

RESEARCH ARTICLE

# Dynamic Portfolio Strategy Using Clustering Approach

Fei Ren<sup>1,2\*</sup>, Ya-Nan Lu<sup>1</sup>, Sai-Ping Li<sup>3</sup>, Xiong-Fei Jiang<sup>4</sup>, Li-Xin Zhong<sup>5</sup>, Tian Qiu<sup>6</sup>

**1** School of Business, East China University of Science and Technology, Shanghai 200237, China, **2** Research Center for Econophysics, East China University of Science and Technology, Shanghai 200237, China, **3** Institute of Physics, Academia Sinica, Taipei 115 Taiwan, **4** College of Information Engineering, Ningbo Dahongying University, Ningbo 315175, China, **5** School of Finance, Zhejiang University of Finance and Economics, Hangzhou 310018, China, **6** School of Information Engineering, Nanchang Hangkong University, Nanchang 330063, China

\* [fren@ecust.edu.cn](mailto:fren@ecust.edu.cn)



## Abstract

The problem of portfolio optimization is one of the most important issues in asset management. We here propose a new dynamic portfolio strategy based on the time-varying structures of MST networks in Chinese stock markets, where the market condition is further considered when using the optimal portfolios for investment. A portfolio strategy comprises two stages: First, select the portfolios by choosing central and peripheral stocks in the selection horizon using five topological parameters, namely degree, betweenness centrality, distance on degree criterion, distance on correlation criterion and distance on distance criterion. Second, use the portfolios for investment in the investment horizon. The optimal portfolio is chosen by comparing central and peripheral portfolios under different combinations of market conditions in the selection and investment horizons. Market conditions in our paper are identified by the ratios of the number of trading days with rising index to the total number of trading days, or the sum of the amplitudes of the trading days with rising index to the sum of the amplitudes of the total trading days. We find that central portfolios outperform peripheral portfolios when the market is under a drawup condition, or when the market is stable or drawup in the selection horizon and is under a stable condition in the investment horizon. We also find that peripheral portfolios gain more than central portfolios when the market is stable in the selection horizon and is drawdown in the investment horizon. Empirical tests are carried out based on the optimal portfolio strategy. Among all possible optimal portfolio strategies based on different parameters to select portfolios and different criteria to identify market conditions, 65% of our optimal portfolio strategies outperform the random strategy for the Shanghai A-Share market while the proportion is 70% for the Shenzhen A-Share market.

## OPEN ACCESS

**Citation:** Ren F, Lu Y-N, Li S-P, Jiang X-F, Zhong L-X, Qiu T (2017) Dynamic Portfolio Strategy Using Clustering Approach. PLoS ONE 12(1): e0169299. doi:10.1371/journal.pone.0169299

**Editor:** Alejandro Raul Hernandez Montoya, Universidad Veracruzana, MEXICO

**Received:** July 16, 2016

**Accepted:** December 14, 2016

**Published:** January 27, 2017

**Copyright:** © 2017 Ren et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data are owned by and available from the Beijing Gildata RESSET Data Technology Co., Ltd, see <http://www.resset.cn/>. Researchers can access the relevant data through this link. We have also provided a Supplementary Information file "S1 File" containing the data sources and the specific URLs and steps to download the raw data used in the paper. Our university purchases the right to use the RESSET database. We use the account and the password offered by the RESSET Company, and log in to the webpage of RESSET to download the data.

**Funding:** This work was partially supported by the National Natural Science Foundation (Nos. 10905023, 71131007, 71371165, and 11501199), Fok Ying Tong Education Foundation Grant (No. 132013), Ningbo Natural Science Foundation (No. 2015A610160), and the Fundamental Research Funds for the Central Universities (2015). Fei Ren designed the study and performed the analysis. Sai-Ping Li prepared the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Introduction

Portfolio management is one of the hottest issues in finance. It primarily concerns with the best combination of securities for specific profits that investors need. The fundamental theory about portfolio optimization can be traced back to the Markowitz framework [1], which selects the allocation of investors' investment based on a mean-variance analysis. A great deal of work has subsequently contributed to the study of portfolios by using a variety of alternative methods, e.g., neural networks [2–4], genetic algorithms [5, 6], simulated annealing [7], Random Matrix Theory (RMT) filtering [8, 9], and hierarchical clustering [10–13]. Among them, hierarchical clustering is one of the most efficient methods for the selection of a basket of stocks for optimal portfolios. In fact, the selection of a set of stocks is a pre-requisite for the Markowitz theory [14, 15], since it is dedicated to the investment proportion of a limited number of selected stocks.

By using the hierarchical clustering method, the correlations between shares are revealed by the topological structure of the constructed stock network [16–25], and can be further applied to portfolio optimization. A minimum spanning tree (MST) [26] description of the correlations between stocks has shown that the stocks included in the minimum risk portfolio (the optimal Markowitz portfolio) tend to lie on the outskirts of the asset network [10, 27]. By extracting the dependency structure of financial equities using both MST and planar maximally filtered graphs (PMFG) methods, it has been found that the portfolios from a selection of peripheral stocks have lower risk and better returns than the portfolios from a selection of central stocks, the centrality/peripherality of which are measured by indices such as degree, betweenness centrality, eccentricity, closeness and eigenvector centrality [11]. A K-means clustering algorithm and its extension C-means clustering algorithm are applied to the classification of stocks. The stocks selected from these classified groups are used for building portfolios, which perform better than the benchmark index [12]. Similar works have been done in India, Taiwan and China stock markets using the same K-means cluster analysis [13, 14]. Furthermore, clusters or communities detected from network graphs also provide useful information for correlations among stocks [19, 22, 28, 29], which has also been used in the stock selection of portfolios [30–32]. This type of method is analogous to the clustering algorithm for their similarities in stock selections from clusters or communities partitioned by specific approaches.

Evidence from recent studies suggests that the topological structures of stock networks are evolving over time and changes markedly during financial crises [22, 28, 29, 33–38]. Therefore, a fixed set of stocks is not a wise choice for the portfolio selection under different market conditions. One possible way of solving this problem is to identify the stock clusters based on the network graphs in different time periods (moving windows) and perform the portfolio selection from the identified clusters in each period [11]. An alternative way is to use the dynamic conditional correlations (DCC) method to estimate time-varying correlations among stock returns based on the Markowitz framework [39–41]. A new estimator called detrended cross-correlation coefficients (DCCA) is also used to describe the correlations between nonlinear dynamic series [42–45]. Other works refer to evolutionary algorithms include [14, 46].

The main motivation of this paper is to propose a new dynamic portfolio strategy based on the time-varying structures of the financial filtered networks in the Chinese stock markets. A moving window with size  $\delta t$  is used to study the variance of stock networks over time  $t$ . We choose the MST method to filter out the network graph in each window for its validity and simplicity, which is generated by connecting the nodes with the most important correlations. The portfolio selection is determined by the network structure in the previous window (selection horizon), which is picked from a selection of peripheral stocks, most diverse

corresponding to the Markowitz portfolio with minimum variance [11], and central stocks, highly correlated and synchronous in price movements. The selected portfolios are subsequently used for investment in the following investment horizon.

The underlying market conditions are further considered in our dynamic portfolio strategy, which comprises the investment strategy together with the portfolio selection. A recent study has verified that accurate price and volatility predictions can be used as a basis for the particular trading strategy adopted for the portfolio [31]. In our study here, we presume that the optimal portfolio can change under different market conditions, and portfolio investments are implemented based on both historical price changes and price predictions in the future. For simplicity, three market conditions: drawup and drawdown trends of the daily price and a relatively stable status in between will be used in the selection and investment horizons. A variety of selected portfolios will be compared under different combinations of market conditions in the two horizons, and the optimal portfolio with the largest profit will be identified under each market condition. To further testify the efficiency of our dynamic portfolio strategy, we will carry out a training process using the first half of the sample data, and use the optimal portfolio obtained from training to make investments using the remaining half of the sample data. As will be shown below, our optimal portfolios outperform the benchmark stock index on average.

