

**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

**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

Nhóm em chân thành cảm ơn thầy Nguyễn Chí Thiện đã tận tình hướng dẫn, gợi ý những tài liệu tham khảo và giải đáp những thắc mắc trong suốt quá trình làm việc. Vì những hạn chế về mặt kiến thức của bản thân, nếu không có những lời hướng dẫn, những kiến thức thầy truyền đạt và những tài liệu tham khảo thầy gợi ý mỗi khi gặp khó khăn trong quá trình làm bài thì nhóm em khó có thể hoàn thành luận văn này. Một lần nữa nhóm em xin chân thành cảm ơn thầy.

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 chứng khoán 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, hiện thực và so sánh kết quả các mô hình toán học giúp cho những nhà đầu tư có thể lựa chọn được những danh mục đầu tư có thể đem lại lợi nhuận tốt.

MỤC LỤC

LỜI CẢM ƠN	3
TÓM TẮT	5
MỤC LỤC.....	1
DANH MỤC KÍ HIỆU VÀ CHỮ VIẾT TẮT	3
DANH MỤC CÁC BẢNG BIỂU, HÌNH VẼ, ĐỒ THỊ	4
CHƯƠNG 1 – MỞ ĐẦU	5
1.1 Đặt vấn đề	5
1.2 Mục tiêu và phạm vi nghiên cứu.....	5
1.2.1 Mục tiêu nghiên cứu	5
1.2.2 Phạm vi nghiên cứu.....	6
CHƯƠNG 2 – TỔNG QUAN	7
2.1 Thu thập dữ liệu	7
2.2 Hiện thực các mô hình toán học của thị trường chứng khoán	8
2.3 So sánh kết quả hiện thực	9
CHƯƠNG 3 – PHƯƠNG PHÁP NGHIÊN CỨU LÝ THUYẾT VÀ THỰC NGHIỆM	10
3.1 Thu thập dữ liệu	10
3.1.1 Các yêu cầu thu thập dữ liệu.....	10
3.1.2 Phương pháp thu thập dữ liệu	10
3.2 Các phương pháp lựa chọn danh mục đầu tư.....	10
3.2.1 Chiến lược danh mục động sử dụng phương pháp phân cụm.....	11
3.2.1.1 Giới thiệu	11
3.2.1.2 Phương thức	11
3.2.1.3 Xây dựng mạng dựa trên thuật toán cây khung nhỏ nhất (MST)	12
3.2.1.4 Lựa chọn danh mục dựa trên các tham số tô pô.....	15
3.2.1.5 Hiện thực phương pháp.....	16
3.2.2 Chiến lược lựa chọn danh mục đầu tư tránh sự lây lan rủi ro trên thị trường chứng khoán.....	16
3.2.2.1 Giới thiệu	16
3.2.2.2 Phương thức.....	17
3.2.2.3 Xây dựng mạng tài chính	18
3.2.2.4 Lựa chọn danh mục đầu tư.....	18

3.2.2.5 Hiện thực phương pháp.....	20
3.2.3 Chiến lược lựa chọn danh mục sử dụng tương quan chéo phụ thuộc q	20
3.2.3.1 Giới thiệu	21
3.2.3.2 Phương thức	21
3.2.3.3 Xây dựng mạng tài chính	22
3.2.3.4 Lựa chọn danh mục đầu tư	22
3.2.3.5 Hiện thực phương pháp.....	23
3.3 So sánh các kết quả thực nghiệm	23
CHƯƠNG 4 – TRÌNH BÀY, ĐÁNH GIÁ, BÀN LUẬN CÁC KẾT QUẢ	26
4.1 Dữ liệu chứng khoán.....	26
4.2 Các chiến lược lựa chọn danh mục đầu tư	26
4.2.1 Chiến lược danh mục động sử dụng phương pháp phân cụm.....	26
4.2.2 Chiến lược lựa chọn danh mục đầu tư tránh sự lây lan rủi ro trên thị trường chứng khoán.....	27
4.2.3 Chiến lược lựa chọn danh mục sử dụng tương quan chéo phụ thuộc q	27
4.3 So sánh kết quả	27
CHƯƠNG 5 – KẾT LUẬN	35
5.1 Kết luận.....	35
5.2 Đề xuất các hướng nghiên cứu tiếp theo.....	35
TÀI LIỆU THAM KHẢO.....	Error! Bookmark not defined.
PHỤ LỤC	Error! Bookmark not defined.

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

$P_i(t)$ Giá đóng cửa của chứng khoán i vào ngày thứ t

r_i chuỗi trả về logarith của các chứng khoán i .

r_j chuỗi trả về logarith của các chứng khoán j

t ngày chọn cổ phiếu

CÁC CHỮ VIẾT TẮT

PMFG Planst Maximally Filtered Graphs

MST Minimum Spanning Tree

BC Betweenness Centrality

E Eigenvector

C Closeness

EC Eigenvector Centrality

D Degree

DANH MỤC CÁC BẢNG BIỂU, HÌNH VẼ, ĐỒ THỊ

DANH MỤC HÌNH

Hình 2.1 Trang web cophieu68.vn.....	8
Hình 2.2 Trang web finance.yahoo.com.....	9
Hình 3.1 Mạng lưới danh mục đầu tư sàn Hà Nội	14
Hình 3.2 Mạng lưới danh mục đầu tư sàn Hồ Chí Minh	14
Hình 4.1 Biểu đồ kết quả của các sàn theo danh mục đầu tư trung tâm	30
Hình 4.2 Biểu đồ kết quả của các sàn theo danh mục đầu tư ngoại vi	33

DANH MỤC BẢNG

Bảng 4.1 Thông tin các dữ liệu đã thu thập được	26
---	----

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 lợi nhuận mong muốn. Nhưng liệu việc đầu tư 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 chứng khoán 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 được ra đời để phục vụ cho việc lựa chọn danh mục đầu tư 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 mô hình toán học của thị trường chứng khoán và các thuật toán giao dịch trên các mô hình.

Thử nghiệm thực nghiệm 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 chiến thuật sẽ được áp dụng trên nhiều thị trường chứng khoán trên toàn thế giới như: Việt Nam, Mỹ, Na Uy, Hàn Quốc, Nhật Bản, Úc...

Tập đầu tư:

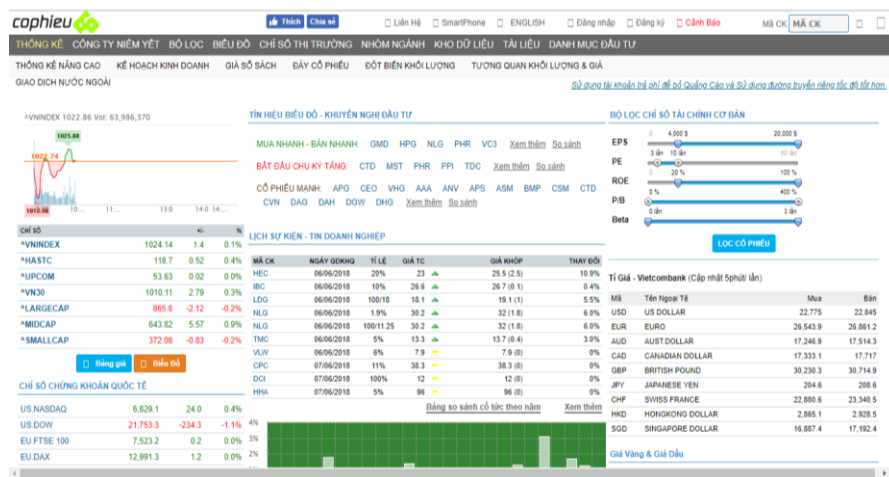
- Các tập đầu tư gồm các chứng khoán của các công ty đã được niêm yết.
- Dữ liệu nghiên cứu trong khoảng thời gian: từ tháng 1 năm 2000 đến tháng 6 năm 2018.

CHƯƠNG 2 – TỔNG QUAN

2.1 Thu thập dữ liệu

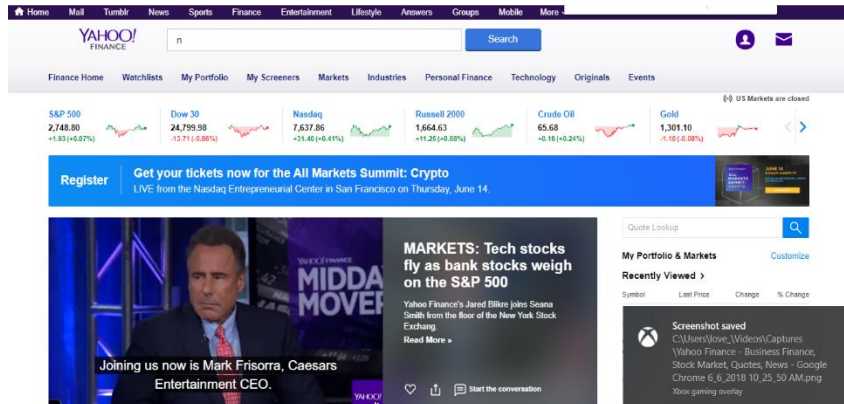
Chúng tôi thu thập dữ liệu chứng khoán của các công ty được niêm yết trên các sàn giao dịch chứng khoán.

