ro lớn hơn và hình thành nhiều rủi ro hơn. Phương pháp này có lợi thế hơn trong việc hình dung danh mục đầu tư. lựa chọn trực tiếp trên bố cục đồ họa của mạng.

Quản lý rủi ro của thị trường chứng khoán bất ổn là ưu tiên hàng đầu cho ngành tài chính. Trong mô hình này, chúng tôi nghiên cứu cách xây dựng một mạng lọc tài chính, cụ thể là Spanning Trees (MST) và Planar Maximally Filtered Graphs (PMFG) có thể được sử dụng để mô tả sự lây lan không đồng nhất của rủi ro trên thị trường tài chính. giảm rủi ro đầu tư bằng cách xây dựng danh mục đầu tư đa dạng. Chúng tôi hãy nhớ rằng các mạng lọc tài chính được xây dựng bằng cách giữ lại các liên kết tương quan cao nhất trong khi hạn chế một số thuộc tính tổng thể của mạng mà không cần chỉ định bất kỳ ngưỡng nào. Cụ thể, MST là một cây bao trùm (một mạng được kết nối không có vòng lặp hoặc chu kỳ) tối đa hóa tổng các mối tương quan trên các kết nối trong cây. Tương tự, PMFG là biểu đồ phẳng chứa MST dưới dạng đồ thị con và giữ lại mối tương quan lớn nhất trên các cạnh4. Cấu trúc liên kết của các mạng này mã hóa hiệu quả cấu trúc phụ thuộc phức tạp của các cổ phiếu tài chính, giải nén các thuộc tính phân cấp và phân nhóm, giảm độ phức tạp của dữ liệu trong khi vẫn giữ các đặc tính cơ bản của tập dữ liệu. Các cổ phiếu được định vị khác nhau trong mạng lọc tài chính thể hiện các kiểu hành vi khác nhau và do đó việc lựa chọn cổ phiếu từ đa số các vùng thay thế của mạng có thể được sử dụng để thiết lập danh mục đa dạng hiệu quả.

Được chấp nhận rộng rãi kể từ lý thuyết danh mục đầu tư hiện đại của Markowitz, nhằm mục đích đa dạng hóa hiệu quả nên chọn cổ phiếu chống tương quan càng tốt và vẫn liên tục chống tương quan theo thời gian. Xác định từ nghiên cứu về hành vi lịch sử trước khi đầu tư, giỏ cổ phiếu có khả năng tốt để duy trì tốt trong giai đoạn đầu tư trong tương lai là rất khó khăn. Thật vậy, cấu trúc của mối tương quan giữa các cổ phiếu đang phát triển theo thời gian và thay đổi rõ rệt trong các cuộc khủng hoảng. Vì lý do này, cách tiếp cận Markowitz thường được áp dụng cho việc lựa chọn các cổ phiếu được xác định bằng cách sử dụng các tiêu chí khác nhau bao gồm cả khu vực công nghiệp và các cân nhắc vĩ mô hoặc kinh tế vi mô khác.

Trong mô hình này, để xác định 'giỏ' cổ phiếu trực tiếp từ cấu trúc phụ thuộc được cung cấp bởi mạng lọc tài chính. Đối với mỗi ngày thị trường t, chúng tôi điều tra hành vi của một lựa chọn N = 300 cổ phiếu có vốn hóa cao và hiệu suất lớn nhất so với năm trước (t thuộc {Δt +1, ..., T - Δt + 1}, Δt = 250 ngày thị trường , xem chi tiết trong phần Phương thức 3.2.2). Cụ thể, chúng tôi tính tương quan qua khung thời gian sáu tháng, giảm ảnh hưởng quá mức của khủng hoảng kinh tế từ xa về tương quan hiện tại bằng cách sử dụng phương pháp tiên tiến lũy thừa 8 (gán trọng số cao hơn cho các sự kiện gần đây hơn và trọng số giảm dần cho các sự kiện trước đây). Sau đó, chúng tôi đã cải thiện bộ ước lượng bằng cách tính toán ma trận tương quan trung bình với độ rút gọn trong khoảng thời gian sáu tháng thu được theo cách này một ước tính mạnh mẽ về các mối tương quan so với năm trước ngày đầu tư t (xem chi tiết trong phần Phương pháp).

Ma trận như vậy cho thấy sự bền bỉ đáng kể, với các giá trị tự tương quan khoảng 50% ngay cả sau một năm. (Sự tự tương quan của ma trận tương quan được định nghĩa là tương quan giữa các vectơ của hệ số tương quan N (N - 1) / 2 tại thời điểm t và tại thời điểm t + τ .) Độ bền cao là một thực tế rất quan trọng cho thấy các phép đo từ quá khứ có thể dự báo tương lai và thứ tự của các mối tương quan được dự kiến sẽ duy trì khá ổn định. Sau đó chúng tôi sử dụng các mối tương quan có trọng số trung bình này với độ co rút để xây dựng các mạng lọc tài chính MST và PMFG3,4,10.

Bây giờ chúng tôi thảo luận về một chiến lược đầu tư hiệu quả có thể được hưởng lợi từ kiến thức về cấu trúc phụ thuộc thị trường như thế nào. Cụ thể, chúng tôi thiết lập danh mục đầu tư bằng cách chọn cổ phiếu từ khu vực ngoại vi của mạng được lọc tài chính và so sánh hiệu suất của các danh mục này với hiệu suất của danh mục được thiết lập bằng cách chọn cổ phiếu trung tâm hoặc cổ phiếu ngẫu nhiên hoặc bằng các phương pháp truyền thống khác. Với mục đích này, trước tiên chúng tôi phân biệt giữa các cổ phiếu nằm trong vùng trung tâm của mạng và những cổ phiếu nằm ở vùng ngoại vi. Nhiều biện pháp trung tâm / ngoại vi đã được đề xuất trong tài liệu 11–16; chúng phản ánh lại các tiêu chí khác nhau và nó không phải là bất thường mà là một đỉnh cho một thước đo kết quả trung tâm và một thước đo khác cho ngoại vị. Đặc biệt, các thước đo trung tâm trên MST và PMFG có xu hướng phân biệt tốt một vài đỉnh trung tâm, kết nối cao, quan trọng và có ảnh hưởng, nhưng chúng kém hiệu quả trong việc xếp hạng các mức ngoại vi khác nhau của các đỉnh không trung tâm. Do đó, chúng tôi đã chấp nhận quan điểm ‘thuyết bất khả tri’ bằng cách xem xét một số của các biện pháp trung tâm / ngoại vi phổ biến nhất :

* bậc (Degree): Bậc của một nút là số lượng các liên kết (cạnh) đến nút đó.
* độ trung tâm trung gian (Betweenness Centrality): Tính toán đường trung tâm ngắn nhất giữa các nút. Một nút có trung tâm trung gian càng lớn thì nó có vai trò quan trọng trong mạng xã hội bởi có thể nó là cầu nối liên kết giữa hai nhóm với nhau và nếu nó bị mất thì việc trao đổi thông tin giữa hai nhóm sẽ không thể thực hiện được.
* Độ gần gũi (Closeness ): Độ gần gũi của mỗi node với các nút khác trong mạng. Nếu một nút có độ gần gũi càng thấp thì để giao tiếp với các nút khác nó muốn thì phải đi qua rất nhiều nút khác trong mạng.
* Độ đặc trưng trung tâm (Eigenvector Centrality): độ đo tương tự như bậc, tuy nhiên thay vì đếm số lượng liên kết đến node đó thì nó quan tâm đến bậc của node mà liên kết với nó. Trong một mạng, khi 2 node có cùng bậc thì đại lượng này sẽ cho biết node nào liên kết với các node quan trọng hơn ở trong mạng.
* Độ lệch tâm (Eigenvector): Trả về độ lệch tâm của các nút trong G.