## Data and Methods

Our daily data include  $181 \times 2$  stocks listed on Shanghai and Shenzhen A-Share markets, which have the largest volumes in two major stock exchanges in mainland China, over the period of 15 years from January 1, 2000 to December 31, 2014. To ensure the continuity and integrity of the data, the stocks selected in our study are most actively traded stocks throughout the sample period. For this purpose, we filter out those stocks which were once suspended from the market for more than 46 trading days, about 1% of a total of 3,627 trading days. This filtering process yields the sample data including 181 A-Share stocks and 643,404 and 639,607 daily records in total for the two markets respectively. The return series of a certain stock  $i$  is computed as  $r_i(t) = \ln P_i(t) - \ln P_i(t-1)$ , where  $P_i(t)$  is the closing price for stock  $i$  on the  $t$ -th day. The price returns for  $181 \times 2$  stocks are calculated, and the effects of corporate actions are eliminated, for instance the cash dividend, the bonus share, and the rights issue.

On a certain day  $t$ , a correlation matrix is calculated by using the Pearson correlation coefficient estimator on the returns series in the window  $\{t - \delta t + 1, \dots, t\}$ , and the stock network constructed by the MST method, see details in network construction in Methods. Ten categories of portfolios are selected respectively from a set of 10% of most peripheral and central stocks in the MST graph, the centrality/peripherality of which are measured by five parameters capturing network topology: degree, betweenness centrality, distance on degree criterion, distance on correlation criterion and distance on distance criterion. For more details, see portfolio selection in Methods. The selected portfolios are used for investment in the following horizon  $\{t + 1, \dots, t + \Delta t\}$ , with an equal weight for each selected stock following [47], in which 1/N portfolio strategy is proven to be more efficient than the mean-variance model. The investment returns of chosen portfolios are calculated under nine combinations of market conditions in the selection and investment horizons. The market conditions include drawup (U), drawdown (D) and stable (S) status identified by trading day criterion, amplitude criterion, "OR" criterion, and "AND" criterion. Please refer to the descriptions on the identification of market conditions in Methods. Since the results obtained from the three criteria are quantitatively similar, we will only present our results based on the trading day criterion here. One then moves to  $t + \varphi$ . The same portfolio strategy is adopted by selecting the portfolio in the

window (horizon)  $\{t + \varphi - \delta t + 1, \dots, t + \varphi\}$ , and then using the selected portfolio for investment in the horizon  $\{t + \varphi + 1, \dots, t + \varphi + \Delta t\}$ . The investment returns of 10 categories of selected portfolios are calculated under nine combinations of market conditions in the two horizons, and the optimal portfolio is obtained by evaluating their average performances for different moving windows.

A suitable choice of  $\delta t$  and  $\varphi$  will indeed help the network capture the information of the original data as much as possible. The larger  $\delta t$  is and the smaller  $\varphi$  is, the more stable the network structure is, and the more the market information is filtered out. On the contrary, the network structure is more volatile and unauthentic [48, 49], though the temporal fluctuations can be easily noticed. Many studies have revealed that in order to ensure stocks to have enough number of trading days to be statistically significant,  $\delta t$  should be larger than the number of sample stocks  $N = 181$  [50, 51]. By careful observations and precise calculations, we choose  $\delta t = 10$  months ( $\approx 200$  days) and  $\varphi = 1$  month ( $\approx 20$  days), thus having 161 daily points to be used for portfolio investments in total. The determination of the optimal values of the parameter  $\varphi$  and the size of investment horizon  $\Delta t$  will be described in detail in network construction and determination of investment horizon in Methods respectively.

### Network construction based on MST method

Denote  $r_i(t)$  and  $r_j(t)$  as the logarithmic returns of stocks  $i$  and  $j$ , the Pearson correlation coefficient between their return series is given by

$$\rho(i, j) = \frac{E[r_i r_j] - E[r_i]E[r_j]}{\sqrt{(E[r_i^2] - E[r_i]^2)(E[r_j^2] - E[r_j]^2)}}, \tag{1}$$

where  $E[\cdot]$  denotes the mathematical expectation of the sequence over time  $t$ . Before the construction of the MST graph, the correlation coefficient is converted into the distance between stocks  $i$  and  $j$  by the following equation

$$d(i, j) = \sqrt{2(1 - \rho(i, j))}, \tag{2}$$

The distance  $d(i, j)$  ranges from 0 to 2, and a small distance corresponds to a large correlation coefficient. For the 181 sample stock datasets in the Shanghai and Shenzhen A-Share markets, a distance matrix with  $181 \times 181$  elements is obtained for each market respectively. The estimation of the correlation matrix has unavoidably associated with a statistical uncertainty, due to the finite length of the return series as well as noise.

We choose the MST method to filter out the network graph in each window so as to eliminate the redundancy and noise while maintaining significant links. By constructing the minimum spanning tree, we effectively reduce the information space from  $n(n - 1)/2$  correlation coefficients to  $n - 1$  tree edges. In other words, the amount of information is being compressed dramatically. The procedure to build the MST network can be carried out as follows: First, arrange the distances between all pairs of stocks in an ascending order. Second, start by matching the nearest nodes. Continue matching according to the ordered list if and only if the graph obtained after the matching is still a tree. Edges maximizing the sum of the correlations over the connections in the tree are more likely to stay by this method. Research works such as references [16, 17, 52–54] have used the MST model to filter networks. Given the data used in our study, we choose Prim algorithm to build our network.

There exists a close relationship between correlation coefficient matrices and MST distance matrices. To investigate their relationship, we calculate the Pearson linear correlation coefficients between the value of the mean, variance, skewness and kurtosis of the elements in both

**Table 1. Relationship between correlation coefficient matrices and MST distance matrices.**

	mean	variance	skewness	kurtosis
Shanghai A-Share market	-0.9849	0.7971	-0.7302	0.2813
Shenzhen A-Share market	-0.9847	0.8011	-0.8624	0.2584

Pearson linear correlation coefficients between four pairs of variables. The four variables include the value of mean, variance, skewness and kurtosis of elements in correlation coefficient matrices and distance matrices for the two markets.

doi:10.1371/journal.pone.0169299.t001

matrices. Generally speaking, the mean value of the elements in the two matrices are anti-correlated, and there exist similar features for skewness. Variance, together with kurtosis of the elements in the two matrices are positively correlated. These characteristics can be expected in view of how distances are constructed from the correlation coefficients. To confirm this, we provide the Pearson linear correlation coefficients of these variables in Table 1. We find that all variables of the elements in the two matrices are strongly correlated except for kurtosis, showing that our MST network contains most of the market information. In general, with a suitable choice of  $\varphi$ , the relationship between correlation coefficient matrices and MST distance matrices will be enhanced. From a series of tests, we find that  $\varphi = 1$  month turns out to be the optimal choice.

### Portfolio selection based on topological parameters

Five parameters are used to measure the centrality and peripherality of nodes in portfolio selection: (I) degree, (II) betweenness centrality, (III) distance on degree criterion, (IV) distance on correlation criterion and (V) distance on distance criterion. We here present a brief introduction of these parameters.

(I) Degree  $K$ , the number of neighbor nodes connected to a node. The larger the  $K$  is, the more the edges that are associated with this node.

(II) Betweenness centrality  $C$ , reflecting the contribution of a node to the connectivity of the network. Denote  $V$  as the set of nodes in the network. For nodes  $i$  and  $j$ ,  $C$  of a node  $k$  can be calculated as

$$C = \sum_{i,j \in V} \frac{\sigma_{ij}(V)}{\sigma_{ij}}, \tag{3}$$

where  $\sigma_{ij}$  is the number of shortest routes from node  $i$  to node  $j$ ,  $\sigma_{ij}(V)$  is a subsector of  $\sigma_{ij}$  whose routes pass through this node  $k$ .

Distance refers to the shortest length from a node to the central node of the network. Here, three types of definitions of a central node are introduced to reduce the error caused by a single method. Therefore three types of distances are described here.