- Dữ liệu chứng khoán của các sàn giao dịch chứng khoán nội địa được chúng tôi thu thập từ website: <http://www.cophieu68.com>



Hình 2.1 Trang web cophieu68.vn

- Dữ liệu chứng khoán của các sàn giao dịch chứng khoán quốc tế được chúng tôi thu thập từ website: <http://www.finance.yahoo.com>



Hình 2.2 Trang web finance.yahoo.com

- Dữ liệu giao dịch hằng ngày của chứng khoán nằm trong khoảng thời gian từ 1 ngày 1 tháng 1 năm 2000 đến ngày 1 tháng 6 năm 2018. Để đảm bảo sự liên tục và toàn vẹn của dữ liệu, các chứng khoán được lựa chọn trong nghiên cứu là chứng khoán có cổ phần tích cực nhất trong giao dịch suốt giai đoạn lấy mẫu. Với mục đích này, chúng tôi sẽ loại bỏ những chứng khoán ngừng giao dịch trong thị trường trong hơn 30 ngày giao dịch.

2.2 Hiện thực các mô hình toán học của thị trường chứng khoán

Động cơ chính của nghiên cứu này là đề xuất một chiến lược đầu có thể đáp ứng hầu hết các thị trường trên thế giới. Trong nghiên cứu này, chúng tôi hiện thực ba chiến lược:

- Chiến lược danh mục động sử dụng phương pháp phân cụm: Bằng cách sử dụng phương pháp phân cụm theo thứ bậc mối quan hệ giữa các chứng khoán chúng tôi có thể phân cụm nhóm chứng khoán tối ưu để chọn danh mục đầu tư mang lại lợi nhuận tốt nhất.
- Chiến lược lựa chọn danh mục đầu tư tránh sự lây lan rủi ro trên thị trường chứng khoán: Các chứng khoán không ảnh hưởng tất cả trên sàn giao dịch vì vậy chúng ta có thể lựa ra các chứng khoán độc lập nhất đưa vào danh mục đầu tư để đạt lợi nhuận cao nhất.

- Chiến lược lựa chọn danh mục sử dụng tương quan chéo phụ thuộc q : Dựa vào sự tương quan chéo của các chứng khoán để tìm ra danh mục đầu tư tốt nhất để đầu tư.

Hiện thực và đầu tư các mô hình trên vào 12 sản giao dịch trên toàn cầu để thống kê kết quả và đề xuất mô hình tốt nhất.

2.3 So sánh kết quả hiện thực

Việc so sánh kết quả của các mô hình được so sánh theo 2 cách:

- So sánh lợi nhuận các chiến lược theo danh mục đầu tư chọn các chứng khoán cùng loại.
- So sánh lợi nhuận theo các khoảng thời gian đầu tư.

Với kết quả so sánh cuối cùng, chúng tôi sẽ bàn luận về hiệu quả của các chiến lược.

CHƯƠNG 3 – PHƯƠNG PHÁP NGHIÊN CỨU LÝ THUYẾT VÀ THỰC NGHIỆM

3.1 Thu thập dữ liệu

3.1.1 Các yêu cầu thu thập dữ liệu

Chúng tôi đã tiến hành thu thập dữ liệu chứng khoán từ nhiều website khác nhau và chúng tôi đã thu được dữ liệu chứng khoán của 12 sàn giao dịch chứng khoán khác nhau trên toàn thế giới. Đó là các sàn giao dịch chứng khoán thuộc các nước như: Việt Nam, Mỹ, Trung Quốc, Canada, Dữ liệu chúng tôi thu thập được chứa thông tin giao dịch hằng ngày của các chứng khoán là: mã chứng khoán, giá mở cửa, giá đóng cửa, số lượng giao dịch,

Để có được đủ dữ liệu phục vụ cho mục đích thử nghiệm các phương pháp lựa chọn danh mục đầu tư, chúng tôi đã tiến hành thu thập dữ liệu giao dịch hằng ngày của những chứng khoán có khối lượng giao dịch lớn nhất của các sàn giao dịch. Và kết quả là chúng tôi đã thu được là thông tin giao dịch hằng ngày nằm trong khoảng từ ngày 1 tháng 1 năm 2000 đến ngày 1 tháng 6 năm 2018.

Sau khi thu thập được dữ liệu, nhằm đảm bảo sự liên tục của dữ liệu chúng tôi sẽ tiến hành loại bỏ các dữ liệu bị dừng giao dịch trên 30 ngày trong suốt quá trình giao dịch trên thị trường.

3.1.2 Phương pháp thu thập dữ liệu

Để thu thập được số dữ liệu đã nêu ở phần trước, chúng tôi đã viết một chương trình tự động tải dữ liệu từ các website. Chương trình được viết bằng ngôn ngữ python và có sử dụng những module như: selenium, os, ...

Trước tiên, chương trình sử dụng thư viện Selenium để lấy được tất cả mã cổ phiếu trên các website như: <https://tradingeconomics.com>, <http://www.cophieu68.com/>.

Tiếp theo sẽ dùng các mã cổ phiếu này để tải lịch sử giao dịch của cổ phiếu đó.

3.2 Các phương pháp lựa chọn danh mục đầu tư

Sau thời gian tìm kiếm chúng tôi đã tìm được ba chiến lược lựa chọn danh mục đầu tư đó là:

- Chiến lược danh mục động sử dụng phương pháp phân cụm [10].
- Chiến lược lựa chọn danh mục đầu tư tránh sự lây lan rủi ro trên thị trường chứng khoán [1].
- Chiến lược lựa chọn danh mục sử dụng tương quan chéo phụ thuộc q [4].

Ba chiến lược này sẽ được mô tả chi tiết ở bên dưới.

3.2.1 Chiến lược danh mục động sử dụng phương pháp phân cụm

3.2.1.1 Giới thiệu

Phân cụm theo thứ bậc là một trong những phương pháp hiệu quả nhất để chọn một tập chứng khoán cho danh mục đầu tư tối ưu. Trên thực tế, việc lựa chọn một bộ chứng khoán 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 chứng khoán được mạng lưới xây dựng chứng khoán biểu diễn bởi cấu trúc tô pô[3][6] và nó áp dụng cho tối ưu hóa danh mục đầu tư. Mô tả về mối tương quan giữa các chứng khoán gồm những chứng khoán 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[8] (MST). Bằng cách khai thác cấu trúc phụ thuộc của các chứng khoán tài chính bao gồm cả thuật toán MST. Người ta thấy rằng danh mục đầu tư từ các chứng khoán ngoại vi có rủi ro thấp hơn và lợi nhuận cao hơn danh mục đầu tư từ các chứng khoán 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 của mạng. Thuật toán phân cụm dữ K-means và thuật toán phân cụm dữ liệu dùng để phân loại cổ phiếu. Các cổ phiếu được lựa chọn từ các nhóm phân loại này được sử dụng để xây dựng danh mục đầu tư

3.2.1.2 Phương thức

Chuỗi trả về logarith của một cổ phiếu nhất định i được tính bằng:

$$r_i(t) = \ln P_i(t) - \ln P_i(t - 1)$$

Trong đó:

- $P_i(t)$: 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 ngày lựa chọn cổ phiếu để đầu tư “ t ”, 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 $\{t - \delta t + 1, \dots, t\}$ và mạng lọc tài chính 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 cổ phiếu 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 ($t, \dots, t + \Delta t$).

Sau đó di chuyển đến $t + \varphi$. Một chiến lược danh mục giống nhau được thông qua bằng cách chọn danh mục đầu tư trong khoảng thời gian $\{t + \varphi - \delta t + 1, \dots, t + \varphi\}$ và sau đó sử dụng danh mục đầu tư đã chọn để đầu tư vào đường thời gian $\{t + \varphi + 1, \dots, t + \varphi + \Delta t\}$. Chúng tôi sẽ tính tổng lợi nhuận trung bình của các tập cổ phiếu trung tâm và ngoại vi sau khi kết thúc thời gian đầu tư.

Sự lựa chọn δt thích hợp thực sự sẽ giúp mạng lọc tài chính nắm bắt được thông tin của dữ liệu ban đầu một cách tốt hơn. Việc lựa chọn được δt tốt sẽ giúp cấu trúc mạng càng ổn định và thông tin thị trường được lọc ra càng chính xác. Nhiều nghiên cứu đã tiết lộ rằng để đảm bảo các cổ phiếu có đủ số ngày giao dịch có ý nghĩa thống kê, δt nên lớn hơn số lượng cổ phiếu mẫu. Chúng tôi chọn $\delta t = 10$ tháng.

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

Giả sử $r_i(t)$ và $r_j(t)$ là chuỗi trả về logarithm của hai cổ phiếu i và j . Sử dụng hệ số tương quan Pearson giữa các chuỗi trả về của chúng bằng công thức:

$$p(i, j) = \frac{E[r_i r_j] - E[r_i]E[r_j]}{\sqrt{(E[r_i^2] - E[r_i]^2)(E[r_j^2] - E[r_j]^2)}} \quad [10, \text{tr.4}]$$

Trong đó:

- $E[\]$: biểu diễn kỳ vọng toán học của dãy theo khoảng thời gian t .
- r_i : chuỗi trả về logarith của các cổ phiếu i .
- r_j : chuỗi trả về logarith 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:

$$d(i, j) = \sqrt{2(1 - p(i, j))} \quad [10, \text{tr.4}]$$

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 cổ phiếu, mỗi một thị trường thu được ma trận khoảng cách.

Chúng tôi chọn phương pháp MST để lọc ra các đồ thị mạng trong các khoảng thời gian 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. Thủ tục xây dựng mạng MST được thực hiện như sau:

- Thứ nhất, sắp xếp khoảng cách giữa tất cả các cặp cổ phiếu theo thứ tự tăng dần.
 - Thứ hai, 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.
- Chúng tôi chọn thuật toán Prim để xây dựng mạng của chúng tôi.