Tính cho cả MST và PMFG trọng số và các đối tác không có trọng số của chúng. Cụ thể, chúng tôi đã xây dựng hai chỉ số trung tâm kết hợp X và Y, nhóm này cùng nhau xếp hạng các biện pháp trước đó. Xét về các phép đo kết hợp này, các giá trị nhỏ (X + Y) được kết hợp với các đỉnh trung tâm trong khi các giá trị lớn được kết hợp với các đỉnh ngoại vi.

Đối với mỗi ngày t, chúng tôi đã xây dựng MST, PMFG được lọc tài chính mạng bằng cách sử dụng các mối tương quan trung bình với độ co rút được tính toán so với năm trước; sau đó chúng tôi chọn m cổ phiếu ngoại vi nhất (với giá trị lớn nhất của X + Y) và thiết lập danh mục đầu tư với một trong hai trọng lượng đồng nhất hoặc trọng lượng Markowitz7, có hoặc không có bán khống. Đối với mỗi danh mục đầu tư, chúng tôi đã quan sát thấy lợi nhuận, được xác định bằng công thức:

* =1 … 250

Hiệu suất của từng khoản đầu tư chiến lược được đo bằng cách tính toán giá trị trung bình (τ) và tiêu chuẩn độ lệch s(τ) của lợi nhuận trong 707 ngày đầu tư. Chúng tôi có sau đó chọn ‘tỷ số tín hiệu thành tiếng ồn’ (còn được gọi là ‘thông tin tỷ lệ ’), Ví dụ, như đại diện cho hiệu suất: chiến lược đầu tư tốt phải liên tục tạo ra lợi nhuận cao liên quan đến biến động nhỏ do đó được đặc trưng bởi tỷ lệ ; ngược lại, chiến lược đầu tư kém tạo ra lợi nhuận nhỏ và biến động lớn hơn (rủi ro lớn hơn) tạo ra các tỷ lệ tín hiệu nhiễu nhỏ. Tại một mức cổ phiếu riêng lẻ, rủi ro được phân phối đồng đều trên đồ thị tài chính.

3.2.2 Phương pháp

Dữ liệu và thu thập hàng ngày của 300 cổ phiếu. Chúng tôi đã nghiên cứu tất cả các cổ phiếu phổ thông thông thường trên thị trường chứng khoán Mỹ trong giai đoạn từ 1981 đến 2010 với tổng số T = 7570 ngày thị trường (dữ liệu từ CRSP23, cổ phiếu phổ thông thông thường của “ Americus Trust Components, Primes and Scores ”, “Quỹ đóng ”, “ Quỹ tín thác đầu tư bất động sản ” đã bị loại khỏi tập dữ liệu). Chúng tôi thực hiện phân tích của chúng tôi về việc di chuyển các khung thời gian của Δt = 250 ngày (một năm thị trường). Giá thiếu liên tiếp ít hơn năm ngày liên tiếp đã được thay thế bằng giá trị trước đó, và cho mỗi ngày t, cổ phiếu có ít hơn Δt quan sát liền kề cho đến khi t và Δt sau khi t bị loại bỏ (lưu ý rằng việc giữ các cổ phiếu này không ảnh hưởng đáng kể kết quả21). Đối với thị trường mỗi ngày, chúng tôi đã chọn 600 cổ phiếu đầu tiên bằng cách viết hoa. Chúng tôi tiếp tục giảm tập dữ liệu bằng cách chỉ giữ lại tập hợp con hàng đầu “hoạt động tốt nhất” của các cổ phiếu trong giai đoạn trước đó. Với mục đích này, cho mỗi mục đích cổ phiếu và mỗi lần chúng tôi tính toán lợi nhuận hàng ngày r(t, 1) và tính lợi nhuận trung bình và độ lệch chuẩn của chúng so với các ngày trước đó. Với tỷ lệ cao nhất của tỷ lệ (với những tỷ lệ trung bình có hiệu suất hàng ngày cao nhất so với các ngày Δt trước đó) với N = 300 cổ phiếu cho mỗi lần.

**Biện pháp phụ thuộc**. Để giảm sự ảnh hưởng quá mức của các sự kiện trên các tương quan hiện tại, chúng tôi sử dụng các trọng số mũ (được định nghĩa là , sao cho 0 và = 1) để các quan sát trong quá khứ đếm ít hơn những cái gần đây8. Ở đây, t = 1,2, ..., τ và θ >.0 là đường thời gian nghiên cứu. Mẫu có trọng số, phương sai, hiệp phương sai và tương quan được xác định từ các mức trung bình có trọng số từ:

Chúng tôi đã sử dụng mức trung bình theo cấp số nhân này để tính toán, đối với mỗi thời gian t, hệ số tương quan trọng của Pearson trong một khung sáu tháng (τ =θ = 125). Đối với mỗi ngày, chúng tôi đã theo dõi các mối tương quan này trong sáu tháng trước và chúng tôi tính giá trị trung bình của chúng với co rút9:

Sự co rút đáng kể cải thiện ý nghĩa số của sự tương quan ma trận (điều kiện số24 được giảm xuống hai bậc độ lớn).

**Các biện pháp trung tâm và ngoại vi**: Chúng tôi tính bậc Degree (D),độ trung tâm trung gian (BC), độ lệch tâm (E), độ gần (C), độ trung tâm riêng (EC), cho cả biểu đồ trọng số và không trọng số cho cả hai MST và PMFG. Đối với trọng số bậc (D) (thường được gọi là cường độ), trọng số độ trung tâm riêng (EC), trọng số giữa đỉnh i và đỉnh j là Trong khi đó, đối với trọng số trung tâm trung gian, độ lệch tâm và độ gần thì được tính là (tức là khoảng cách Euclide). Các phương pháp này được sắp xếp tương ứng theo thứ tự giảm dần cho các thước đo trung tâm (D,BC, EC) và thứ tự tăng dần cho các phép đo ngoại vi (E, C). Sau đó, đối với mỗi biện pháp gắn thứ hạng 25 đã được tính toán sao cho các đỉnh trung tâm đã được chỉ định thứ hạng cao hơn và các đỉnh ngoại vi thấp hơn. Tất cả các biện pháp này về tính trung tâm / ngoại vi rõ ràng là không độc lập và thực sự tất cả đều có kết quả tích cực hoặc liên quan lẫn nhau

nơi chúng tôi biểu thị với xếp hạng gắn liền của mức độ của bậc (D) . đối tác không trọng số của nó; đối với tất cả các biện pháp khác, chúng tôi sử dụng ký hiệu tương ứng (BC, E, C, EC) thay vì D. Hai phương pháp lai này phân biệt giữa:

* các đỉnh nhiều kết nối với các đỉnh nhiều kết nối khác (nhỏ X, nhỏ Y);
* các đỉnh nhiều kết nối với các đỉnh ít kết nối (nhỏ X,lớn Y);
* các đỉnh ít kết nối với các đỉnh nhiều kết nối (lớn X, nhỏ Y);
* các đỉnh ít kết nối với các đỉnh ít kết nối (lớn X,lớn Y).