(III) Distance on degree criterion  $D_{degree}$ , a central node is the node that has the largest degree.

(IV) Distance on correlation criterion  $D_{correlation}$ , a central node is the node with the highest value of the sum of correlation coefficients with its neighbors;

(V) Distance on distance criterion  $D_{distance}$ , a central node is the node that gives the smallest value for the mean distance.

We use the parameters defined above to select the portfolios. Nodes with the largest 10% of degree or betweenness centrality are chosen to be in the central portfolio, and nodes whose degree equals to 1 or betweenness centrality equals to 0 are chosen to be in the peripheral portfolio. Similarly, we define the nodes rank in the top 10% by distance as the stocks of the

peripheral portfolios, and the bottom 10% as the stocks of the central portfolios. The difference in the definitions results from a simple reason: In an MST network, the number of peripheral nodes (i.e., leaf nodes of a network) whose degree equals to 1 and betweenness centrality equals to 0, is much larger than 10% of the total nodes. We need to mention that it makes no difference to our results if we select randomly from these peripheral nodes so as to equal to the number in each portfolio.

The central portfolios and the peripheral portfolios represent two opposite sides of correlation and agglomeration. Generally speaking, central stocks play a vital role in the market and impose strong influence on other stocks, whereas the correlations between peripheral stocks are weak and contain more noise than central stocks. We have learned in our study that the two kinds of portfolios have their own features under different market conditions.

### Determination of investment horizon

In this subsection, we will discuss the optimal choice of the length of investment horizons  $\Delta t$ . In general, the length of the investment horizons cannot be too long, otherwise the topological properties of the network will change and the selected central or peripheral portfolios will change accordingly. On the other hand, the length of the investment horizons cannot be too short, or the returns will be greatly influenced by market noises or exogenous events. Here, we compare the profits gained in different time horizons, namely 1 month, 5 months, 10 months and 15 months.

Sharpe ratio, defined as the ratio of the expected value of the excess returns to its standard deviation [55], is widely used to evaluate the performance of portfolios in practice [56, 57]. Table 2 shows Sharpe ratios of portfolios with different investment horizons, in which there is no particular length that could maximize Sharpe ratio in all circumstances. What is more, a one-way ANOVA is used to test the equality of excess returns gained in different lengths of investment horizons. In Table 2, the  $p$ -values for different horizons are all insignificant, which indicates that the choice of the horizon length does not affect the portfolio returns. In order to match the length of selection horizons and investment horizons, and to facilitate our identification of the market conditions, we choose  $\Delta t = 10$  months for portfolio investments.

**Table 2. Comparison between excess returns for different lengths of investment horizons.**

		Shanghai A-Share market					Shenzhen A-Share market				
		1 month	5 month	10 month	15 month	$p$ -value	1 month	5 month	10 month	15 month	$p$ -value
$K$	peripheral	-0.0113	0.0024	-0.0005	-0.0023	0.9968	-0.0084	0.0025	-0.0007	-0.0028	0.9964
	central	0.0455	0.0112	0.0043	0.0002	0.9652	0.0444	0.0120	0.0061	0.0030	0.9892
$C$	peripheral	-0.0113	0.0024	-0.0005	-0.0023	0.9968	-0.0084	0.0025	-0.0007	-0.0028	0.9964
	central	0.0387	0.0058	0.0016	-0.0015	0.9635	0.0281	0.0074	0.0028	0.0002	0.9915
$D_{degree}$	peripheral	-0.0379	-0.0488	-0.0724	-0.1228	0.9978	-0.0281	-0.0048	-0.0054	-0.0072	0.9954
	central	0.0349	0.0386	0.0446	0.0367	0.9941	0.0389	0.0155	0.0069	0.0033	0.9923
$D_{correlation}$	peripheral	-0.0317	-0.0109	-0.0079	-0.0084	0.9991	-0.0385	-0.0043	-0.0052	-0.0067	0.9911
	central	0.0291	0.0061	0.0041	0.0016	0.9958	0.0450	0.0121	0.0067	0.0033	0.9912
$D_{distance}$	peripheral	-0.0351	-0.0091	-0.0067	-0.0079	0.9979	-0.0356	-0.0024	-0.0048	-0.0068	0.9876
	central	0.0195	0.0075	0.0038	0.0010	0.9984	0.0268	0.0107	0.0052	0.0020	0.9966

Sharpe ratio of portfolios for different lengths of investment horizons and ANOVA test of portfolios' excess returns. Portfolios listed include central and peripheral portfolios selected with respect to  $K$ ,  $C$ ,  $D_{degree}$ ,  $D_{correlation}$  and  $D_{distance}$  for the Shanghai and Shenzhen A-Share markets.

doi:10.1371/journal.pone.0169299.t002

### Identification of market conditions

Market conditions, which describe the general trend of the market index over a specific horizon, are measured by using four criteria: (I) trading day criterion, (II) amplitude criterion, (III) “OR” criterion, and (IV) “AND” criterion. The market indices corresponding to the Shanghai A-Share market and the Shenzhen A-Share market are the Shanghai A-Share Index and the Shenzhen A-Share Index respectively.

(I) Trading day criterion. The ratio  $r_d$  of the number of days with rising index to the total number of trading days in a specific time window is given by

$$r_d = \frac{N_i^+}{N_i}, \tag{4}$$

where  $N_i^+$  is the number of days in which the closing price is larger than that of the previous day and  $N_i$  is the total number of trading days in the  $i$ -th time window. The ratio  $r_d$  ranges from 0 to 1. A large value of  $r_d$  represents a drawup condition while a small value of  $r_d$  represents a drawdown condition. With the thresholds  $\theta_+$  and  $\theta_-$ , we identify a drawup condition if  $r_d > \theta_+$ , a drawdown condition if  $r_d < \theta_-$ , and a stable condition if  $\theta_- \leq r_d \leq \theta_+$ .

(II) Amplitude criterion. The ratio  $r_f$  of the sum of the amplitudes of the trading days with rising index to the sum of the amplitudes of the total trading days in a specific time window is given by

$$r_f = \frac{\sum_{t \in T_i^+} |P(t) - P(t-1)|}{\sum_{t \in T_i} |P(t) - P(t-1)|}, \tag{5}$$