Hình 3.2 Mạng lưới danh mục đầu tư sàn Hồ Chí Minh

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

Năm tham số được sử dụng để đo trung tâm và ngoại vi nhằm lựa chọn ra tập cổ phiếu trung tâm và cổ phiếu ngoại vi. Chúng tôi sẽ giới thiệu ngắn gọn các thông số này như sau:

- 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 bằng công thức:

$$C = \sum_{i, j \in V} \frac{\sigma_{ij}(V)}{\sigma_{ij}} \quad [10, \text{tr.4}]$$

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

- Khoảng cách dựa trên bậc (D_{degree}): Nút trung tâm là nút có bậc lớn nhất.
- Khoảng cách dựa trên sự tương quan ($D_{\text{correlation}}$): 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 (D_{distance}): Nút trung tâm là nút có giá trị trung bình khoảng cách tới các nút kề nhỏ nhất.

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 chọn ra một nút trung tâm dựa theo Degree, Correlation và Distance. Sau đó chúng tôi sẽ tính khoảng cách từ nút trung tâm đến tất cả các nút.

Chúng tôi sẽ chọn 10% số cổ phiếu có khoảng cách lớn nhất sẽ là tập cổ phiếu ngoại vi, 10% số cổ phiếu có khoảng cách nhỏ nhất sẽ là tập cổ phiếu 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 của các cổ phiếu. Cổ phiếu trung tâm đóng một vai trò quan trọng trên thị trường, ảnh hưởng lớn tới chỉ số thị trường và gây áp lực mạnh lên các cổ phiếu khác, trong khi mối tương quan giữa các cổ phiếu ngoại vi ít ảnh hưởng và chịu tác động của chỉ số thị trường.

3.2.1.5 Hiện thực phương pháp

Chúng tôi sẽ trình bày các bước để chọn được các danh mục đầu tư từ một tập hợp tất cả các cổ phiếu trong một sàn giao dịch chứng khoán như sau:

Vào một ngày “t” nhất định chúng tôi muốn dùng phương pháp phân cụm để lựa chọn danh mục đầu tư, chúng tôi sẽ thực hiện các bước như sau:

B1: Đọc từ file tất cả các dữ liệu cổ phiếu của một sàn nhất định.

B2: Sử dụng chuỗi giá đóng cửa trong khoảng thời gian $\{t - \delta t + 1, \dots, t\}$ để tính chuỗi trả về “r” cho tất cả các cổ phiếu. (phụ lục 7)

B3: Sử dụng chuỗi trả về “r” để tính hệ số tương quan Pearson giữa các cặp cổ phiếu. Sau đó xây dựng ma trận tương quan mà mỗi ô trong ma trận đại diện cho hệ số tương quan Pearson của một cặp cổ phiếu. (phụ lục 9)

B4: Sử dụng ma trận tương quan để xây dựng mạng tài chính “MST”. (phụ lục 12)

B5: Dùng các tham số tô pô của cây khung nhỏ nhất để chọn ra hai danh mục đầu tư là trung tâm và ngoại vi.

Sau đó dùng hai danh mục này để đầu tư.

3.2.2 Chiến lược lựa chọn danh mục đầu tư tránh sự lây lan rủi ro trên thị trường chứng khoán

3.2.2.1 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 các cổ phiếu có hiệu quả làm giảm rủi ro đầu tư. Họ cho rằng các khoản đầu tư vào các khu vực ngoại vi gồm các cổ phiếu kém kết nối trong mạng lọc tài chính sẽ cải thiện lợi nhuận. Ngược lại, đầu tư vào tập hợp con của các chứng khoán trung tâm gồm các chứng khoán có nhiều liên kết 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 (Minimum Spanning Trees) và đồ thị phẳng (Planar Maximally Filtered Graphs).

Xây dựng một mạng lọc tài chính, cụ thể là Minimum Spanning Trees (MST) và Planar Maximally Filtered Graphs (PMFG). MST là 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 phân loại 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 độ tương quan của các cổ phiếu qua khung thời gian sáu tháng. Sau đó chúng tôi sử dụng các mối tương quan để xây dựng các mạng lọc tài chính PMFG. Chúng tôi thiết lập danh mục đầu tư bằng cách chọn cổ phiếu từ khu vực ngoại vi và trung tâm của mạng được lọc tài chính. 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. Với mục đích này, chúng tôi phân biệt giữa các cổ phiếu nằm trong vùng trung tâm của mạng và những cổ phiếu nằm ở vùng ngoại vi.

3.2.2.2 Phương thức

Đối với mỗi danh mục đầu tư, chúng tôi chọn khung thời gian $\delta t = 125$, chuỗi trả về “r” bằng công thức sau:

$$r_i(t) = (P_i(t - \tau) - P_i(t - 125)) / P_i(t - 125) \quad [10]$$

Với $\tau = 0, 1, \dots, 125 - 1$. Chúng tôi giảm tập dữ liệu bằng cách chỉ giữ lại tập hợp mã cổ phiếu hàng đầu “hoạt động tốt nhất” của các cổ phiếu trong giai đoạn trước đó. Chúng tôi sẽ tính tỉ lệ $\frac{E[r]_{\Delta t}}{s_{\Delta t}}$ cho từng mã cổ phiếu, sau đó chúng tôi sẽ loại một nửa số mã cổ phiếu có tỷ lệ thấp nhất. Việc loại những cổ phiếu có tỷ lệ thấp sẽ giúp chúng tôi giữ lại được một nửa số cổ phiếu đang có hiệu quả tốt trong khoảng thời gian Δt .

Chúng tôi dùng những mã cổ phiếu còn lại để xây dựng ma trận tương quan Pearson. Sau đó sẽ dùng ma trận tương quan Pearson để xây dựng PMFG và lựa chọn danh mục đầu tư.

3.2.2.3 Xây dựng mạng tài chính

Sau khi tính được ma trận tương quan Pearson, chúng tôi sẽ dùng ma trận tương quan để xây dựng đồ thị PMFG. Thuật toán sẽ được thực hiện như sau:

- Sắp xếp tất cả các cặp cạnh có độ tương quan giảm dần để có được một danh sách có thứ tự \mathbf{l}_{sort} .
- Thêm một cạnh giữa cặp nút i và j dựa trên thứ tự của \mathbf{l}_{sort} chỉ khi đồ thị vẫn phẳng sau khi cạnh được thêm vào.
- Một đồ thị G được hình thành với số cạnh $N = 3(n - 2)$.

Đồ thị G chúng tôi thu được đại diện cho mạng tài chính mà mỗi đỉnh trong đồ thị tương ứng với một cổ phiếu trong sàn, các cạnh nối giữa các đỉnh là giá trị tương quan giữa hai đỉnh.

3.2.2.4 Lựa chọn danh mục đầu tư

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 có thể được sử dụng để xây dựng một danh mục đầu tư đa dạng hiệu quả làm giảm rủi ro đầu tư. Chúng tôi thấy rằng các khoản đầu tư vào các khu vực ngoại vi gồm các chứng khoán có kết nối yếu trong

mạng tài chính có hiệu quả tốt hơn so với các chứng khoán trung tâm. Cụ thể là cây khung tối thiểu (Minimum Spanning Trees) và đồ thị phẳng lọc tối đa phẳng (Planar Maximally Filtered Graphs). Ngược lại, đầu tư vào tập hợp con của các chứng khoán trung tâm, các chứng khoán có nhiều kết nối cao được đặc trưng bởi rủi ro lớn hơn và hình thành nhiều rủi ro hơn.

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 PMFG. Sau đó, đối với mỗi biện pháp xếp hạng 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

$$X = \frac{C_D^W + C_D^u + C_{BC}^W + C_{BC}^u - 4}{4*(N-1)} \quad [1, \text{tr.7}]$$

$$Y = \frac{C_E^W + C_E^u + C_C^W + C_C^u + C_{EC}^W + C_{EC}^u - 6}{6*(N-1)} \quad [1, \text{tr.7}]$$

nơi chúng tôi biểu thị với C_D^W xếp hạng gần liền của mức độ của bậc (D) . C_D^u 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.

- 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 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 đỉnh 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, để chọn được trung tâm và ngoại vi, chúng tôi nhận thấy những cổ phiếu trung tâm có giá trị X lớn hơn những cổ phiếu ngoại vi nên chúng tôi sẽ tính cho tỉ lệ X/Y cho từng cổ phiếu.

Cổ phiếu nằm ở trung tâm PMFG sẽ có tỉ lệ X/Y lớn và ngược lại.

Chúng tôi chọn ra hai danh mục đầu tư trung tâm và ngoại vi từ PMFG, mỗi tập có số lượng bằng 10% số cổ phiếu ban đầu.

3.2.2.5 Hiện thực phương pháp

Vào một ngày t nhất định chúng tôi muốn dùng phương pháp phân cụm để lựa chọn danh mục đầu tư, chúng tôi sẽ thực hiện các bước như sau:

B1: Đọc từ file tất cả các dữ liệu cổ phiếu của một sản nhất định.

B2: Sử dụng chuỗi giá đóng cửa trong khoảng thời gian $\{t - \delta t + 1, \dots, t\}$ để tính chuỗi trả về “r” cho tất cả các cổ phiếu. (phụ lục 7)

B3: Chúng tôi sẽ dùng chuỗi trả về “r” để tính tỉ lệ $\mathbf{E}[\mathbf{r}]_{\Delta t} / \mathbf{s}_{\Delta t}$ và sẽ loại đi một nửa số cổ phiếu có tỉ lệ $\mathbf{E}[\mathbf{r}]_{\Delta t} / \mathbf{s}_{\Delta t}$ thấp nhất.

B4: Sử dụng chuỗi trả về “r” để tính hệ số tương quan Pearson giữa các cặp cổ phiếu. Sau đó xây dựng ma trận tương quan mà mỗi ô trong ma trận đại diện cho hệ số tương quan Pearson của một cặp cổ phiếu. (phụ lục 9)

B5: Sử dụng ma trận tương quan để xây dựng mạng tài chính “PMFG”.(phụ lục 13)

B6: Dùng các tham số tô pô của cây khung nhỏ nhất để chọn ra hai danh mục đầu tư là trung tâm và ngoại vi.