Do đó, chúng tôi được coi là các biện pháp lai của trung tâm tổng và sự khác biệt giữa X và Y. Giá trị của X + Y là nhỏ đối với các đỉnh trung tâm và lớn cho các đỉnh ngoại vi; trong khi giá trị của X - Y là lớn nếu đỉnh có vài kết nối quan trọng và nó nhỏ nếu nó có nhiều kết nối không quan trọng.

3.3 Phân tích loại bỏ xu hướng tương quan chéo phụ thuộc q của thị trường chứng khoán

3.3.2 Giới thiệu

Phân tích mối tương quan chéo giữa các tài sản tài chính khác nhau đã trở nên vô cùng hấp dẫn [1, 2] kể từ khi các nhà nghiên cứu bắt đầu báo cáo vi phạm của Efficient Market Hypothesis (EMH). Ngay từ đầu, các phân tích chéo đã dựa vào các công cụ tuyến tính như sự tương quan Pearson, đòi hỏi sự cố định trong dữ liệu, nhưng các tập dữ liệu tài chính thế giới thực hiếm khi cố định. Để xem xét tính phi tuyến tính và phi cơ sở trong dữ liệu trong thế giới thực, các phương pháp mới dựa trên xu hướng thoái hóa đã được đề xuất, trong đó phổ biến nhất là phân tích biến động xu hướng (DFA). Được thúc đẩy bởi DFA được áp dụng cho một chuỗi thời gian, khái quát hóa của nó được gọi là phân tích biến động tương quan chéo (DCCA) đã được đề xuất để định lượng các mối tương quan chéo tầm xa giữa một cặp tín hiệu không cố định [4]. DFA và DCCA sau đó được mở rộng bằng các phiên bản đa phương thức của họ: MFDFA và MFDCCA, tương ứng [5–7]. DFA, DCCA và các đối tác đa chiều của họ đã được áp dụng qua một loạt các hệ thống bao gồm các hệ thống sinh học, tài chính đến vật lý [8–10]. Gần đây, một hệ số tương tự với hệ số Pearson, hệ số tương quan chéo được giải phóng p(s) đã được giới thiệu trong Ref [11]. Hệ số này áp dụng cho các tín hiệu không cố định định lượng mức ý nghĩa của các mối tương quan giữa các dao động của tín hiệu không dừng cố định tại một thang đo nổi bật s [12]. Gần đây, hệ số DCCA ρ(s) đã được sử dụng rộng rãi để nghiên cứu mối tương quan chéo phi tuyến tính giữa các chuỗi thời gian tài chính [8, 13–16]. Mặc dù thành công của hệ số p(s), nó có một số hạn chế khi các tương quan chéo được định lượng trong số các dao động ở độ lớn khác nhau. Một phần mở rộng của p(s), hệ số tương quan chéo p(q, s), q ∈ R, dựa trên các hàm biến động q-dependent từ MFDFA và MFDCCA [5, 7, 17]. Kwapien et al. gần đây đã chỉ ra rằng phương pháp này có thể được áp dụng cho việc phân tích dữ liệu thực nghiệm từ các hệ thống phức tạp tự nhiên như hệ thống vật lý, sinh học, xã hội và tài chính. Trọng tâm của chúng tôi ở đây là trên thị trường tài chính.

Đối với các chuỗi thời gian trả về đó, chúng tôi tạo ra ma trận tương quan chéo q-dependent C(q, s).Chúng tôi tính toán các thuộc tính thống kê của ma trận tại các đơn đặt hàng khác nhau và các thang thời gian khác nhau. Khi phân tích ma trận tương quan chéo Pearson, chúng tôi phân tích động lực riêng và biến đổi riêng của ma trận và thấy rằng mối tương quan chéo của biến động thị trường chứng khoán ở các cường độ khác nhau thể hiện cấu trúc và động lực độc đáo. Những biến động lớn luôn bị chi phối bởi một vài nhóm ngành, nhưng những biến động nhỏ thể hiện một hành vi khác. Sau đó chúng tôi trình bày các ma trận tương quan chéo dưới dạng các mạng phức tạp và sử dụng phương pháp đồ thị tối đa được tinh lọc (PMFG) để xây dựng các mạng tương quan và phân tích các đặc điểm cơ bản của chúng. Các mạng PMFG cho những dao động nhỏ hơn không đồng nhất so với những gì thu được cho những biến động lớn. Sử dụng chỉ số trung tâm, chúng tôi phân loại các cổ phiếu là trung tâm hoặc ngoại vi theo xếp hạng trung tâm của họ. Áp dụng điều này để tối ưu hóa danh mục đầu tư, chúng tôi thấy rằng danh mục đầu tư của các cổ phiếu ngoại vi có mức lợi nhuận cao hơn so với một trong những cổ phiếu trung tâm và được lựa chọn ngẫu nhiên.

3.3.3 Phương thức

3.3.3.1 Phân tích tương quan chéo d-dependent

Hệ số tương quan chéo q-dependent có thể thu được từ phương trình sau:

(i) Chúng tôi xem xét một chuỗi thời gian và , i = 1 … l. Chúng tôi đã tích hợp chuỗi thời gian này và nhận được hai chuỗi thời gian mới.

= , k = 1 … l.

= , k = 1 … l.

(ii) Chúng tôi chia và vào 2 = 2 x int(l/s) các ô không chồng chéo chiều dài s từ đầu và cuối của 2 chuỗi thời gian tích hợp. Sau đó chúng tôi tính toán xu hướng địa phương cho từng phân đoạn v(v = 0, 1, ..., 2) bằng một hình vuông nhỏ nhất và trừ nó ra khỏi và để xua tan chuỗi tích hợp. Sau đó chúng tôi tìm thấy các tín hiệu còn lại X, Y bằng với sự khác biệt giữa các tín hiệu tích hợp và đa thức bậc thứ tự trang bị cho các thiết bị này.

Hiệp phương sai và phương sai của X và Y trong hộp v được định nghĩa:

trong đó Z đại diện cho X hoặc Y.

(iii) Sau đó, chúng tôi xác định các chức năng biến động của order q và scale s

Hệ số tương quan chéo q-dependent giữa và được định nghĩa:

Khi q = 2 chúng tôi khôi phục hệ số tương quan chéo của p (s) [11]. Hệ số tương quan chéo q-dependent được giới hạn trong [−1, 1] khi q >= 0. Hệ số có thể có giá trị tùy ý khi q < 0. Ở đây chúng tôi tập trung vào trường hợp khi q > 0. Số mũ q hoạt động như một bộ lọc. Khi q > 2 hộp có biến động lớn đóng góp cho ρ (q, s) nhiều nhất, nhưng khi q <2 hộp có giá trị tương đối nhỏ chiếm ưu thế trong hàm biến động, do đó đóng góp nhiều nhất cho p(q, s).

3.3.3.2 Random matrix theory

Sau khi giới thiệu hệ số tương quan chéo q-dependent, chúng tôi bây giờ xây dựng các ma trận tương quan chéo C (q, s) tại các orders q và detrending scales s khác nhau. Nếu chúng tôi giả định các ma trận tương quan là ngẫu nhiên, thì lý thuyết ma trận ngẫu nhiên có thể được sử dụng như một điểm chuẩn để định lượng đến mức độ của các thuộc tính của ma trận tương quan chéo q-dependent lệch khỏi dự đoán của ma trận thuần túy ngẫu nhiên. Lý thuyết ma trận ngẫu nhiên đã được áp dụng rộng rãi để điều tra các hiện tượng tập thể trong các thị trường tài chính [2, 1, 19–27]. Đánh giá toàn diện là được cung cấp trong Ref. [28].