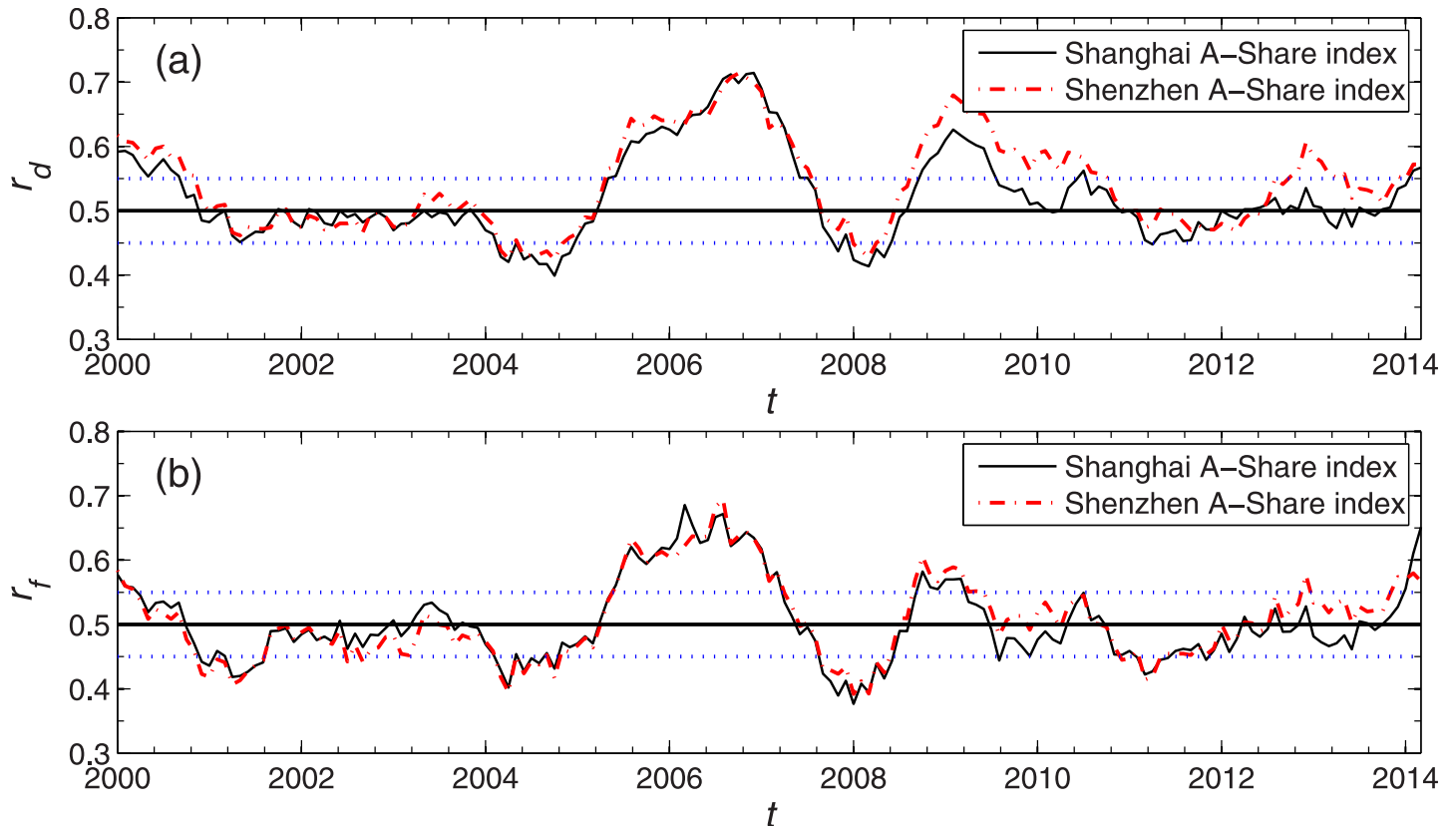
where  $T_i^+$  is the set of days in which the closing price is larger than that of the previous day,  $T_i$  is the set of all the trading days in the  $i$ -th time window, and  $P(t)$  is the closing price on the  $t$ -th day. Similarly, with the thresholds  $\theta_+$  and  $\theta_-$ , we identify a drawup condition if  $r_f > \theta_+$ , a drawdown condition if  $r_f < \theta_-$ , and a stable condition if  $\theta_- \leq r_f \leq \theta_+$ .

(III) “OR” criterion. We identify a drawup condition if  $r_d > \theta_+$  or  $r_f > \theta_+$ , and a drawdown condition if  $r_d < \theta_-$  or  $r_f < \theta_-$ . A stable condition is identified if  $\theta_- \leq r_d \leq \theta_+$  and  $\theta_- \leq r_f \leq \theta_+$ . Situations like  $r_d > \theta_+$  and  $r_f < \theta_-$ , or  $r_f > \theta_+$  and  $r_d < \theta_-$  do not exist for the thresholds that we choose.

(IV) “AND” criterion. We identify a drawup condition if  $r_d > \theta_+$  and  $r_f > \theta_+$ , and a drawdown condition if  $r_d < \theta_-$  and  $r_f < \theta_-$ . A stable condition is identified if  $\theta_- \leq r_d \leq \theta_+$  or  $\theta_- \leq r_f \leq \theta_+$ . Situations like  $r_d > \theta_+$  and  $r_f < \theta_-$ , or  $r_f > \theta_+$  and  $r_d < \theta_-$  do not exist for the thresholds that we choose.

The ratios  $r_d$  and  $r_f$  over time are shown in Fig 1, where each point on the curve is measured by using the market indices in a 10-month horizon following this point. In the figure, the patterns of  $r_d$  and  $r_f$  show some differences, which cause slight distinctions in the identification of market conditions. In our study, we choose  $\theta_+ = 0.55$ ,  $\theta_- = 0.45$  as the thresholds. Theoretically, other choices of thresholds will also work in our study. For a larger value of  $\theta_+$  and a smaller value of  $\theta_-$ , the number of samples of drawup or drawdown conditions is not statistically sufficient for the lack of data. For a smaller value of  $\theta_+$  and a larger value of  $\theta_-$ , the drawup and drawdown conditions cannot be distinctly identified. Other suitable choices of thresholds around  $\theta_+ = 0.55$ ,  $\theta_- = 0.45$  have also been studied and the results do not change significantly.

Market conditions identified based on the trading day criterion for the Shanghai and Shenzhen A-Share markets are shown in Fig 2. The upright triangle indicates a drawup horizon from the current time to 10 months later. Similarly the inverted triangle indicates a drawdown horizon and the cross symbol indicates a stable horizon. It appears that most of the drawup and drawdown conditions identified in our study are proper and sustainable in both markets.



**Fig 1. Ratios  $r_d$  and  $r_f$ .** Ratios  $r_d$  and  $r_f$  as a function of time  $t$ , together with the thresholds  $\theta_+ = 0.55$  and  $\theta_- = 0.45$  depicted by blue dotted lines.

doi:10.1371/journal.pone.0169299.g001

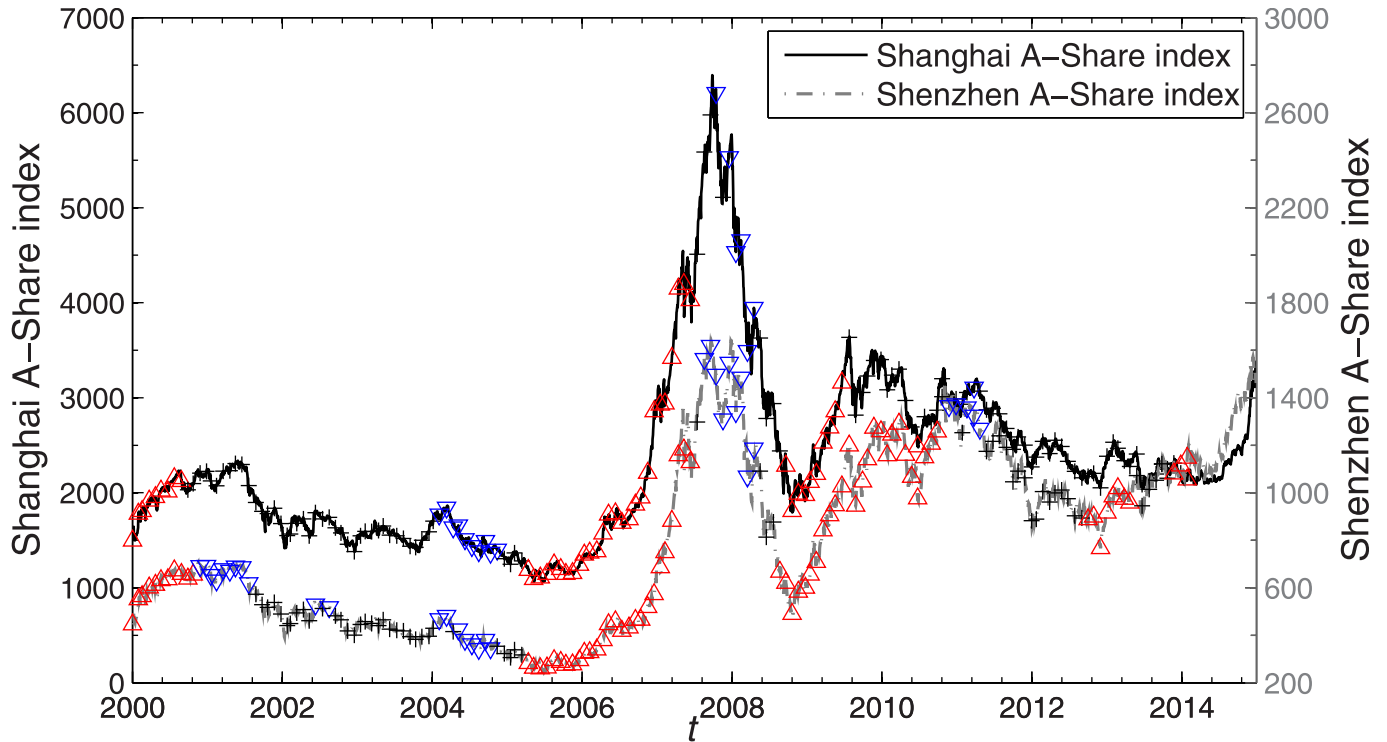
For each time window we get three possible market conditions. We thus get nine combinations of market conditions in the selection horizon and the following investment horizon: drawup in both selection horizon and investment horizon (UU), drawup in the selection horizon and stable in the investment horizon (US), drawup in the selection horizon and drawdown in the investment horizon (UD), stable in the selection horizon and drawup in the investment horizon (SU), stable in both selection and investment horizons (SS), stable in the selection horizon and drawdown in the investment horizon (SD), drawdown in the selection horizon and drawup in the investment horizon (DU), drawdown in the selection horizon and stable trend in the investment horizon (DS), and drawdown in both selection and investment horizons (DD).