Sau đó dùng hai danh mục này để đầu tư.

3.2.3 Chiến lược lựa chọn danh mục sử dụng tương quan chéo phụ thuộc q

3.2.3.1 Giới thiệu

Phân tích mối tương quan chéo giữa các tài sản tài chính khác nhau đã trở nên vô cùng hấp dẫn kể từ khi các nhà nghiên cứu bắt đầu báo cáo vi phạm của Efficient Market Hypothesis (EMH)[2][7]. Các tính chất của ma trận tương quan chéo phụ thuộc q của thị trường chứng khoán đã được phân tích bằng cách sử dụng lý thuyết ma trận ngẫu nhiên và mạng phức tạp. Các cấu trúc tương quan của các biến động ở các độ lớn khác nhau có các đặc tính độc đáo. Các tương quan chéo giữa các biến động nhỏ mạnh 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 phụ thuộc q này và khám phá một số nhận dạng mới. Bằng cách sử dụng các mạng tương quan phụ thuộc vào q 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.

3.2.3.2 Phương thức

Vào một ngày t , chúng tôi sẽ dùng dữ liệu trong khoảng $\{t - \delta t + 1, \dots, t\}$ để lựa chọn danh mục đầu tư, với $\delta t = 1500$ ngày.

Hệ số tương quan chéo phụ thuộc q được tính như sau:

(i) Chúng tôi sẽ tính chuỗi trả về ‘ r ’ của hai cổ phiếu như sau:

$$r_x(t) = P_x(t) - E[x], \quad t = 1 \dots \delta t \quad [4, \text{tr.3}]$$

$$r_y(t) = P_y(t) - E[x], \quad i = 1 \dots \delta t \quad [4, \text{tr.3}]$$

(ii) Chúng tôi chia 2 chuỗi đã được tích hợp thành các phân đoạn có chiều dài S .

Sau đó chúng tôi tính xu hướng địa phương cho từng phân đoạn. Sau đó chúng tôi tính tín hiệu còn lại của mỗi phân đoạn bằng tổng của các tín hiệu tích hợp sau khi được loại bỏ xu hướng:

$$X_v(s, i) = \sum_{j=1}^i X^x(vs + j) - P_{X,s,v}^m(j) \quad [4, \text{tr.2}]$$

$$Y_v(s, i) = \sum_{j=1}^i X^y(vs + j) - P_{Y,s,v}^m(j) \quad [4, \text{tr.2}]$$

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

$$f_{XY}^2(s, v) = \frac{1}{s} \sum_{i=1}^s X_v(s, i) Y_v(s, i) \quad [4, \text{tr.2}]$$

$$f_{ZZ}^2(s, v) = \frac{1}{s} \sum_{i=1}^s Z_v^2(s, i) \quad [4, \text{tr.2}]$$

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

$$F_{XY}^q(s) = \frac{1}{2M_s} \sum_{v=0}^{2M_s-1} \text{sgn}[f_{XY}^2(s, v)] |f_{XY}^2(s, v)|^{q/2} \quad [4, \text{tr.3}]$$

$$F_{ZZ}^q(s) = \frac{1}{2M_s} \sum_{v=0}^{2M_s-1} [f_{XY}^2(s, v)]^{q/2} \quad [4, \text{tr.3}]$$

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

$$p(q, s) = \frac{F_{XY}^q(s)}{\sqrt{F_{XX}^q(s) F_{YY}^q(s)}} \quad [4, \text{tr.3}]$$

3.2.3.3 Xây dựng mạng tài chính

Chúng tôi sẽ tính ma trận tương quan chéo mà mỗi ô trong ma trận ứng với giá trị tương quan chéo của một cặp cổ phiếu. Sau khi tính được ma trận tương quan tương quan chéo dựa trên q và s , chúng tôi sẽ dùng ma trận để xây dựng PMFG. Thuật toán sẽ được thực hiện như sau:

- Sắp xếp tất cả các cặp cạnh có độ tương quan giảm dần để có được một danh sách có thứ tự \mathbf{l}_{sort} .
- Thêm một cạnh giữa cặp nút i và j dựa trên thứ tự của \mathbf{l}_{sort} chỉ khi đồ thị vẫn phẳng sau khi cạnh được thêm vào.
- Một đồ thị G được hình thành với số cạnh $N = 3(n - 2)$.

3.2.3.4 Lựa chọn danh mục đầu tư

Chúng tôi chỉ sử dụng chỉ số phụ thuộc q vào mạng PMFG để lựa chọn chứng khoán và xác định 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 sử dụng một số biện pháp trung tâm thực hiện rất tốt. Ở đây đầu tiên chúng tôi tính điểm số trung tâm được xác định bởi

$$X = \frac{C_D^W + C_D^u + C_{BC}^W + C_{BC}^u - 4}{4 * (N - 1)}$$

$$Y = \frac{C_E^W + C_E^u + C_C^W + C_C^u + C_{EC}^W + C_{EC}^u - 6}{6 * (N - 1)}$$

Tính cho cả MST và PMFG trọng số và các đối tác không có trọng số của chúng. Cụ thể, để chọn được trung tâm và ngoại vi, chúng tôi nhận thấy những cổ phiếu trung tâm có giá trị X lớn hơn những cổ phiếu ngoại vi nên chúng tôi sẽ tính cho tỉ lệ X/Y cho từng cổ phiếu.

Cổ phiếu nằm ở trung tâm PMFG sẽ có tỉ lệ X/Y lớn và ngược lại.

Chúng tôi chọn ra hai danh mục đầu tư trung tâm và ngoại vi từ PMFG, mỗi tập có số lượng bằng 10% số cổ phiếu ban đầu.

3.2.3.5 Hiện thực phương pháp

Vào một ngày t nhất định chúng tôi muốn dùng phương pháp phân cụm để lựa chọn danh mục đầu tư, chúng tôi sẽ thực hiện các bước như sau:

B1: Đọc từ file tất cả các dữ liệu cổ phiếu của một sàn nhất định.

B2: Sử dụng chuỗi giá đóng cửa trong khoảng thời gian $\{t - \delta t + 1, \dots, t\}$ để tính chuỗi trả về “r” cho tất cả các cổ phiếu. (phụ lục 7)

B4: Sử dụng chuỗi trả về “r” để tính hệ số tương quan chéo giữa các cặp cổ phiếu. Sau đó xây dựng ma trận tương quan mà mỗi ô trong ma trận đại diện cho hệ số tương quan chéo của một cặp cổ phiếu. (phụ lục 10)

B5: Sử dụng ma trận tương quan để xây dựng mạng tài chính “PMFG”. (phụ lục 13)

B6: Dùng các tham số tô pô của cây khung nhỏ nhất để chọn ra hai danh mục đầu tư là trung tâm và ngoại vi.

Sau đó dùng hai danh mục này để đầu tư.

3.3 So sánh các kết quả thực nghiệm

3.3.1 Chọn 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 vì khoảng thời gian đầu tư ảnh hưởng trực tiếp đến lợi nhuận của các cổ phiếu đem lại cho nhà đầu tư. 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 khi đầu tư trong khoảng thời gian ba tháng, mười tháng.

3.3.2 Lựa chọn phương pháp tạo danh mục đầu tư tối ưu

3.3.2.1 Chiến lược danh mục động sử dụng phương pháp phân cụm

Chúng tôi hiện thực hóa tất cả các danh mục đầu tư theo các tham số tô pô của mạng tài chính. So sánh các danh mục theo các tham số với nhau và đưa ra kết luận danh mục theo tham số tối ưu nhất.

3.3.2.2 Lựa chọn danh mục đầu tư tránh sự lây lan rủi ro trên thị trường chứng khoán

Chúng tôi áp dụng đầu tư cho cả hai tập danh mục trung tâm và ngoại vi để so sánh kết quả để biết được chiến lược áp dụng với loại danh mục đầu tư nào là tốt nhất.

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

Với mỗi mạng tài chính đã xây dựng được ở phương pháp này chúng tôi sẽ chọn ra chỉ số phụ thuộc q lần lượt bằng 1, 2, 3, 4 tương ứng với khung thời gian tương ứng là 50, 100, 150, 200.

So sánh lợi nhuận trung bình của các danh mục đầu tư này và như thế chúng tôi sẽ tìm được chiến lược tối ưu nhất.

3.3.3 Phương pháp so sánh các chiến lược

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 vì khoảng thời gian đầu tư ảnh hưởng trực tiếp đến lợi nhuận của các cổ

phiếu đem lại cho nhà đầu tư. 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 khi đầu tư trong khoảng thời gian mười tháng.

Chúng tôi sẽ so sánh kết quả của các chiến lược. Bằng cách so sánh tổng lợi nhuận trung bình của danh mục đầu tư tối ưu nhất của từng chiến lược. Từ đó chúng ta sẽ rút ra chiến lược đầu tư tốt nhất áp dụng cho nhiều sản phẩm giao dịch.

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

4.1 Dữ liệu chứng khoán

Chúng tôi đã thu thập được dữ liệu cổ phiếu của mười hai sàn giao dịch. Trong đó bao gồm các sàn chứng khoán nội địa và các sàn nước ngoài.