Chúng tôi xem xét ma trận tương quan ngẫu nhiên được tạo từ chuỗi thời gian, ví dụ: chuỗi thời gian trả về , I = 1 … l.

C =

Trong đó R là một ma trận N × L chứa N chuỗi thời gian trả về của độ dài L với giá trị trung bình bằng không và phương sai đơn vị, không đồng nhất với nhau. Hàm phân phối xác suất của các giá trị riêng của một ma trận ngẫu nhiên có thể được viết một cách phân tích trong giới hạn N, L 🡪 ∞ với Q = > 1

P(ℷ) =

Khi và là các giá trị riêng của C(q, s). Và và được tính bởi công thức:

= 1 + 2

Phương trình 11 là chính xác cho các phần tử ma trận phân tán Gaussian. Nếu các bản phân phối giá trị riêng biệt lệch khỏi dự đoán của 11, điều đó báo hiệu sự tồn tại của tương quan lẫn nhau trong chuỗi thời gian.

Chúng tôi phân hủy các ma trận tương quan chéo q-dependent với các giá trị riêng , k = 1 … N và eigenvectors , k = 1 … N cung cấp thông tin về hành vi tập thể của thị trường chứng khoán. Ở đây chúng tôi sử dụng tỷ lệ nghịch đảo tham gia để định lượng nghịch đảo của số thành phần eigenvector đóng góp đáng kể.

Tỷ lệ tham gia nghịch đảo (IPR) được xác định:

Ở đây là lth thành phần của the eigenvector tương ứng với eigenvalue . Ý nghĩa của có thể được minh họa bằng hai trường hợp giới hạn, (i) một vector với các thành phần giống nhau = 1 / có = 1 / N, trong khi (ii) một vector với một thành phần = 1 và số còn lại không có = 1. Chúng tôi cũng xác định tỷ lệ tham gia (PR) là 1 / , tương đương với những người đóng góp đáng kể cho eigenvalue . Trong lý thuyết ma trận ngẫu nhiên, kỳ vọng của IPR là

3.3.3.3 Planar maximally filtered graph

Như đã đề xuất trong Ref[17], chúng tôi sử dụng cách tiếp cận mạng phức tạp để phân tích ma trận tương quan chéo q-dependent. Chúng tôi sử dụng phương pháp đồ thị tối đa được tinh lọc (PMFG) để xây dựng mạng dựa trên ma trận tương quan C(q, s). Thuật toán được thực hiện như sau,

(i) Sắp xếp tất cả (q, s) theo thứ tự giảm dần để có được danh sách theo thứ tự .

(ii) Thêm một cạnh giữa các nút i và j dựa trên thứ tự trong chỉ khi đồ thị vẫn phẳng sau khi cạnh được thêm vào.

(iii) Một đồ thị G(q, s) được hình thành với các cạnh = 3 (N - 2) dưới sự ràng buộc của độ phẳng.

Như được mô tả trong Ref. [18], PMFGs không chỉ giữ cho tổ chức phân cấp của cây bao trùm tối thiểu (MST) mà còn tạo ra các dòng. Chúng tôi tính toán các thông số topo cơ bản như hệ số phân cụm C, độ dài đường đi ngắn nhất L và tính tổng hợp A. Chúng tôi cũng áp dụng một chỉ số không đồng nhất γ [29] để đo tính không đồng nhất của PMFGs được xác định bởi:

Ở đây là các degree của node i và j liên kết bởi cạnh {}.

3.3.5 Ứng dụng

Chúng tôi hiện đang tìm hiểu khả năng sử dụng các mạng PMFG phụ thuộc vào q để cải thiện hiệu suất tối ưu hóa danh mục đầu tư theo khuôn khổ danh mục đầu tư của Markowitz [33]. Đầu tiên, chúng tôi giới thiệu ngắn gọn lý thuyết danh mục đầu tư của Markowitz và sau đó chúng tôi sử dụng một số chỉ số trung tâm để chọn danh mục đầu tư từ các mạng PMFG. Xem xét một danh mục đầu tư (t) của các cổ phiếu có tỷ suất hoàn vốn , i = 1. . . m, m là kích thước danh mục đầu tư, nghĩa là số lượng cổ phiếu trong danh mục đầu tư. Lợi nhuận trên (t) của cổ phiếu là:

Trong đó là phần của tài sản đầu tư vào cổ phiếu i. Các phân số được chuẩn hóa sao cho = 1. Nguy cơ giữ danh mục đầu tư (t) có thể được định lượng bằng phương sai:

Với là tương quan chéo Pearson giữa và , và và là độ lệch chuẩn của và . Để tìm danh mục đầu tư tối ưu, chúng tôi tối đa hóa lợi tức của danh mục đầu tư = dưới sự ràng buộc rằng rủi ro trên danh mục đầu tư là một số giá trị cố định . Tối đa hóa tùy thuộc vào hai ràng buộc này tương đương với một bậc hai vấn đề tối ưu hóa

Với là ma trận hiệp phương sai của ma trận hồi quy R được đề cập trong ngữ cảnh trước (bây giờ với kích thước L × m). Tham số q là dung sai rủi ro q ∈ [0, ∞). Nếu chúng tôi thiết lập q lớn, chúng tôi có khả năng chịu rủi ro cao, dẫn đến lợi nhuận kỳ vọng lớn. Các danh mục đầu tư tối ưu có thể được biểu diễn dưới dạng cốt truyện của lợi nhuận như là một hàm của nguy cơ được gọi là biên giới hiệu quả. Ở đây chúng tôi không sử dụng hệ số tương quan chéo q-dependent trong chỉ số rủi ro . Chúng tôi chỉ sử dụng các mạng PMFG phụ thuộc vào q để lựa chọn cổ phiếu m và sau đó lý thuyết danh mục đầu tư Markowitz truyền thống được sử dụng để định lượng hiệu suất của danh mục đầu tư. Nó đã chỉ ra rằng danh mục đầu tư được lựa chọn từ các mạng PMFG được xây dựng từ ma trận tương quan chéo Pearson sử dụng một số biện pháp trung tâm thực hiện rất tốt [34]. Ở đây đầu tiên chúng tôi tính điểm số trung tâm được xác định bởi