## Results

### Evolution of network structures

Practically speaking, the network structure is evolving over time and changes remarkably during crises. Some evolutionary characteristics of the market can be found from the topological parameters of the networks. First, the average of the correlation coefficients among all stocks reflect the overall connections of the spanning tree, which is shown in Fig 3. To learn more about the details of the evolution of correlation coefficients in each window, the correlation coefficients ranging within one standard deviation of the average value are also shown in the figure. The average of the correlation coefficients rises sharply during the periods of market crashes in 2001 and 2008. As the market recovers, the average of the correlation coefficients declines correspondingly. This finding is consistent with the conclusions illustrated by



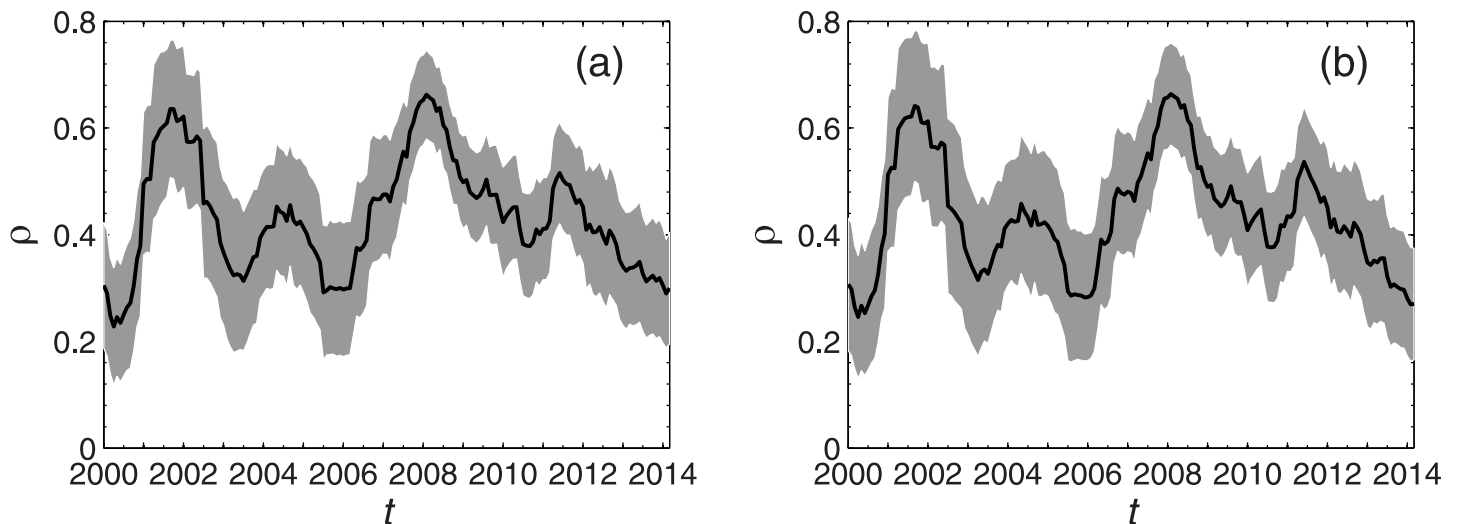


**Fig 2. Illustration of market condition.** Market conditions identified based on trading day criterion for the Shanghai and Shenzhen A-Share markets. Market conditions include drawup (upright triangle), drawdown (inverted triangle) and stable (cross) conditions.

doi:10.1371/journal.pone.0169299.g002

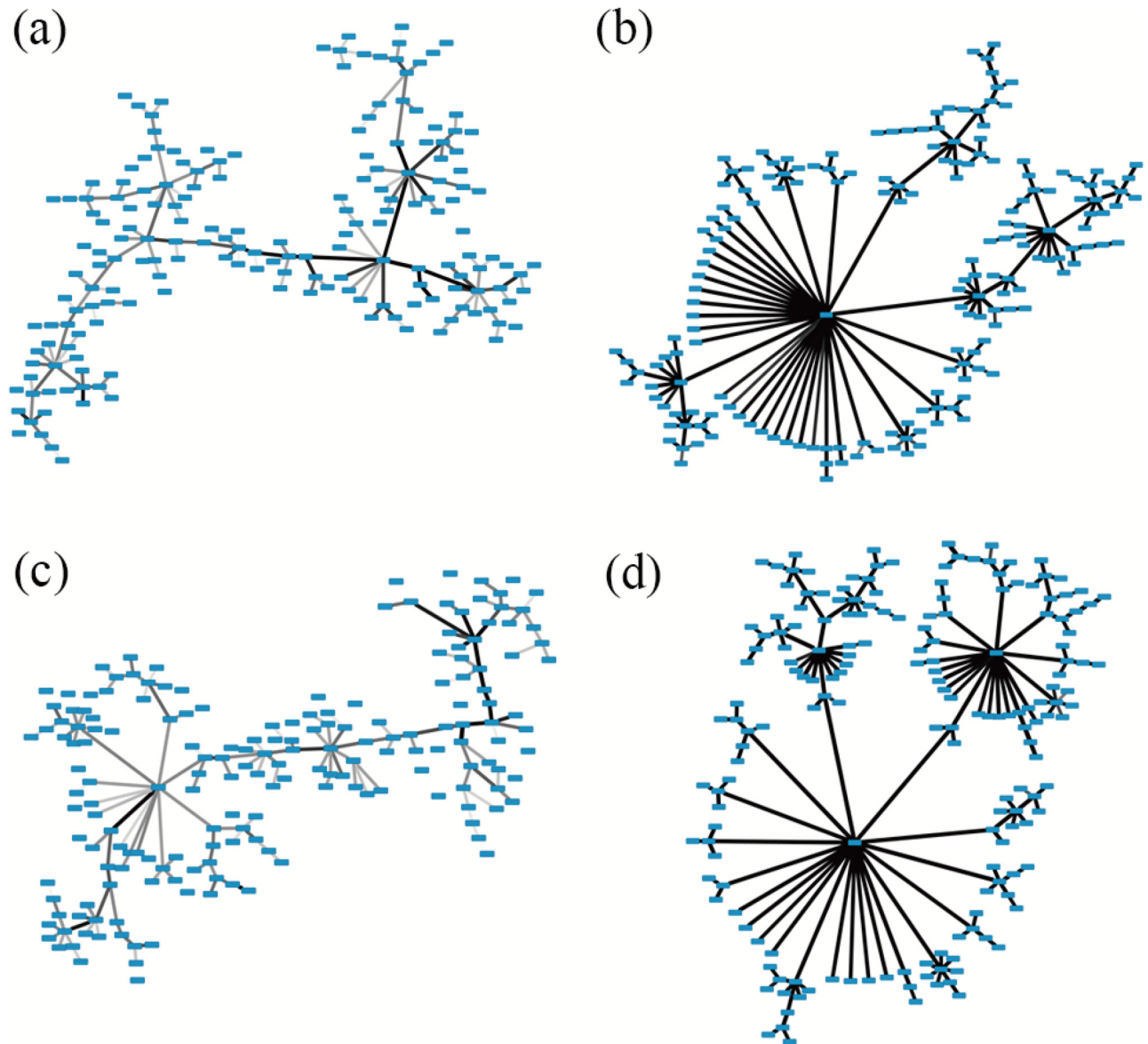
previous research works, indicating that the connection between stocks will be enhanced in a crisis [10, 34].