Thông tin thu thập dữ liệu của các sàn như sau :

Tên sàn giao dịch	Quốc gia	Số cổ phiếu
HOSE	Việt Nam	90
HNX	Việt Nam	60
NYSE	Mỹ	200
AMEX	Mỹ	119
OSLOBORS	Na Uy	116
NASDAQ	Mỹ	112
AustraliaS&P200	Úc	58
KOSPI	Hàn Quốc	228
NIKKEI225	Nhật	173
TSX	Canada	153
XU100	Thổ Nhĩ Kỳ	64
SHANGHAI	Trung Quốc	57

Bảng 4.1 Thông tin các dữ liệu đã thu thập được

4.2 Các chiến lược lựa chọn danh mục đầu tư

Sau khi hiện thực được ba chiến lược chúng tôi sẽ tiến hành đầu tư liên tục tất cả các chiến lược trong khoảng thời gian từ ngày 1 tháng 6 năm 2014 đến ngày 1 tháng 6 năm 2017. Bởi vì mỗi chiến lược có nhiều cách để lựa chọn danh mục đầu tư cho nên chúng tôi sẽ tìm hiểu xem cách lựa chọn danh mục đầu tư nào đem lại hiệu quả tốt nhất.

4.2.1 Chiến lược danh mục động sử dụng phương pháp phân cụm

Chiến lược này chúng ta sẽ có năm cách để lựa chọn cổ phiếu khác nhau dựa vào tham số tô pô của “MST”.

Kết quả thực nghiệm của tất cả các phương pháp chọn danh mục đầu tư (phụ lục 4). Sau khi xem xét chúng tôi lựa chọn được cách dùng tham số tô pô Dcorrelation. Chúng tôi sẽ lựa chọn kết quả của Dcorrelation để so sánh với kết quả của các chiến lược khác.

4.2.2 Chiến lược lựa chọn danh mục đầu tư tránh sự lây lan rủi ro trên thị trường chứng khoán

Xem kết quả thực nghiệm ở phụ lục 3.

4.2.3 Chiến lược lựa chọn danh mục sử dụng tương quan chéo phụ thuộc q

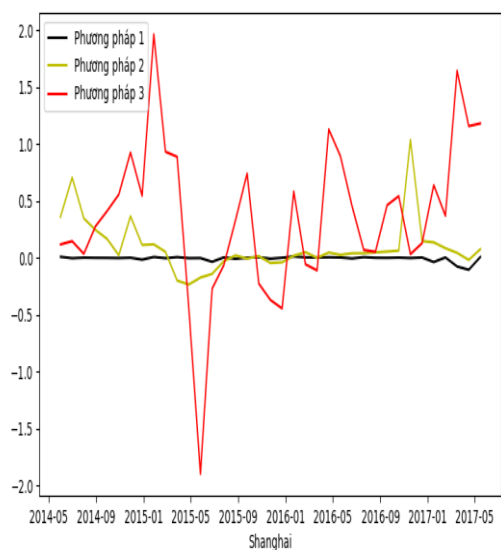
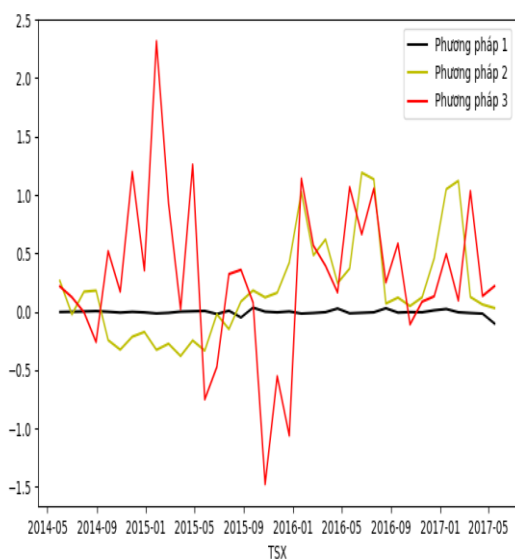
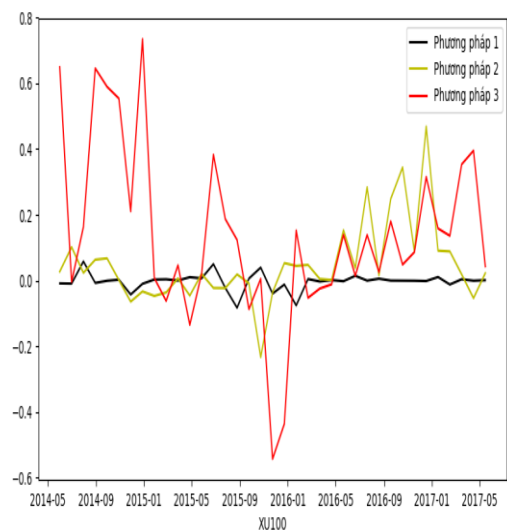
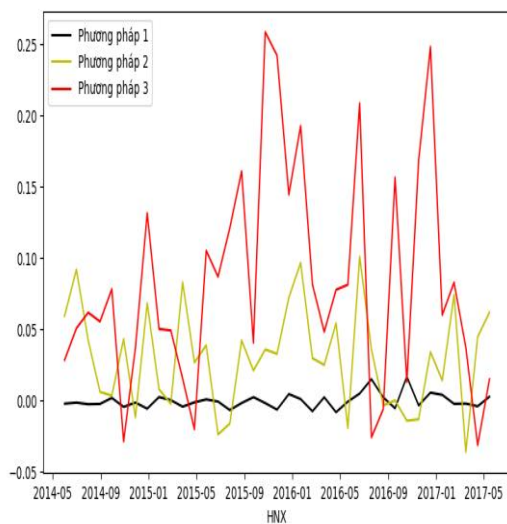
Chiến lược này phụ thuộc vào các cặp “ q, s ” khác nhau thì thu được các “PMFG” khác nhau và sẽ thu được các danh mục đầu tư tốt xấu khác nhau.

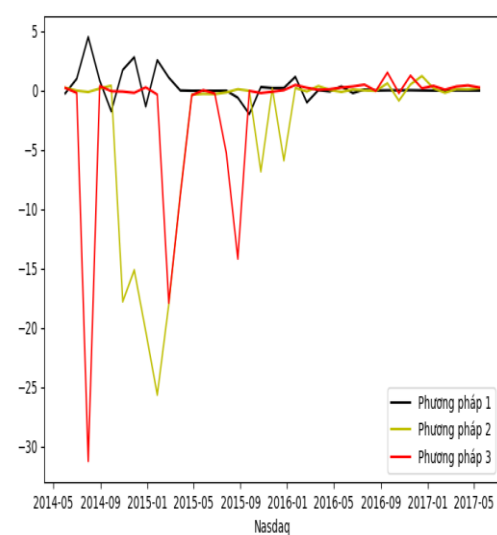
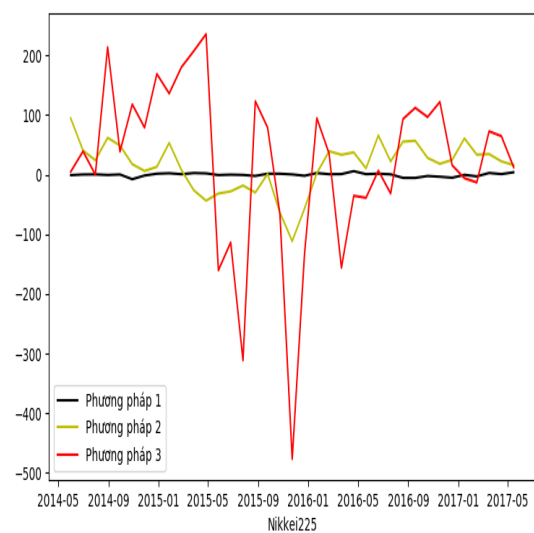
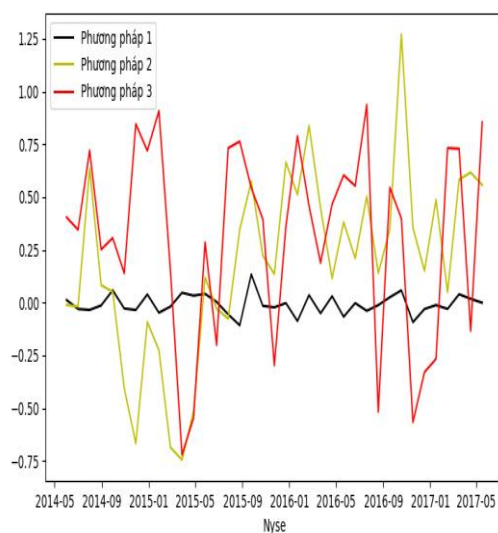
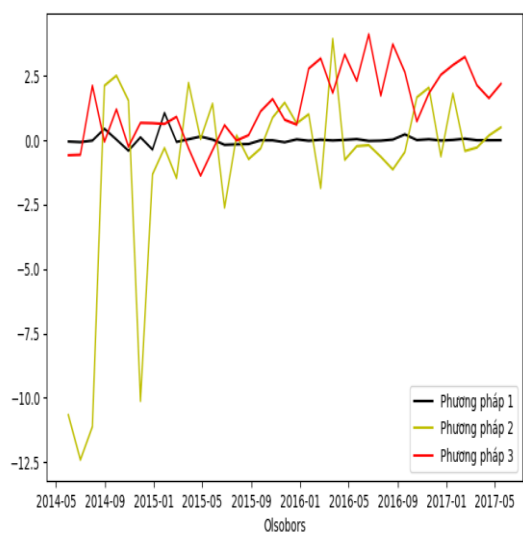
Chúng tôi sẽ xem xét kết quả thực nghiệm (phụ lục 1 và 2) để tìm ra cặp “ q, s ” có hiệu quả tốt nhất.

Sau khi xem xét chúng tôi chọn $q = 2$ và $s = 150$. Việc chọn $q = 2$ vì chúng tôi nhận thấy kết quả đầu tư chênh lệch rất ít khi các giá trị “ q ” thay đổi. Việc chọn $s = 150$ vì khi $s = 150$ sẽ đem lại lợi nhuận tốt hơn, ổn định hơn và ít bị lỗ hơn các giá trị “ s ” khác.