CHƯƠNG 4 – TRÌNH BÀY, ĐÁNH GIÁ, BÀN LUẬN CÁC KẾT QUẢ

4.1 Chiến lược danh mục động sử dụng giải thuật Clustering

***4.1.1 Sự phát triển của mạng cấu trúc***

Thực tế, cấu trục mạng đang phát triển và thay đổi theo thời gian. Các đặc điểm tiến hóa của thị trường có thể được tìm thấy qua các thông số tô pô. Trung bình của hệ số tương quan của tất cả các cổ phiếu phản ánh tổng kết nối của cây khung nhỏ nhất.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Bậc | C | D\_correlation | D\_distance | D\_distance |
| VNINDEX | central | -0.010 | -0.076 | -0.044 | -0.070 | -0.125 |
| peripheral | -0.166 | -0.166 | -0.005 | 0.009 | 0.036 |
| HNXINDEX | central | -0.061 | -0.061 | -0.025 | -0.058 | -0.086 |
| peripheral | 0.005 | 0.005 | 0.052 | 0.030 | 0.025 |
| NYSE | central | -0.337 | -0.3901 | -0.18303 | -0.477 | -0.246 |
| peripheral | -0.300 | -0.300 | -0.076 | -0.062 | -0.027 |
| AMEX | central | -0.223 | -0.159 | -0.389 | -0.151 | -0.152 |
| peripheral | -0.078 | -0.07866 | -0.0175 | -0.005 | -0.131 |
| OLSOSBORS | central | -2.19974 | -4.756 | -2.671 | -2.739 | -3.168 |
| peripheral | -0.43202 | -0.432 | 0.690 | 0.573 | 0.457 |
| NASDAQ | central | -0.486 | -0.486 | -0.409 | -0.409 | -1.080 |
| peripheral | -0.428 | -0.428 | -1.019 | -1.019 | -0.407 |
| AEX | central | 0.019 | -0.0279 | -0.040 | -0.040 | -0.039 |
| peripheral | -0.02 | -0.027 | -0.00625 | -0.006 | -0.009 |
| CAC40 | central | -0.047 | -0.07 | -0.066 | -0.066 | -0.076 |
| peripheral | -0.059 | -0.059 | -0.09689 | -0.155 | -0.075 |
| EURO100 | central | -0.040 | -0.073 | -0.05638 | -0.050 | -0.043 |
| peripheral | -0.038 | -0.038 | -0.01905 | -0.009 | 0.010 |
| IBEX35 | central | -0.00714 | -0.008 | 0.002675 | 0.002 | -0.0008 |
| peripheral | -0.02141 | -0.021 | -0.01467 | -0.014 | -0.013 |
| NIKKEI225 | central | -2.03991 | -4.991 | -6.181 | -8.091 | 11.450 |
| peripheral | 3.325 | 2.123 | -4.541 | -10.201 | -11.501 |
| TSX | central | -0.006 | -0.039 | 0.026 | 0.106 | -0.021 |
| peripheral | -0.069 | 0.069 | 0.032 | 0.005 | -0.011 |
| AustraliaS&P200 | central | -0.027 | -0.036 | -0.00492 | -0.008 | -0.004 |
| peripheral | -0.021 | -0.021 | -0.0046 | 0.024 | -0.004 |
| NZX50 | central | -0.051 | -0.008 | -0.0029 | -0.016 | -0.002 |
| peripheral | -0.042 | -0.008 | -0.002 | 0.002 | -0.002 |
| Shanghai | central | -0.033 | -0.051 | -0.007 | -0.050 | 0.002 |
| peripheral | -0.028 | -0.028 | 0.001 | -0.110 | -0.008 |
| KOSPI | central | -185.029 | -238.946 | -151.989 | 57.079 | 48.326 |
| peripheral | -378.684 | -378.684 | 580.219 | -373.791 | -361.656 |
| SSEC50 | central | -0.004 | -0.010 | -0.003 | -0.008 | 0.006 |
| peripheral | -0.022 | -0.022 | -0.087 | -0.116 | -0.005 |
| BOVEPA | central | -0.05143 | -0.051 | -0.083 | -0.056 | -0.023 |
| peripheral | -0.05188 | -0.001 | -0.020 | -0.094 | -0.024 |

Bảng 3 Kết quả đầu tư của 20 thị trường theo điều kiện thị trường

Chúng tôi tiếp tục tìm ra chiến lược có xác suất lớn nhất để đạt được nhiều lợi nhuận hơn so với chiến lược ngẫu nhiên, tức là, chiến lược này vượt trội hơn chiến lược ngẫu nhiên trong hầu hết các khoảng thời gian đầu tư. Chiến lược này sử dụng Dcorrelation làm tham số để chọn danh mục đầu tư và xác định các điều kiện thị trường dựa trên tiêu chí ngày giao dịch cho cả hai thị trường

4.2 Lây lan rủi ro trên thị trường chứng khoán

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | 30 | 90 | 150 |
| VNINDEX | central | 3.3876 | 1.5762 | 1.25511 |
| peripheral | 5.0389 | 2.2365 | 2.0510 |
| HNXINDEX | central | 0.33612 | -0.1988 | 0.51055 |
| peripheral | 1.3761 | 0.14138 | 0.48976 |
| NYSE | central | 4.472612 | 1.892711 | 4.651575 |
| peripheral | 11.0986 | 2.18256 | 6.19652 |
| AMEX | central | 0.60087 | 0.9714 | 1.9483 |
| peripheral | -1.468 | -2.192 | 8.0120 |
| OLSOSBORS | central | -16.61524 | -7.37290 | 16.25635 |
| peripheral | -30.454389 | -15.45416 | 12.88803 |
| NASDAQ | central | 23.309475 | 23.30342 | 1.3840351 |
| peripheral | -185.3177 | -64.27049 | -39.2883 |
| AEX | central | 2.6948 | 0.6806 | 1.44836 |
| peripheral | 0.2272 | -0.4128 | -0.0655 |
| CAC40 | central | 3.536 | 0.6028 | 0.8622 |
| peripheral | 2.3002 | -0.040 | 1.6112 |
| EURO100 | central | 2.3833 | 0.4924 | 0.6603 |
| peripheral | 0.6272 | 0.1323 | -0.14233 |
| IBEX35 | central | 0.56957 | 0.00850 | 0.042264 |
| peripheral | -0.7392 | -0.20614 | 0.065786 |
| NIKKEI225 | central | 47.5796 | 63.87336 | 360.753041 |
| peripheral | 152.0093 | -165.56981 | 476.28450 |
| TSX | central | -3.89666 | -1.10475 | 1.74492 |
| peripheral | 8.98258 | 3.117144 | 4.315287 |
| AustraliaS&P200 | central | -0.0385 | 0.2564 | 0.1846 |
| peripheral | 0.9001 | 0.0536 | 0.0077 |
| NZX50 | central | 0.149827 | 0.06441 | 0.07107 |
| peripheral | 0.09074 | 0.0786 | 0.09139 |
| Shanghai | central | 3.632252 | 0.589639 | 1.80959 |
| peripheral | 2.4357 | 1.24527 | 1.46586 |
| KOSPI | central | 37928.9811 | 11019.49779 | 11191.15907 |
| peripheral | 2375.4624 | 408.9918 | 8014.1457 |
| SSEC50 | central | 1.39173 | 0.327147 | 0.474984 |
| peripheral | 1.06037 | 0.657822 | 0.83656 |
| BOVEPA | central | -0.2909 | 0.1791 | 0.3729 |
| peripheral | 0.0719 | 0.3972 | 0.1739 |
| XU100 | central | 2.319801 | 0.7061 | 1.925644 |
| peripheral | 3.79650 | 0.9660 | 2.1631 |

***4.2.1 Hiệu suất trung bình của các danh mục đầu tư khác nhau***

Các danh mục đầu tư các nhau được chạy thử nghiệm và cho ra kết quả:

* Đầu tư trung tâm thì có 14 trên 19 thị trường lời trên tổng 3 khoảng thời gian
* Đầu tư ngoại vi thì có 11 trên 19 thị trưởng lời lời trên tổng 3 khoảng thời gian

***4.2.2 Hiệu suất trên các khoảng thời gian***

* Khoảng thời gian 30 ngày thì:
  + trung tâm có 15 trên 19 danh mục lời
  + trung tâm có 15 trên 19 danh mục lời
* Khoảng thời gian 90 ngày thì:
  + trung tâm có 16 trên 19 danh mục lời
  + trung tâm có 12 trên 19 danh mục lời
* Khoảng thời gian 150 ngày thì:
  + trung tâm có 19 trên 19 danh mục lời
  + trung tâm có 11 trên 19 danh mục lời