To better understand the variance of the network structure over time  $t$ , we choose two typical examples of MST networks beginning at January 1, 2000 and January 1, 2008, when stock



**Fig 3. Evolution of average correlation coefficient.** Evolution of the average of the correlation coefficients in the Shanghai A-Share market (a) and Shenzhen A-Share market (b). The average of the correlation coefficients are shown by the black solid lines in the center, and correlation coefficients ranging within one standard deviation of the average are shown in the grey area.

doi:10.1371/journal.pone.0169299.g003

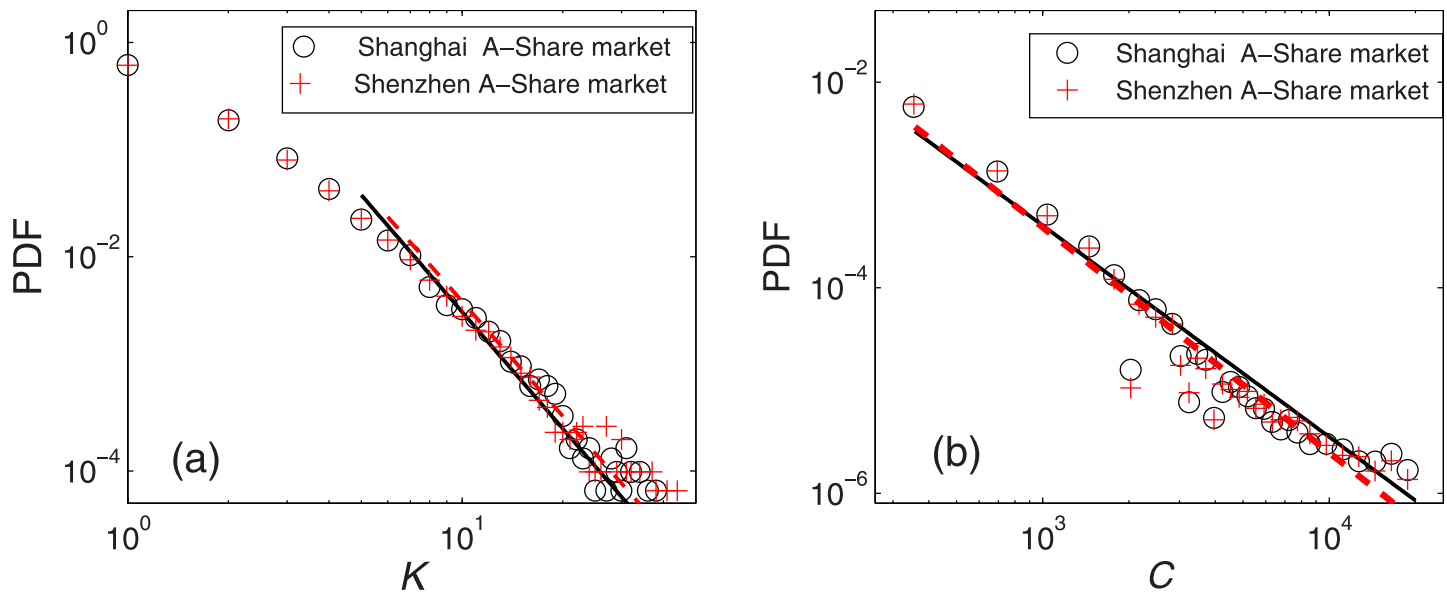


**Fig 4. Evolution of network structure.** (a) MST network for the Shanghai A-Share market in the period from January 1, 2000 to October 31, 2000; (b) MST network for the Shanghai A-Share market in the period from January 1, 2008 to October 31, 2008; (c) MST network for the Shenzhen A-Share market in the period from January 1, 2000 to October 31, 2000; (d) MST network for the Shenzhen A-Share market in the period from January 1, 2008 to October 31, 2008. Distances between stocks are indicated by line width: A thicker line represents a shorter distance while a thinner line represents a longer distance.

doi:10.1371/journal.pone.0169299.g004

prices are stable in the former window but volatile in the latter. Their network structures are shown in Fig 4. We can see that the distances between stocks are much smaller, and the corresponding network shrinks to a large extent during the 2008 market crisis.

Other parameters, namely degree, betweenness centrality, distance on degree criterion, distance on correlation criterion and distance on distance criterion, are more specific than the correlation coefficient, can therefore tell us more information about the stocks and networks. These parameters will be used for selecting portfolios in our later study.



**Fig 5. PDFs of  $K$  and  $C$ .** PDF of degree  $K$  and betweenness centrality  $C$  in the Shanghai A-Share market (a) and the Shenzhen A-Share market (b). The fitted lines with exponents estimated by a maximum likelihood estimation method proposed by Newman are for the Shanghai A-Share (black solid line) and Shenzhen A-Share markets (red dashed line).

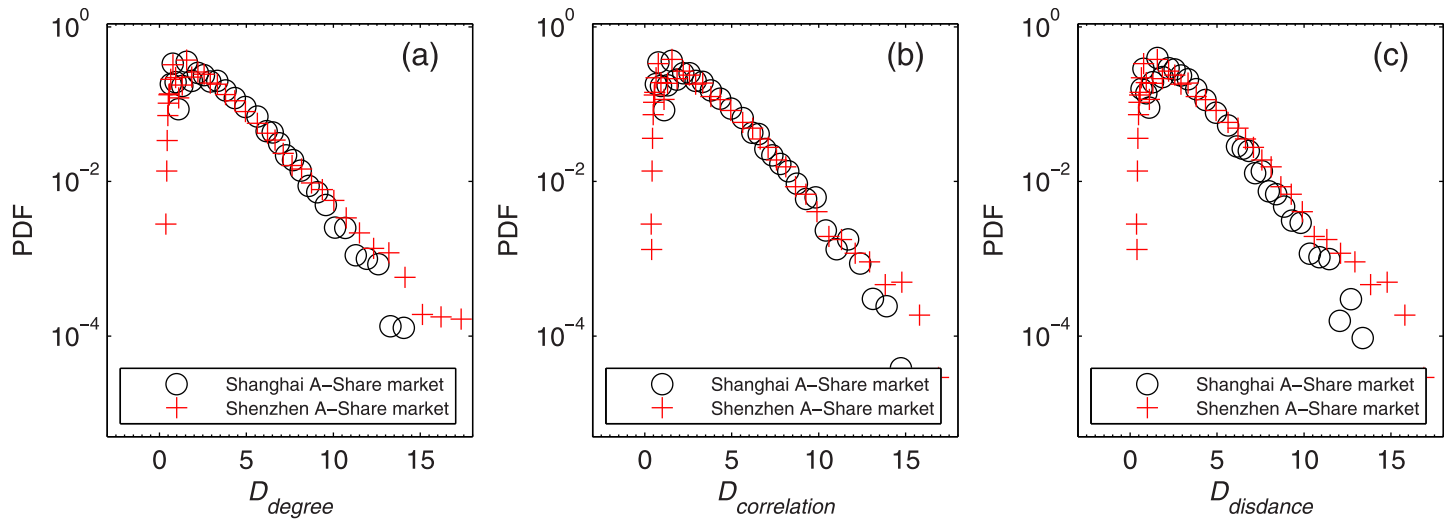
doi:10.1371/journal.pone.0169299.g005

The parameter degree  $K$  describes the relationship between a stock and its neighbors. Many studies have found that the degree in empirical networks follows a power-law distribution [48, 53, 58], where the distribution of  $K$  takes the form  $P(K) \sim K^{-\alpha}$ . Here, we observe a power-law behavior for the probability density function (PDF) of  $K$  in Fig 5. A maximum likelihood estimation method proposed by Newman is used to fit the distributions [59]. The exponents  $\alpha = 3.6202$  and  $3.5597$  are obtained for  $P(K)$  in the Shanghai and Shenzhen A-Share markets respectively. Another parameter, the betweenness centrality  $C$  reflects the contribution of a stock to the connectivity of the whole network, resembling the degree to some extent. The PDF of  $C$ , also shown in Fig 5, shows a power-law behavior. The exponents  $\alpha = 2.0562$ , and  $2.1927$  are estimated for  $P(C)$  in the Shanghai and Shenzhen A-Share markets respectively. Results in the figure show the scale free property of MST networks in the two markets, and also the intense connection between stocks existing in a small number of central stocks, the volatility of which might have a great impact on its neighbors.