4.3 So sánh kết quả

Sau đây chúng tôi xem xét các biểu đồ bên dưới và đưa ra những đánh giá về hiệu quả của các chiến lược lựa chọn danh mục đầu tư.





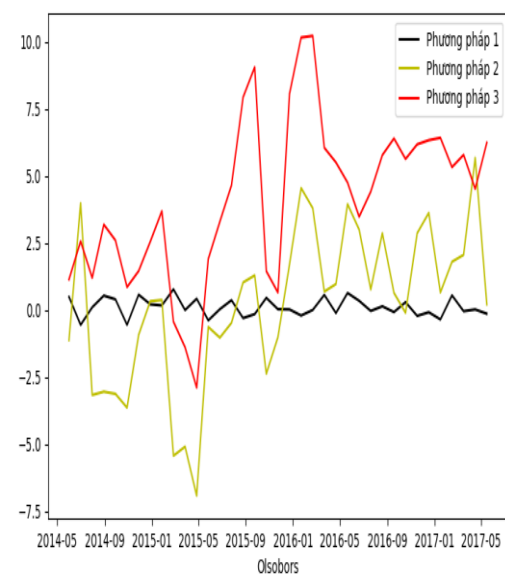
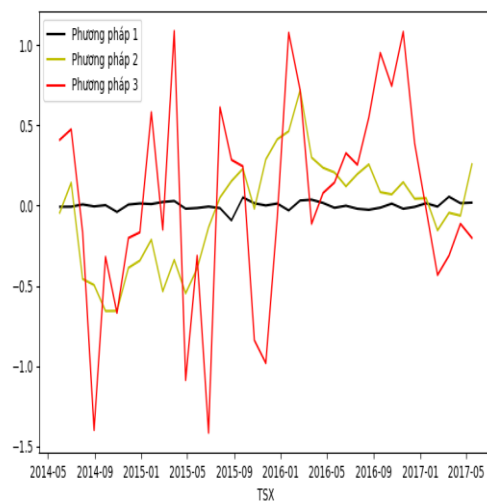
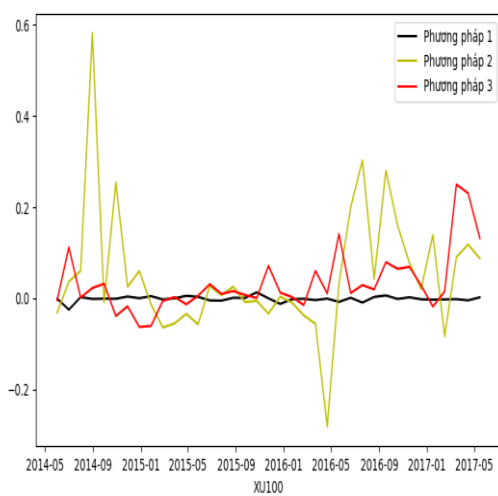


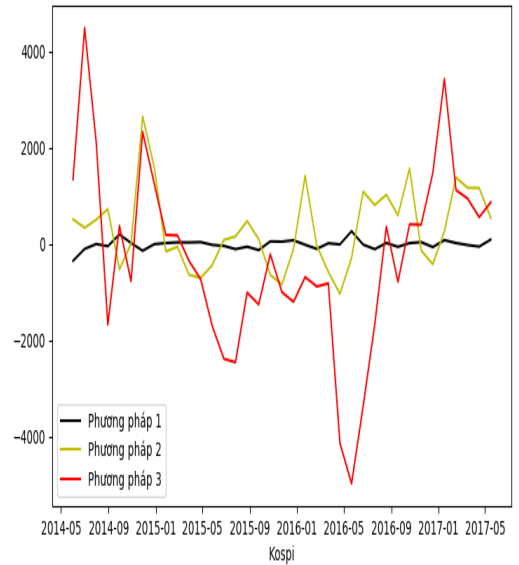
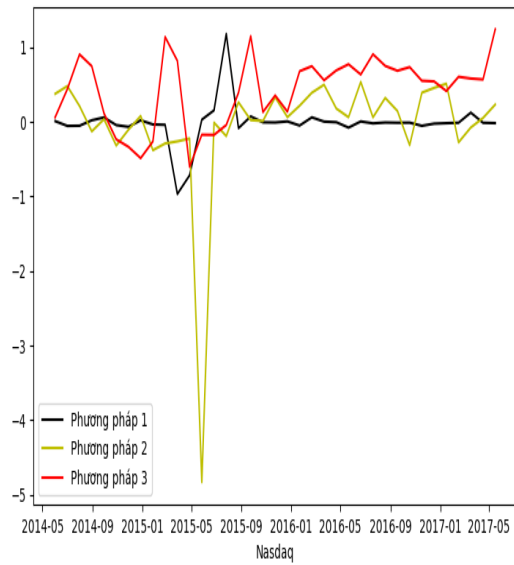
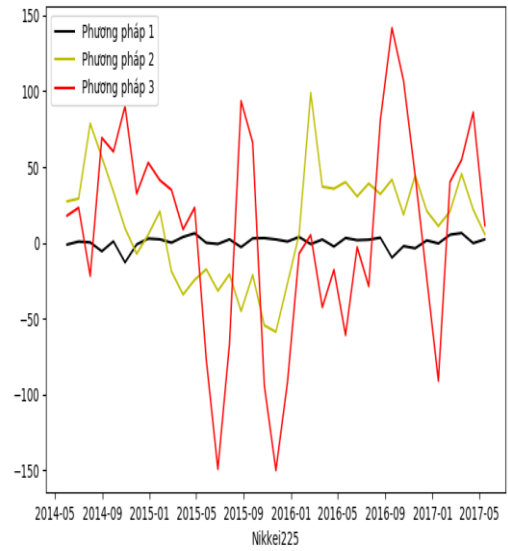
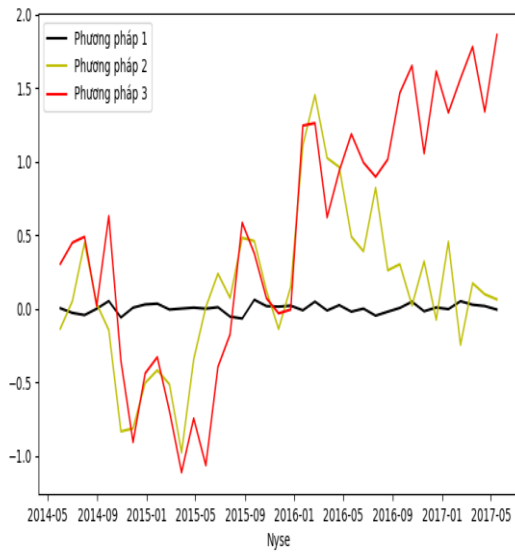
Hình 4.1 Biểu đồ kết quả của các sàn theo danh mục đầu tư trung tâm

Kết quả theo danh mục đầu tư trung tâm theo đồ thị cho chúng ta thấy sự hiệu quả của cả ba chiến lược.

- Chiến lược một không hiệu quả cho với danh mục trung tâm.

- Hai chiến lược còn lại nhận được kết quả tốt, đạt lợi nhuận đáng kể.







Hình 4.2 Biểu đồ kết quả của các sản theo danh mục đầu tư ngoại vi

Các biểu đồ kết quả của danh mục ngoại vi cho chúng ta thấy chiến lược đầu tư gặp nhiều biến động hơn, chịu nhiều rủi ro hơn.

Sau khi nhìn tổng quan các kết quả thực nghiệm của tất cả các sản. Chúng tôi nhận thấy chiến lược danh mục động sử dụng phương thức phân cụm thực sự không đem lại hiệu quả. Hai chiến lược còn lại đem lại hiệu quả tốt. Bên cạnh đó, có những giai đoạn phải chịu nhiều rủi ro lớn.

Chiến lược danh mục động sử dụng phương thức phân cụm: Chiến lược an toàn, luôn ổn định không gặp nhiều rủi ro nhưng cũng không mang lại hiệu quả mà nhà đầu tư mong muốn.

Chiến lược lựa chọn danh mục đầu tư tránh sự lây lan rủi ro trên thị trường cũng đem lại lợi nhuận không quá cao nhưng không gặp nhiều rủi ro.

Chiến lược lựa chọn danh mục sử dụng tương quan chéo phụ thuộc q đem lại lợi nhuận lớn vượt trội hơn hai chiến lược trên nhưng chịu rủi ro cũng lớn nhưng việc chịu rủi ro không đáng kể so với hiệu quả đạt được.

CHƯƠNG 5 – KẾT LUẬN

5.1 Kết luận

Luận văn “Hiện thực và so sánh các mô hình toán học trên thị trường chứng khoán” đã thực hiện được mục tiêu nghiên cứu.

- Thứ nhất, luận văn đã tổng quan các mô hình cũng như các phương pháp. Luận văn đã tổng quan được các nghiên cứu trên mười hai thị trường chứng khoán trên thế giới.
- Thứ hai, Luận văn đã đề xuất được danh mục đầu tư có lợi nhuận trên tất cả các thị trường trên thế giới. Kết quả thực nghiệm cho thấy sự biến động của các mô hình theo từng thị trường cũng như mức độ phụ thuộc của loại cổ phiếu.

5.2 Đề xuất các hướng nghiên cứu tiếp theo

Để tiếp tục phát triển thị trường chứng khoán và giảm thiểu rủi ro trên thị trường chứng khoán nên các hướng nghiên cứu trong tương lai có thể thực hiện:

- Mở rộng sự tương quan của các thị trường chứng khoán chỉ không chỉ là sự tương quan của các cổ phiếu trong một sàn.
- Phát triển ứng dụng có thể tự động tham gia giao dịch đem lợi nhuận về cho người sử dụng.