***4.2.3 Khả năng trả về âm.***

Khả năng trả về âm cao nhất là 9 trên 19

***4.2.4 Khả năng thu nhập cao hơn.***

Khả năng trả về âm cao nhất là 19 trên 19

4.3 Phân tích loại bỏ xu hướng tương quan chéo phụ thuộc q của thị trường chứng khoán

**Thảo luận**

Chúng tôi đã chỉ ra rằng các biểu đồ được lọc tài chính có thể được sử dụng để chọn danh mục đầu tư có rủi ro thấp hơn và lợi nhuận tốt hơn so với những người thu được các phương pháp truyền thống khác. Điều này đã đạt được bằng cách xác định đầu tiên ma trận tương quan phù hợp, sau đó xây dựng MST và PMFG đồ thị được lọc tài chính và cuối cùng thiết lập các chỉ số thích hợp để chọn danh mục đầu tư làm bằng chứng khoán ở trung tâm hoặc ngoại vi vùng. Chúng tôi đã định lượng hiệu suất đầu tư theo sử dụng một loạt các biện pháp, bao gồm: tỷ lệ “tín hiệu đến tiếng ồn” giữa lợi nhuận trung bình và độ lệch chuẩn của chúng; danh mục đầu tư phương sai; xác suất để thu được lợi nhuận lớn hơn; khả năng không âm trả về; lợi nhuận trung bình và tỷ lệ thông tin Sharpe (xem Thông tin hỗ trợ). Tất cả các kết quả đều cho thấy danh mục đầu tư thiết lập từ một lựa chọn các cổ phiếu ngoại vi có rủi ro thấp hơn và tốt hơn lợi nhuận so với danh mục đầu tư được thiết lập từ một lựa chọn cổ phiếu trung tâm. Nghèo nàn biểu diễn của danh mục đầu tư trung tâm có thể là kết quả của thực tế là trung tâm của mạng có nhiều khả năng phải chịu bất ngờ nhiễu loạn do hiệu ứng đàn: trong thời kỳ bùng nổ và treo hệ thống được tương quan cao và đồng thời các nhà đầu tư vội vàng trong cùng một hướng, mua hoặc bán, tương ứng. Do đó, danh mục đầu tư có chứa các công ty ở trung tâm của những tâm trạng không hợp lý có nhiều khả năng mang rủi ro lớn hơn. Hiệu quả đa dạng hóa là có thể nếu danh mục đầu tư bao gồm các cổ phiếu được đặc trưng bởi cả hai mối tương quan thấp và tín hiệu nhiễu tín hiệu trả về tỉ lệ dự kiến cao. Chúng tôi đã chỉ ra rằng các chứng khoán này nằm trong các vùng ngoại vi của đồ thị được lọc tài chính. Có một phạm vi rộng lớn của khả năng ứng dụng và thử nghiệm cho hiện tại phương pháp trong nhiều lĩnh vực khác nhau bao gồm thị trường ngoại vi

***4.3.1 Data description***

Chúng tôi sử dụng hàm logarit được xác định là:

= ln(t+1) - ln(t)

Với (t) là giá đóng cửa điều chỉnh hàng ngày của cổ phiếu i tại thời điểm t. Sau đó chúng tôi sử dụng phương pháp trước đây để tính toán hệ số tương quan chéo q-dependent giữa bất kỳ chuỗi thời gian trả về và và lấy C(q, s) là ma trận N x N. Các mục ma trận của C(q, s) là các hệ số tương quan (q, s) giữa tất cả các cặp cổ phiếu. Chúng tôi thiết lập q và the detrending scale s [30, 1000] ngày giao dịch với = 40. Chúng tôi cũng thực hiện tính toán tương tự trên chuỗi thời gian trả lại xáo trộn và chuỗi thời gian ngẫu nhiên được mô phỏng và sử dụng chúng làm mô hình tham chiếu.

***4.3.2. Cross-correlation matrix analysis***

Với một loạt các ma trận tương quan chéo C (q, s) theo thứ tự q và bậc thang khác nhau, chúng tôi phân tích phân bố xác suất của các giá trị tương quan chéo, tức là các mục tam giác trên của ma trận tương quan. Đầu tiên, chúng tôi hiển thị cốt truyện của ma trận cho các bậc q khác nhau và các thang đo giảm dần trong Hình 1, đặt các mục chéo thành 0 để hiển thị tốt hơn. Sức mạnh của tương quan trung bình sẽ tăng nhẹ khi tỷ lệ tăng, nhưng sẽ giảm khi thứ tự multifractal q tăng lên. Chúng tôi sắp xếp các hàng và cột của ma trận tương quan theo khu vực chính thức và phân vùng con của S & P500. Lưu ý các cấu trúc ngành và phân ngành phụ trong các ma trận tương quan. Khi q <2 cấu trúc ngành rõ rệt hơn nhiều. Hình 2 cho thấy sự phân bố của các phần tử ma trận P (ρ) cho sáu giá trị khác nhau của q và sáu giá trị khác nhau của thang đo s. Chúng tôi có thể quan sát thấy sự phân bố của các ma trận trở nên ngày càng lệch sang bên trái và chiều rộng của các đỉnh phân phối khi thứ tự multifractal q tăng lên. Sự phân bố xác suất của hệ số tương quan chéo phụ thuộc q cho chuỗi thời gian trả về lệch đáng kể so với phân phối xáo trộn, và điều này có thể cung cấp thông tin chính xác về mối tương quan chéo giữa các độ dao động khác nhau. Các phân phối xáo trộn và mô phỏng trùng với nhau. Do đó, cấu trúc tương quan chéo khác nhau là kết quả của sự tương quan phi tuyến giữa các độ dao động khác nhau. Ngoài ra, khi q> 2 phân phối trở nên tương đối gần với trường hợp xáo trộn. Chúng tôi tính toán bốn khoảnh khắc thứ tự đầu tiên của ma trận tương quan để minh họa sự thay đổi trong phân phối tương quan chéo.

Hình 3 cho thấy bốn khoảnh khắc thứ tự đầu tiên của sự phân bố hệ số tương quan tại các đơn đặt hàng khác nhau q và các vảy giảm dần s. Mối tương quan chéo trung bình giảm khi thứ tự đa chiều tăng lên, cho thấy sự tương quan chéo giữa các biến động lớn là tương đối yếu. Từ phương sai, sai lệch và kurtosis chúng tôi thấy một sự chuyển đổi rõ ràng trong phân phối. Các hệ số tương quan chéo cho các đơn đặt hàng đa hướng lớn và nhỏ q phần lớn khác nhau, cho thấy các cấu trúc tương quan khác nhau giữa các độ dao động khác nhau.