Distance, which is the total length from a node to the central node of the network, has three categories according to the choice of central node in our study, namely distance on degree criterion  $D_{degree}$ , distance on correlation criterion  $D_{correlation}$  and distance on distance criterion  $D_{distance}$ . Their PDFs are shown in Fig 6. Although the three categories of distances differ from each other, their distributions as well as the stocks selected by these criteria share similar behavior. As can be seen from Fig 6, few stocks are very distant from the central node, while most stocks are at a relatively short distance from the central node. In addition, we note that the MST networks change over time, its central node therefore also changes.

### Comparison between portfolio strategies under different market conditions

In this subsection, we compare the returns of the central and peripheral portfolios to look for the optimal portfolio among these portfolios. The portfolios are selected by using five parameters, i.e.,  $K$ ,  $C$ ,  $D_{degree}$ ,  $D_{correlation}$  and  $D_{distance}$  in the selection horizon, and the returns of the



**Fig 6. PDF of distance.** (a) PDF of distance based on degree criterion  $D_{degree}$ , (b) PDF of distance based on correlation criterion  $D_{correlation}$  and (c) PDF of distance based on distance criterion  $D_{distance}$  in the Shanghai A-Share and Shenzhen A-Share markets.

doi:10.1371/journal.pone.0169299.g006

selected portfolios are calculated in the following investment horizon. The length of the investment horizon is set to be 10 months as discussed in Determination of investment horizon in Methods. In this paper, the selection horizon lasts from January 1, 2000 to February 31, 2014, and the investment horizon lasts from November 1, 2000 to December 31, 2014. We thus have a total of 161 daily points to be used for portfolio investments.

We calculate the returns of the central and peripheral portfolios, and use the returns of random portfolios as a benchmark. A random portfolio is defined as a randomly selected portfolio containing 10% of the total stocks. We first classify the samples of returns of selected portfolios and random portfolios into groups according to the nine combinations of market conditions identified using thresholds  $\theta_+ = 0.55$ ,  $\theta_- = 0.45$  based on the trading day criterion. For each combination of market conditions, we calculate the average return of each stock in the group of selected portfolios and the average return of each stock in the group of random portfolios. The difference between the average returns of selected portfolios and random portfolios is defined as the excess return. Furthermore, a one-way ANOVA is used to test the equality of the excess returns between central portfolios and peripheral portfolios under the same market condition. Null hypothesis of the one-way ANOVA, in which the excess returns of central portfolios and peripheral portfolios are equal, are tested at a specific significance level. If the null hypothesis is rejected, the excess returns of the two portfolios are significantly different. If the null hypothesis cannot be rejected, there is no significant difference between the excess returns of the central portfolios and the peripheral portfolios.

The results of the one-way ANOVA test and the excess returns of the central and the peripheral portfolios for the Shanghai and Shenzhen A-Share markets are listed in Tables 3 and 4 respectively. Notice that if the sample number is less than 11 under a specific combination of market conditions, we do not show the result of this combination for the lack of data. In addition, if there is no significant difference between the excess returns of the central portfolios and the peripheral portfolios, the results are also not shown. In Table 3, we can see from the  $p$ -value that the excess returns between the central and peripheral portfolios of the listed groups are all significantly different at the 10% level, and most of them are significantly different at the 5% level. By comparing the excess returns of the central and the peripheral portfolios under all listed combinations of market conditions, we find that the central portfolios are

**Table 3. Comparison between excess returns of central and peripheral portfolios for Shanghai A-Share market.**

Parameter	Market condition	Num	f-value	p-value	excess returns	
					central	peripheral
K	US	24	15.17	0.00**	<b>0.04±0.01</b>	0.00±0.00
	SS	74	6.75	0.01**	<b>0.02±0.01</b>	0.00±0.00
	SD	11	5.79	0.03**	<b>0.02±0.02</b>	0.00±0.00
	DU	11	3.87	0.06*	<b>0.03±0.02</b>	0.00±0.00
$D_{degree}$	US	24	5.69	0.02**	<b>0.02±0.01</b>	-0.03±0.02
	SU	12	4.27	0.05*	<b>0.02±0.02</b>	-0.04±0.02
	SS	74	21.39	0.00**	<b>0.02±0.01</b>	-0.03±0.01
	SD	11	3.57	0.07*	-0.02±0.02	<b>0.04±0.02</b>
$D_{correlation}$	DU	11	53.08	0.00**	<b>0.05±0.02</b>	-0.08±0.02
	US	24	5.33	0.03**	<b>0.02±0.01</b>	-0.03±0.02
	SU	12	4.27	0.05*	<b>0.02±0.02</b>	-0.04±0.02
	SS	74	20.82	0.00**	<b>0.02±0.01</b>	-0.03±0.01
$D_{distance}$	DU	11	52.77	0.00**	<b>0.05±0.02</b>	-0.08±0.02
	US	24	3.98	0.05*	<b>0.01±0.01</b>	-0.03±0.01
	SU	12	4.18	0.05*	<b>0.01±0.02</b>	-0.03±0.02
	SS	74	19.81	0.00**	<b>0.03±0.01</b>	-0.03±0.01
	SD	11	4.36	0.05*	-0.02±0.02	<b>0.04±0.02</b>
DU	11	56.84	0.00**	<b>0.05±0.02</b>	-0.08±0.02	

Excess returns of central and peripheral portfolios in the Shanghai A-Share market are compared based on one-way ANOVA. Excess returns of central and peripheral portfolios are selected with respect to five parameters, i.e.,  $K$ ,  $C$ ,  $D_{degree}$ ,  $D_{correlation}$  and  $D_{distance}$ , under different combinations of market conditions based on trading day criterion. The listed variables include the number of samples (Num),  $f$ -value and  $p$ -value of one-way ANOVA, excess returns of central and peripheral portfolios under each combination of market conditions. Results which are not significant or calculated with less than 11 samples are not shown. (\*indicates significance at 10% level, \*\*indicates significance at 5% level. Figures after ± indicate standard error).

doi:10.1371/journal.pone.0169299.t003

more profitable except for two cases. More specifically, when the market is stable in the investment horizon, and whether the market is stable or drawup in the selection horizon, the excess returns of central portfolios are significantly larger than those of peripheral portfolios. When the market is drawdown in the selection horizon and drawup in the investment horizon, or when the market is stable in the selection horizon and drawup in the investment horizon, the central portfolios outperform the peripheral portfolios when using  $K$ ,  $D_{degree}$ ,  $D_{correlation}$  and  $D_{distance}$  as parameters. Besides, when the market is stable in the selection horizon and drawdown in the investment horizon, the peripheral portfolios gain more than the central portfolios when using  $D_{degree}$  and  $D_{distance}$  as parameters. However, under the same market condition and using  $K$  as the parameter, the optimal portfolio is the central portfolio with a relatively small gap between the excess returns of the two portfolios. In Table 4 for the Shenzhen A-Share market, we find that the results generally coincide with those for the Shanghai A-Share market. The central portfolios outperform the peripheral portfolios under every combination of market conditions when there exist significant difference between their returns. What is more, the central portfolios selected by using  $K$ ,  $D_{degree}$ ,  $D_{correlation}$  and  $D_{distance}$  outperform the peripheral portfolios when the market goes through a rising trend in both selection and investment horizons.