TÀI LIỆU THAM KHẢO

Tiếng Anh

1. F. Pozzi, T. Di Matteo & T. Aste. (2013), Spread of risk across financial markets: better to invest in the peripheries.
2. Hadrien Salat, Roberto Murcio, Elsa Arcaute. (2016), Multifractal methodology. *Physica A*.
3. Laurent Laloux, Pierre Cizeau, Jean-Philippe Bouchaud, Marc Potters. (1999), Noise Dressing of Financial Correlation Matrices. *Physical Review Letters*, pp. 1467–1470.
4. Longfeng Zhaoa, Wei Lia, Andrea Fenub, Boris Podobnikb, Yougui Wang, H. Eugene Stanley. (2017), The q-dependent detrended cross-correlation analysis of stock market.
5. Mantegna RN. (1999), Hierarchical structure in financial markets, pp. 193–197.
6. Onnela JP, Kaski K, Kertész J.(2004), Clustering and information in correlation based financial networks. *Eur Phys J B*. pp. 353–362.
7. Tumminello, M., Aste, T., Di Matteo, T. & Mantegna, R. N. (2005), A tool for filtering information in complex systems, *PNAS*, pp. 10421–10426.
8. Vasiliki Plerou, Parameswaran Gopikrishnan, Bernd Rosenow, Lu'is Nunes Amaral, H. Stanley. (1999), Universal and Nonuniversal Properties of Cross Correlations in Financial Time Series. *Physical Review Letters*, pp. 1471–1474.
9. West DB. (1996), *Introduction to Graph Theory*, Prentice-Hall.
10. Ya-Nan Lu, Sai-Ping Li, Xiong-Fei Jiang, Li-Xin Zhong, Tian Qiu. (2017), Dynamic Portfolio Strategy Using Clustering Approach.

PHỤ LỤC

Phụ lục 1: Bảng kết quả danh mục đầu tư trung tâm của phương pháp lựa chọn danh mục sử dụng tương quan chéo phụ thuộc q .

	Q = 1	Q = 2	Q = 3	Q = 4	
HOSE	3.75463	3.75463	3.75463	3.75463	S = 50
	3.52848	3.52848	3.52848	3.52848	S = 100
	2.90714	2.90714	2.90714	2.90714	S = 150
	3.40786	3.40786	3.40786	3.40786	S = 200
HNX	0.26285	0.26285	0.26285	0.26285	S = 50
	0.20398	0.20398	0.20398	0.20398	S = 100
	0.48087	0.48087	0.48087	0.48087	S = 150
	0.12701	0.12701	0.12701	0.12701	S = 200
NYSE	7.82341	7.82341	15.61300	15.61300	S = 50
	11.52869	13.55722	13.55722	13.55722	S = 100
	12.09821	18.45782	18.45782	18.45782	S = 150
	9.06077	12.99458	12.99458	12.99458	S = 200
AMEX	5.32541	5.32541	5.32541	5.32541	S = 50
	6.30394	6.30394	6.30394	6.30394	S = 100
	6.37189	6.37189	6.37189	6.37189	S = 150
	6.61549	6.61549	6.61549	6.61549	S = 200
OLSOBORS	145.82633	145.82633	145.82633	145.82633	S = 50
	141.07013	141.07013	141.07231	141.07231	S = 100
	154.66823	154.66823	153.67921	153.67921	S = 150
	148.20766	148.20766	148.20766	148.90263	S = 200
NASDAQ	23.09840	23.09840	23.09840	23.09840	S = 50

	14.74788	14.74788	14.74788	14.74788	S = 100
	15.79620	15.79620	15.79620	15.79620	S = 150
	16.25060	16.25060	16.25060	16.25060	S = 200
NIKKEI225	395.63899	309.28129	309.28129	309.2812 9	S = 50
	386.05138	295.09345	295.09345	295.09345	S = 100
	107.31043	107.31043	107.31043	107.31043	S = 150
	217.72988	217.72988	217.72988	217.72988	S = 200
TSX	-0.74459	-0.74459	4.52798	4.52798	S = 50
	-1.48634	-1.48634	2.96280	2.96280	S = 100
	-0.53690	-0.53690	2.29487	2.29487	S = 150
	-0.88548	1.86861	1.86861	1.86861	S = 200
XU100	2.25714	2.25714	2.25714	2.25714	S = 50
	1.32738	1.32738	1.32738	1.32738	S = 100
	1.23405	1.23405	1.23405	1.23405	S = 150
	0.33403	0.33403	0.334036	0.334036	S = 200
S&P	0.74358	0.74358	0.74358	0.74358	S = 50
	-0.02359	-0.02359	-0.02359	-0.02359	S = 100
	-0.15276	-0.15276	-0.15276	-0.15276	S = 150
	0.77099	0.77099	0.77099	0.77099	S = 200
Shanghai	1.14483	1.14483	1.14483	1.14483	S = 50
	1.05625	1.05625	1.05625	1.05625	S = 100
	1.07556	1.07556	1.07556	1.07556	S = 150
	1.11260	1.11260	1.11260	1.11260	S = 200
KOSPI	18103.78019	18103.78019	18103.78019	18103.78019	S = 50
	17160.86931	17160.86931	17160.86931	17160.86931	S = 100

	-10209.18720	-10209.18720	-10209.18720	-10209.18720	S = 150
	8329.49593	8329.49593	8329.49593	8329.49593	S = 200

Phục lục 2: Bảng kết quả danh mục đầu tư ngoại vi của phương pháp lựa chọn danh mục sử dụng tương quan chéo phụ thuộc q .

	Q = 1	Q = 2	Q = 3	Q = 4	
HOSE	6.14503	6.14503	6.14503	6.14503	S = 50
	7.14467	7.14467	7.14467	7.14467	S = 100
	5.59229	5.59229	5.59229	5.59229	S = 150
	5.01294	5.01294	5.01294	5.01294	S = 200
HNX	2.38231	2.38231	2.38231	2.38231	S = 50
	2.14277	2.14277	2.14277	2.14277	S = 100
	3.08094	3.08094	3.08094	3.08094	S = 150
	2.91629	2.91629	2.91629	2.91629	S = 200
NYSE	2.72682	2.72682	11.38577	11.38577	S = 50
	8.49763	9.75832	9.75832	9.75832	S = 100
	6.13988	11.53866	11.53866	11.53866	S = 150
	4.58272	15.20124	15.20124	15.20124	S = 200
AMEX	-5.04066	-5.04066	-5.04066	-5.04066	S = 50
	0.71831	0.71831	0.71831	0.71831	S = 100
	11.53041	11.53041	11.53041	11.53041	S = 150
	9.31617	9.31617	9.31617	9.31617	S = 200
OLSOBO RS	44.10619	44.10619	44.10619	44.10619	S = 50
	44.66543	44.66543	44.66543	44.66543	S = 100
	50.48123	50.48123	50.45633	50.45633	S = 150

	40.45728	40.45728	40.45728	40.52709	S = 200
NASDAQ	-117.76741	-117.76741	-117.76741	-117.76741	S = 50
	-116.82778	-116.82778	-116.82778	-116.82778	S = 100
	-72.03583	-72.03583	-72.03583	-72.02034	S = 150
	-14.21345	-14.21345	-14.21345	-14.21345	S = 200
NIKKEI2 25	863.38163	646.28915	646.28915	646.28915	S = 50
	472.53663	647.96939	650.33392	647.96939	S = 100
	458.81801	458.81801	458.81801	458.81801	S = 150
	566.22697	566.22697	566.22697	566.22697	S = 200
TSX	3.21242	3.21242	11.36908	11.36908	S = 50
	8.28843	8.28843	14.40499	14.40499	S = 100
	5.97006	5.97006	10.94591	10.94591	S = 150
	7.04196	14.71386	14.71386	14.71386	S = 200
XU100	3.57569	3.57569	3.57569	3.57569	S = 50
	3.20124	3.20124	3.20124	3.20124	S = 100
	5.15228	5.15228	5.15228	5.15228	S = 150
	7.12078	7.12078	7.12078	7.12078	S = 200
S&P	1.46032	1.46032	1.46032	1.46032	S = 50
	1.75989	1.75989	1.75989	1.75989	S = 100
	1.00531	1.00531	1.00531	1.00531	S = 150
	-0.53236	-0.53236	-0.53236	-0.53236	S = 200
Shanghai	6.27927	6.27927	6.27927	6.27927	S = 50
	5.23429	5.23429	5.23429	5.23429	S = 100
	5.36583	5.36583	5.36583	5.36583	S = 150
	9.32067	9.32067	9.32067	9.32067	S = 200

KOSPI	-8324.21816	-8324.21816	-8324.21816	-8324.21816	S = 50
	-4836.98342	-4836.98342	-4836.98342	-4836.98342	S = 100
	14651.55059	14651.55059	14651.55059	14651.55059	S = 150
	1480.34823	1480.34823	1480.34823	1480.34823	S = 200

Phục lục 3: Bảng kết quả danh mục đầu tư trung tâm và ngoại vị của chiến lược lựa chọn danh mục đầu tư tránh sự lây lan rủi ro trên thị trường.