Để phân tích thông tin chính hãng được thực hiện bởi các ma trận tương quan chéo phụ thuộc q, chúng tôi phân hủy các ma trận tương quan chéo và sắp xếp các giá trị riêng , k = 1 … 401 theo thứ tự tăng dần với các eigenvectors tương ứng của chúng, k = 1 … 401. Hình 4 và Hình 5 cho thấy sự phân bố của các giá trị riêng biệt số lượng lớn và lệch các giá trị riêng, tương ứng. Hình 4 chỉ cho các giá trị riêng nhỏ hơn 2. Các đường màu đen và màu xanh là các bản phân phối giá trị riêng cho các ma trận tương quan chéo phụ thuộc q ban đầu và kịch bản xáo trộn. Các đường màu đỏ là các bản phân phối riêng biệt được dự đoán bởi lý thuyết ma trận ngẫu nhiên. Chúng tôi cũng mô phỏng chuỗi thời gian 401 sử dụng phân phối Gaussian. Các đường màu xanh lá cây là các giá trị riêng của số lượng lớn từ các ma trận tương quan chéo phụ thuộc q được tính toán bằng chuỗi thời gian Gaussian mô phỏng. Chúng tôi thấy rằng sự phân bố số lượng riêng biệt của chuỗi thời gian xáo trộn và chuỗi thời gian mô phỏng xấp xỉ giống nhau. Điều này xác nhận rằng độ lệch của phân phối riêng là kết quả của sự tương quan chéo phi tuyến tính. Giới hạn dưới và trên của các giá trị riêng được dự đoán bởi RMT là = 0.47 và = 1.73. Sự phân bố các giá trị riêng của số lượng lớn cho tương quan chéo phụ thuộc q ban đầu khác với dự đoán lý thuyết ma trận ngẫu nhiên. Lưu ý rằng khi q> 2 phân phối giá trị riêng biệt cho các ma trận tương quan chéo ban đầu và ma trận xáo trộn tiếp cận dự đoán ma trận ngẫu nhiên. Hình 5 cho thấy các giá trị riêng biệt lệch cho các ma trận tương quan chéo ban đầu (đen), các kết quả xáo trộn (màu xanh), và các kết quả mô phỏng (màu xanh lục). Hành vi của những giá trị riêng biệt lệch khác với giá trị của q và s khác nhau. Các giá trị q lớn và các giá trị nhỏ có xu hướng gây ra các giá trị riêng biệt lớn hơn. Lưu ý rằng các giá trị riêng biệt lệch cho q = 4 lớn và nhỏ s = 70 đặc biệt là thông thoáng. Ngược lại, khi q = 0,4 và s = 830 chỉ có giá trị riêng biệt lớn nhất tiếp tục đi chệch khỏi các giá trị riêng biệt xáo trộn và mô phỏng. Điều này chỉ ra rằng các dao động nhỏ chỉ có thời gian đặc trưng rất ngắn. Hiệu ứng trung bình dài hạn của các dao động nhỏ bằng mức ồn. Nói chung, thứ tự multifractal lớn và quy mô nhỏ detrending s làm cho cấu trúc ngành (deviating eigenvalues) và chế độ thị trường (lớn nhất eigenvalue) tách ra từ mức độ tiếng ồn. Các giá trị q lớn và các giá trị nhỏ có xu hướng gây ra các giá trị riêng biệt lớn hơn. Lưu ý rằng các giá trị riêng biệt lệch cho q = 4 lớn và nhỏ s = 70 đặc biệt là thông thoáng. Ngược lại, khi q = 0,4 và s = 830 chỉ có giá trị riêng biệt lớn nhất tiếp tục đi chệch khỏi các giá trị riêng biệt xáo trộn và mô phỏng. Điều này chỉ ra rằng các dao động nhỏ chỉ có thời gian đặc trưng rất ngắn. Hiệu ứng trung bình dài hạn của các dao động nhỏ bằng mức ồn. Nói chung, thứ tự multifractal lớn và quy mô nhỏ detrending s làm cho cấu trúc ngành (deviating eigenvalues) và chế độ thị trường (lớn nhất eigenvalue) tách ra từ mức độ tiếng ồn.

Bốn giá trị riêng biệt đầu tiên cho các đơn đặt hàng nhiều thứ tự q và các vảy tăng dần được biểu diễn trong Hình 6. Các giá trị riêng lớn nhất cho q <2 xấp xỉ bằng với thứ tự của kích thước hệ thống. Hành vi của các giá trị riêng lớn nhất tương tự như tương quan chéo trung bình trong hình 3 (a). Điều này hỗ trợ kết luận rằng lớn nhất

eigenvalue tương ứng với chế độ thị trường được mô tả bởi nhiều nghiên cứu [19, 1] và nó giảm khi giá trị q tăng lên. Vì vậy, chế độ thị trường tại q nhỏ là cực kỳ mạnh mẽ hơn, mà có vẻ phản trực giác. Chúng tôi cũng quan sát thấy rằng bốn giá trị riêng đầu tiên tăng khi tăng quy mô tăng dần.

Người ta tin rằng những giá trị riêng biệt lệch khỏi dự đoán của lý thuyết ma trận ngẫu nhiên có chứa một số thông tin chính hãng liên quan đến ngành hoặc ngành như được mô tả trong Ref. [19, 21]. Để tìm ra thông tin ẩn được thực hiện bởi những giá trị riêng biệt lệch tại các đơn đặt hàng khác nhau và các thang chia nhỏ, trước tiên chúng tôi phân chia các cổ phiếu 401 thành các nhóm ngành được dán nhãn l = 1 … 24 (N1 stocks each) theo mã nhóm ngành của các cổ phiếu do GICS cung cấp. Sau đó chúng tôi xây dựng một ma trận chiếu P, với các phần tử = 1 / nếu cổ phiếu i thuộc nhóm ngành công nghiệp l và = 0 nếu không phải. Đối với bộ riêng , sự đóng góp = của mỗi ngành công nghiệp có thể thu được. Hình 7 cho thấy sự đóng góp của từng nhóm ngành với các giá trị nhỏ nhất và nhỏ thứ hai , k = 1, , k = 2. Các đường màu đỏ (k = 1) và bầu trời xanh (k = 2) là giá trị đóng góp sau ảnh hưởng của giá trị riêng lớn nhất cho được xóa. Các đường màu xanh là giá trị đóng góp trung bình cho các ma trận tương quan được tính toán bằng chuỗi thời gian xáo trộn. Mô hình tham chiếu này cho chúng tôi biết độ lệch của từ mức độ nhiễu. Có 24 nhóm ngành chính cho 401 cổ phiếu: Bán lẻ, hàng gia dụng và cá nhân, năng lượng, tài chính đa dạng, ngân hàng, bảo hiểm, bất động sản, dược phẩm, công nghệ sinh học & khoa học đời sống, thiết bị chăm sóc sức khỏe & dịch vụ, hàng hóa vốn, Giao thông vận tải, Phần mềm & Dịch vụ, Thương mại & Chuyên nghiệp Dịch vụ, Vật liệu, Công nghệ Phần cứng & Thiết bị, Chất bán dẫn & Thiết bị bán dẫn, Dịch vụ viễn thông, Tiện ích, từ trái sang phải. Nó được chỉ ra rằng đối với các giá trị nhỏ nhất và nhỏ thứ hai, và , sự đóng góp đến từ một vài nhóm ngành và cho các nhóm ngành này mạnh hơn nhiều so với mức độ ồn. Sự đóng góp của ngành công nghiệp của các giá trị riêng biệt lớn , k = 399, 400 được trình bày trong Hình 8. sự đóng góp cho các giá trị riêng biệt lớn cũng đến từ một vài nhóm ngành và mạnh hơn nhiều so với mức độ ồn. Cho có nhiều ngành đóng góp đáng kể với một mẫu hỗn hợp. Đóng góp chính đến từ Tài chính đa dạng, Ngân hàng, Bất động sản và Tiện ích. Nhưng những đóng góp cho luôn đến từ Năng lượng và Tiện ích.

Như thể hiện trong hình 4, có rất nhiều giá trị riêng nhỏ trong dự đoán của lý thuyết ma trận ngẫu nhiên. Hình 9 cho thấy sự đóng góp của mỗi nhóm ngành công nghiệp với các giá trị riêng , k = 200, 250 sâu bên trong khu vực khối lượng riêng. Theo dự kiến, trong khu vực này, các giá trị riêng không thể hiện được một mô hình đáng kể nào. Mức đóng góp của từng nhóm ngành giống như trong chuỗi thời gian xáo trộn. Đối với cả hai , k = 200 và , k = 250, không có các nhóm ngành đóng góp rõ ràng.

Như đã giải thích ở trên, tỷ lệ tham gia nghịch đảo định lượng nghịch đảo của số thành phần eigenvector đóng góp đáng kể. Ở đây chúng tôi đưa ra tỷ lệ tham gia nghịch đảo của các ma trận tương quan chéo q-dependent ở các đơn đặt hàng nhiều bậc khác nhau và các thang đo giảm dần trong Fig.10. Chúng tôi trình bày tỷ lệ tham gia nghịch đảo mà không có giá trị riêng biệt lớn nhất để trực quan hóa tốt hơn. Lưu ý rằng có sự chuyển đổi trong IPR cho thứ tự đa chiều nhỏ và lớn q. Khi q ≤ 2, các giá trị riêng nhỏ chiếm ưu thế bởi tỷ trọng tương đối nhỏ của các cổ phiếu có IPR lớn hơn. Nó có thể được xác nhận bằng tỷ lệ tham gia 1 / trong Hình 12, là tỷ lệ tham gia cho các giá trị nhỏ nhỏ hơn 50. Đối với các giá trị riêng trung bình và lớn, tỷ lệ tham gia lớn hơn 200.

Hình 11 cho thấy tỷ lệ tham gia 1 / cho giá trị riêng lớn nhất. Tỷ lệ tham gia lớn nhất cho q <2 là 376 tiếp cận kích thước hệ thống N = 401. Khi q ≥ 2, tỷ số tham gia cho giá trị riêng lớn nhất giảm nhanh và có giá trị 200. Sự khác biệt nổi bật trong số đóng góp của các giá trị riêng lớn nhất cho các biến động khác nhau ngụ ý rằng hành vi tập thể của các dao động nhỏ (q < 2) là đồng nhất hơn (tỷ lệ tham gia lớn). Hình 12 cho thấy bản đồ nhiệt của tỷ lệ tham gia 1 / ở các đơn đặt hàng nhiều bậc khác nhau q khi s = 50, 210, 410, 810. k là nhãn của giá trị riêng λk. Khi q ≥ 2, tỷ lệ tham gia của các giá trị riêng nhỏ (nhỏ k) là rất nhỏ cho thấy rằng các giá trị riêng nhỏ chứa thông tin hữu ích. Chỉ có một tập hợp rất nhỏ cổ phiếu đóng góp vào giá trị riêng nhỏ nhất. Chúng tôi có thể xác minh điều này bằng cách sử dụng sự đóng góp thành phần eigenvector trong hình 7. Khi q ≥ 2 các giá trị riêng nhỏ bị chi phối bởi một vài sector. Điều này có ý nghĩa liên quan đến tối ưu hóa danh mục đầu tư. Nói chung, mô hình hành vi tập thể đối với các dao động nhỏ khác với các biến động lớn.

***4.3.3 PMFG analysis***

Biểu đồ được lọc tối đa phẳng (PMFG) đã được sử dụng để phân tích cấu trúc và động lực của thị trường chứng khoán trong thời điểm khủng hoảng [30, 31], và nó có hiệu quả nắm bắt các cấu trúc ngành. Ở đây chúng tôi xây dựng các mạng PMFG sử dụng ma trận tương quan chéo q-dependent. Hình 13 cho thấy các mạng được xây dựng bằng cách sử dụng thuật toán PMFG. Cơ cấu ngành cho q nhỏ là rõ ràng hơn so với q cho q. Gần đây Kawpen và cộng sự [32] xây dựng các cây bao trùm tối thiểu bằng cách sử dụng ma trận tương quan phụ thuộc q. Một số cấu trúc ẩn được tìm thấy bằng cách sử dụng các tập dữ liệu phút. Ở đây chúng tôi thấy rằng khi q ≤ 2, một kho trung tâm xuất hiện, nhưng khi q> 2, mức độ không đồng nhất trở nên yếu. Đặc biệt, khi q ≤ 2 các nút màu xanh đậm (các hình thức cổ phiếu ngành Tài chính) rất gần nhau. Tuy nhiên, khi q> 2, các liên kết giữa tài chính cổ phiếu ngành nới lỏng. Những đặc điểm này đồng thuận với kết quả từ [32] trong đó chúng phát hiện ra một ngôi sao như cấu trúc cây bao trùm tối thiểu khi q ≤ 2.

Để định lượng ảnh hưởng của các biến động trên PMFG tại các đơn đặt hàng nhiều bậc khác nhau q và các thang đo giảm dần, chúng tôi tính toán số lượng topo của PMFG. Số lượng topo của PMFG được trình bày trong Fig14. Hình 14 (a) cho thấy hệ số phân cụm C của PMFG tăng khi thứ tự multifractal q tăng lên. Hệ số phân cụm là lớn khi quy mô giảm dần ngắn. Chiều dài đường đi ngắn nhất L được thể hiện trong Fig14 (b). Chiều dài đường đi ngắn nhất là lớn cho q lớn và s ngắn. Fig14 (c) là chỉ số không đồng nhất H [29], định lượng mức độ không đồng nhất của PMFG. Nó tương tự với chỉ số pháp luật quyền lực của mạng không có quy mô. Được biết, tính không đồng nhất của mạng BA là 0,11. Chúng tôi nhận thấy rằng đối với q nhỏ, tính không đồng nhất của mạng PMFG lớn hơn mạng BA. Điều này có nghĩa là cấu trúc của các mạng PMFG đối với các đơn đặt hàng nhỏ multifractal q là cực kỳ không đồng nhất. chúng tôi cũng cho thấy sự tương đồng A của PMFG ở Fig14 (d). Sự tương đồng tiêu cực đối với q <2 cho thấy một gợi ý về cấu trúc không phân định trong đó các cổ phiếu trung tâm có xu hướng kết nối với các cổ phiếu có mức độ nhỏ. Khi q> 2 phương pháp tiếp cận tương đối 0. Điều này cho thấy rằng đối với q lớn các kết nối được phân bố đều hơn (xem Hình 13). Trong mạng có q> 2 mức độ của các hub trung tâm nhỏ hơn các hub trong mạng với q <2. Từ biến thể của đại lượng topo, chúng tôi có thể phỏng đoán rằng đối với các dao động nhỏ (q nhỏ) trong thời gian ngắn

quy mô (nhỏ), tồn tại một số cổ phiếu hàng đầu. Nhưng đối với các biến động lớn (q lớn) và quy mô thời gian dài (lớn), các cổ phiếu có mối tương quan thống nhất. Tóm lại, từ những đại lượng topo đó, sự thay đổi cấu trúc rõ ràng là điều hiển nhiên cho thấy sự khác biệt về hành vi tập thể giữa các dao động của các cường độ khác nhau ở các thang thời gian khác nhau.

CHƯƠNG 5 – KẾT LUẬN

**TÀI LIỆU THAM KHẢO**

**Tiếng Anh**

* 1. Mantegna RN. (1999), Hierarchical structure in financial markets, pp. 193-197.
  2. Ya-Nan Lu, Sai-Ping Li, Xiong-Fei Jiang, Li-Xin Zhong, Tian Qiu. (2017), Dynamic Portfolio Strategy Using Clustering Approach,
  3. Longfeng Zhaoa, Wei Lia, Andrea Fenub, Boris Podobnikb, Yougui Wang, H. Eugene Stanley. (2017), The q-dependent detrended cross-correlation analysis of stock market
  4. F. Pozzi, T. Di Matteo & T. Aste. (2013), Spread of risk across financial markets: better to invest in the peripheries