		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

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
Austral- iaS&P200	central	-0.0385	0.2564	0.1846
	peripheral	0.9001	0.0536	0.0077
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
XU100	central	2.319801	0.7061	1.925644
	peripheral	3.79650	0.9660	2.1631

Phục lục 4: Bảng kết quả thuận toán 1 3 tháng

		K	C	D_correlation	D_degree	D_distance
AMEX	central	-0.02267	-0.02941	0.0397	-0.7746	-0.00828
	peripheral	-0.0573	-0.0029	-0.0175	-0.00793	-0.04286
Austral- iaS&P200	central	-0.0151	-0.004719	-0.00682	-0.0065	-0.001137

	peripheral	-0.119	-0.0635	-0.02477	-0.0495	-0.03302
HNXIN-DEX	central	-0.0352	-0.061	0.017403	-0.0205	0.017403
	peripheral	0.0012	0.005	-0.00724	0.01267	-0.0072
VNINDEX	central	-0.00861	-0.076	-0.0074719	-0.0888	-0.01169
	peripheral	-0.0331	-0.166	0.00548	0.06768	0.00548
KOSPI	central	375.8154	-238.946	-306.830	580.3652	276.1231
	peripheral	1043.6692	-378.684	90.066	381.069	1633.8159
NASDAQ	central	-0.3296	-0.19866	-0.59527	-0.47095	0.08133
	peripheral	-0.4918	-0.49188	9.7959	9.4970	-0.481
NIK-KEI225	central	35.57861	29.9940	15.3560	22.1948	13.9215
	peripheral	1.2509	1.25095	-4.5879	9.2697	-3.0673
NYSE	central	-0.15488	-0.228855	0.057734	-0.3061	0.05773
	peripheral	-0.1875	-0.300	-0.27123	-0.1792	-0.2712
OLSOS-BORS	central	0.05788	-0.98614	3.88571	1.0191	0.41983
	peripheral	0.0270	0.027056	0.77488	-0.5882	-0.1294
Shanghai	central	-0.030043	-0.033091	0.01256	-0.0215	0.00184
	peripheral	-0.155924	-0.15592	-0.2439834	-0.2980	-0.03094
TSX	central	-0.031763	-0.01399	0.018656	-0.07798	-0.01951

	peripheral	0.05698	0.05698	-0.16592	-0.04549	-0.0591
XU100	central	-0.02689	-0.03316	-0.040654	-0.04694	-0.01767
	peripheral	-0.069	-0.0696	-0.11364	-0.07764	-0.06781

Phục lục 5: Bảng kết quả danh mục đầu tư của chiến lược danh mục động sử dụng phương pháp phân cụm.

		K	C	D_correlat ion	D_degr ee	D_distance
AMEX	central	-0.02267	-0.02941	0.0397	-0.7746	-0.00828
	peripheral	-0.0573	-0.0029	-0.0175	-0.00793	-0.04286
Austral- iaS&P200	central	-0.0151	-0.004719	-0.00682	-0.0065	-0.001137
	peripheral	-0.119	-0.0635	-0.02477	-0.0495	-0.03302
HNXIN- DEX	central	-0.0352	-0.061	0.017403	-0.0205	0.017403
	peripheral	0.0012	0.005	-0.00724	0.01267	-0.0072
VNINDEX	central	-0.00861	-0.076	-0.0074719	-0.0888	-0.01169
	peripheral	-0.0331	-0.166	0.00548	0.06768	0.00548
KOSPI	central	375.8154	-238.946	-306.830	580.3652	276.1231
	peripheral	1043.6692	-378.684	90.066	381.069	1633.8159
NASDAQ	central	-0.3296	-0.19866	-0.59527	-0.47095	0.08133
	peripheral	-0.4918	-0.49188	9.7959	9.4970	-0.481

NIK-KEI225	central	35.57861	29.9940	15.3560	22.1948	13.9215
	peripheral	1.2509	1.25095	-4.5879	9.2697	-3.0673
NYSE	central	-0.15488	-0.228855	0.057734	-0.3061	0.05773
	peripheral	-0.1875	-0.300	-0.27123	-0.1792	-0.2712
OLSOS-BORS	central	0.05788	-0.98614	3.88571	1.0191	0.41983
	peripheral	0.0270	0.027056	0.77488	-0.5882	-0.1294
Shanghai	central	-0.030043	-0.033091	0.01256	-0.0215	0.00184
	peripheral	-0.155924	-0.15592	-0.2439834	-0.2980	-0.03094
TSX	central	-0.031763	-0.01399	0.018656	-0.07798	-0.01951
	peripheral	0.05698	0.05698	-0.16592	-0.04549	-0.0591
XU100	central	-0.02689	-0.03316	-0.040654	-0.04694	-0.01767
	peripheral	-0.069	-0.0696	-0.11364	-0.07764	-0.06781

Phụ lục 6 : Script python “Tải mã cổ phiếu” từ các website

```
def crawl_stockID_in_tradingeconomic(driver, market_name, save_to):
    stockIDs = []
    driver.get('https://tradingeconomics.com/' + market_name + '/stock-market')
    showmore = driver.find_element_by_css_selector('svg')
    showmore.click()

    div_contains_table = driver.find_element_by_class_name('table-minimize')

    rows = div_contains_table.find_elements_by_css_selector('tr')
    for r in rows[1:]:
        txt = r.text
        stockID = txt.split(' ')[0]
        print(stockID)
```

```
stockIDs.append(stockID)
```

```
save_list_stockID_to_file(stockIDs, save_to)
```

Phụ lục 7: Script python “Tải dữ liệu cổ phiếu” từ các website

```
for i in stockIDs:
    driver.get('https://finance.yahoo.com/quote/' + i + '.IS/history?')
    time.sleep(1)
    try:
        a = driver.find_element_by_link_text('Download Data')
        a.click()
    except NoSuchElementException:
        print(i)
        continue

    time.sleep(1)
```

Phụ lục 8: Script python “Tính chuỗi r”

```
def calculate_r(time_series):
    r = []
    for i in range(0, len(list) - 1):
        difference = math.log(list[i] + 0.01, 2.718) - math.log(list[i + 1] +
0.01, 2.718)
        r.append(difference)

    return r
```

```
def calculate_r(time_series):
    r = []
    price_in_first_day = stock_returns[0]
    for i in range(1, len(stock_returns)):
        difference = (stock_returns[i] - price_in_first_day) /
price_in_first_day
        r.append(difference)

    return r
```

```
def calculate_r(time_series):
```

```

r = []
for i in range(0, len(time_series) - 1):
    difference = math.log(time_series[i] + 0.01, 2.718) -
                math.log(time_series[i + 1] + 0.01, 2.718)
    r.append(difference)

return r

```

Phụ lục 9: Script python “Tính giá trị kì vọng”

```

def calculate_expected(l):
    expected = 0
    for item in l:
        expected += item

    return expected / len(l)

```

Phụ lục 10: Script python “Xây dựng ma trận tương quan Pearson”

```

def build_pearson_correlation_matrix(stocks):
    matrix = np.empty([len(stocks), len(stocks)], dtype=np.float)
    for i in range(0, len(stocks)):
        for j in range(i, len(stocks)):
            if i == j:
                matrix[i][i] = 0
                continue
            p = pearsonr(stocks[i].r, stocks[j].r)[0]
            matrix[i][j] = round(p, 6)
            matrix[j][i] = round(p, 6)
    return matrix

```

Phụ lục 11: Script python “Tính hệ số tương quan chéo”

```

def qdependent_cc_coefficient(dcca_stock_1, dcca_stock_2, Q, S):
    lenght_integrated_ts = min(len(dcca_stock_1.intgrated_ts),
                                len(dcca_stock_2.integrated_ts))
    segment_count = int(lenght_integrated_ts / S)
    segments = [Segment()] * segment_count
    for i in range(0, segment_count):
        # Get integrated ts in segment
        s1_integrated = dcca_stock_1.integrated_ts[i * S : i * S + S]

```

```

        s2_integrated = dcca_stock_2.integrated_ts[i * S : i * S + S]
s1_rs = calculate_residual_signals(s1_integrated)
s2_rs = calculate_residual_signals(s2_integrated)
# calculate variance and covariance
s1_variance = calculate_variance(s1_rs, S)
s2_variance = calculate_variance(s2_rs, S)
covariance = (1 / S) * sum(s1_rs[i] * s2_rs[i] for i in range(0, S))
segments[i].set_s1_variance(s1_variance)
segments[i].set_s2_variance(s2_variance)
segments[i].set_covariance(covariance)
#Defined the fluctuation functions
F_covariance = (1 / segment_count) *
sum(np.sign(s.covariance) *
abs(s.covariance)**(Q/2) for s in segments)
F_s1_variance = (1 / segment_count) *
sum(s.s1_variance**(Q/2) for s in segments)
F_s2_variance = (1 / segment_count) *
sum(s.s2_variance**(Q/2) for s in segments)
return F_covariance / (F_s1_variance * F_s2_variance)**(1/2.0)

```

Phụ lục 12: Script python “Xây dựng ma trận tương quan chéo”

```

def build_crosscorrelation_matrix(dcca_stocks, Q, S):
    matrix = np.empty([len(dcca_stocks), len(dcca_stocks)], dtype=np.float)
    for i in range(0, len(dcca_stocks)):
        for j in range(i, len(dcca_stocks)):
            if i == j:
                matrix[i][i] = 0
                continue
            p = qdependent_cc_coefficient(dcca_stocks[i], dcca_stocks[j], Q, S)
            matrix[i][j] = round(p, 6)
            matrix[j][i] = round(p, 6)
    return matrix

```

Phụ lục 13: Script python “Xây dựng MST”

```

def build_MST(distance_matrix):
    G = nx.Graph()
    for stock in stocks:
        G.add_node(stock.ticker)

```

```

prim(distance_matrix)
for i in range(1, len(parent)):
    G.add_edge(stocks[parent[i]].ticker, stocks[i].ticker, weight=key[i])
return G

```

Phụ lục 14: Script python “Xây dựng PMFG”

```

def build_PMFG(G):
    sorted_edges = sort_graph_edges(G)
    PMFG_graph = compute_PMFG(sorted_edges, len(G.nodes))
    return PMFG_graph

```