We here give the PDF of the returns of individual stocks in the central and peripheral portfolios for the Shanghai and Shenzhen A-Share markets in Figs 7 and 8 respectively. Under most market conditions, the peak of the distribution for the central portfolios in Fig 7 is on the

**Table 4. Comparison between excess returns of central and peripheral portfolios for Shenzhen A-Share market.**

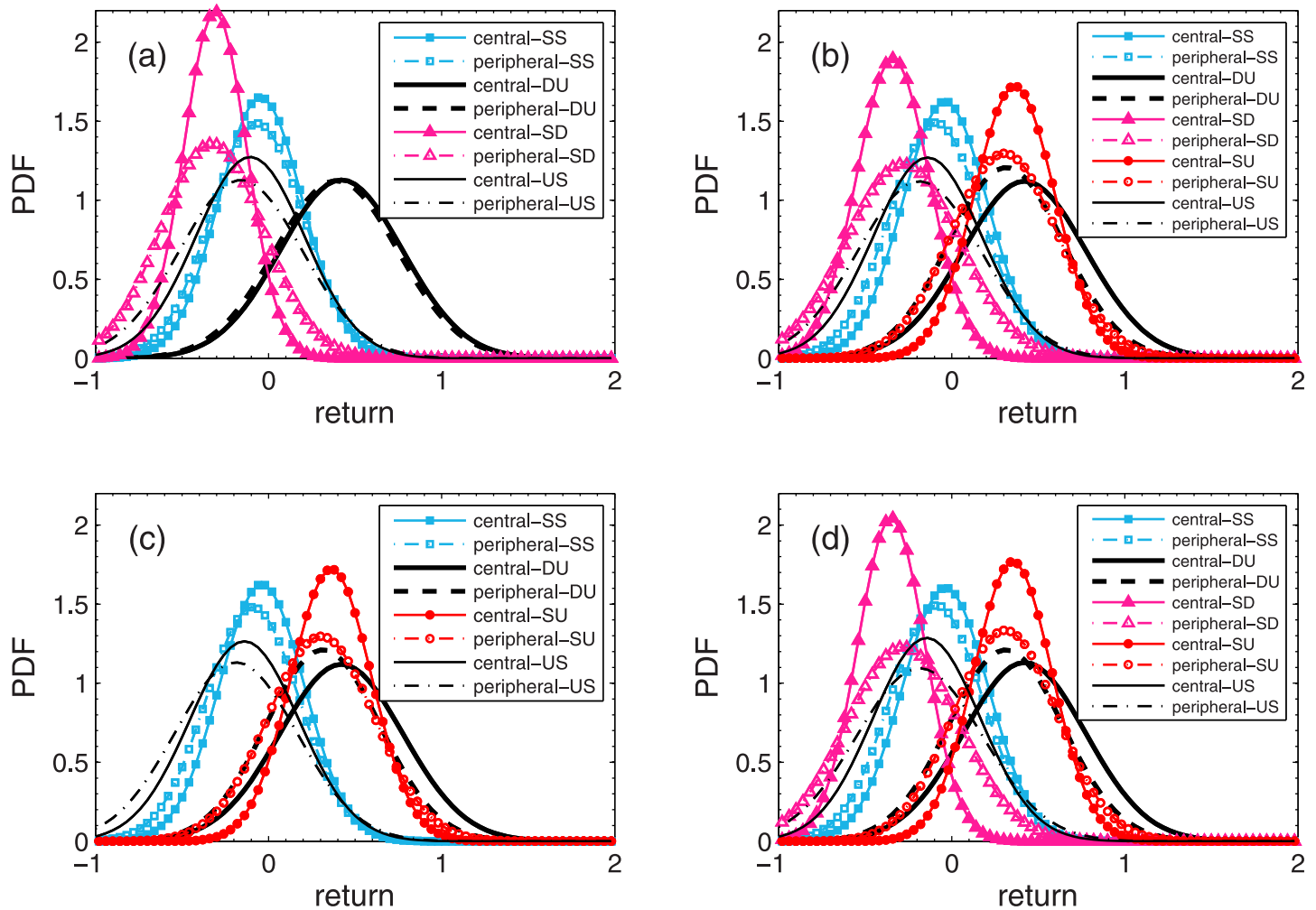
Parameter	Market condition	Num	f-value	p-value	excess returns	
					central	peripheral
K	UU	37	2.87	0.09*	<b>0.02±0.01</b>	0.00±0.00
	US	32	9.75	0.00**	<b>0.03±0.01</b>	-0.01±0.00
	SU	20	3.93	0.05*	<b>0.03±0.01</b>	0.00±0.00
	SS	50	19.83	0.00**	<b>0.02±0.01</b>	-0.01±0.00
C	SS	50	8.00	0.01**	<b>0.01±0.01</b>	-0.01±0.00
$D_{degree}$	UU	37	10.35	0.00**	<b>0.02±0.01</b>	-0.04±0.01
	SU	20	20.66	0.00**	<b>0.05±0.01</b>	-0.05±0.02
	SS	50	8.50	0.00**	<b>0.02±0.01</b>	-0.01±0.01
$D_{correlation}$	UU	37	3.06	0.08*	<b>0.00±0.01</b>	-0.03±0.01
	US	32	2.82	0.10*	<b>0.03±0.01</b>	-0.01±0.01
	SU	20	20.66	0.00**	<b>0.05±0.02</b>	-0.05±0.02
	SS	50	11.91	0.00**	<b>0.03±0.01</b>	-0.02±0.01
$D_{distance}$	UU	37	6.62	0.01**	<b>0.01±0.01</b>	-0.03±0.01
	SU	20	4.76	0.04**	<b>0.02±0.01</b>	-0.04±0.02
	SS	50	10.63	0.00**	<b>0.02±0.01</b>	-0.02±0.01

Excess returns of central and peripheral portfolios in the Shenzhen A-Share market are compared based on one-way ANOVA. Excess returns of central and peripheral portfolios are selected with respect to five parameters, i.e.,  $K$ ,  $C$ ,  $D_{degree}$ ,  $D_{correlation}$  and  $D_{distance}$ , under different combinations of market conditions based on trading day criterion. The listed variables include the number of samples (Num),  $f$ -value and  $p$ -value of one-way ANOVA, excess returns of central and peripheral portfolios under each combination of market conditions. Results which are not significant or calculated with less than 11 samples are not shown. Figures after ± indicate standard error. (\*indicates significance at 10% level, \*\*indicates significance at 5% level. Figures after ± indicate standard error).

doi:10.1371/journal.pone.0169299.t004

right side of the peak for the peripheral portfolios, indicating that the returns of the central portfolios are on the average larger than those of the peripheral portfolios. However, when portfolios are selected with respect to  $D_{degree}$  or  $D_{distance}$ , the peak of the distribution for the peripheral portfolios is on the right side of the peak for the central portfolios under market conditions of SD. These results are consistent with the results from Table 3. We can also see from Fig 7 that the returns of the central portfolios distribute at a relatively narrow range compared with the returns of the peripheral portfolios under most market conditions, showing the close connection among stocks in the central portfolios. Fig 8 shows the PDFs of the returns of individual stocks for the Shenzhen A-Share market. The peak of the distribution for the central portfolios is on the right side of the peak for the peripheral portfolios under all market conditions, indicating that the returns of the central portfolios are on the average larger than those of the peripheral portfolios. Meanwhile, all the returns of the central portfolios have narrow distributions when compared with the returns of the peripheral portfolios, showing the diversification of returns of stocks in the peripheral portfolios.

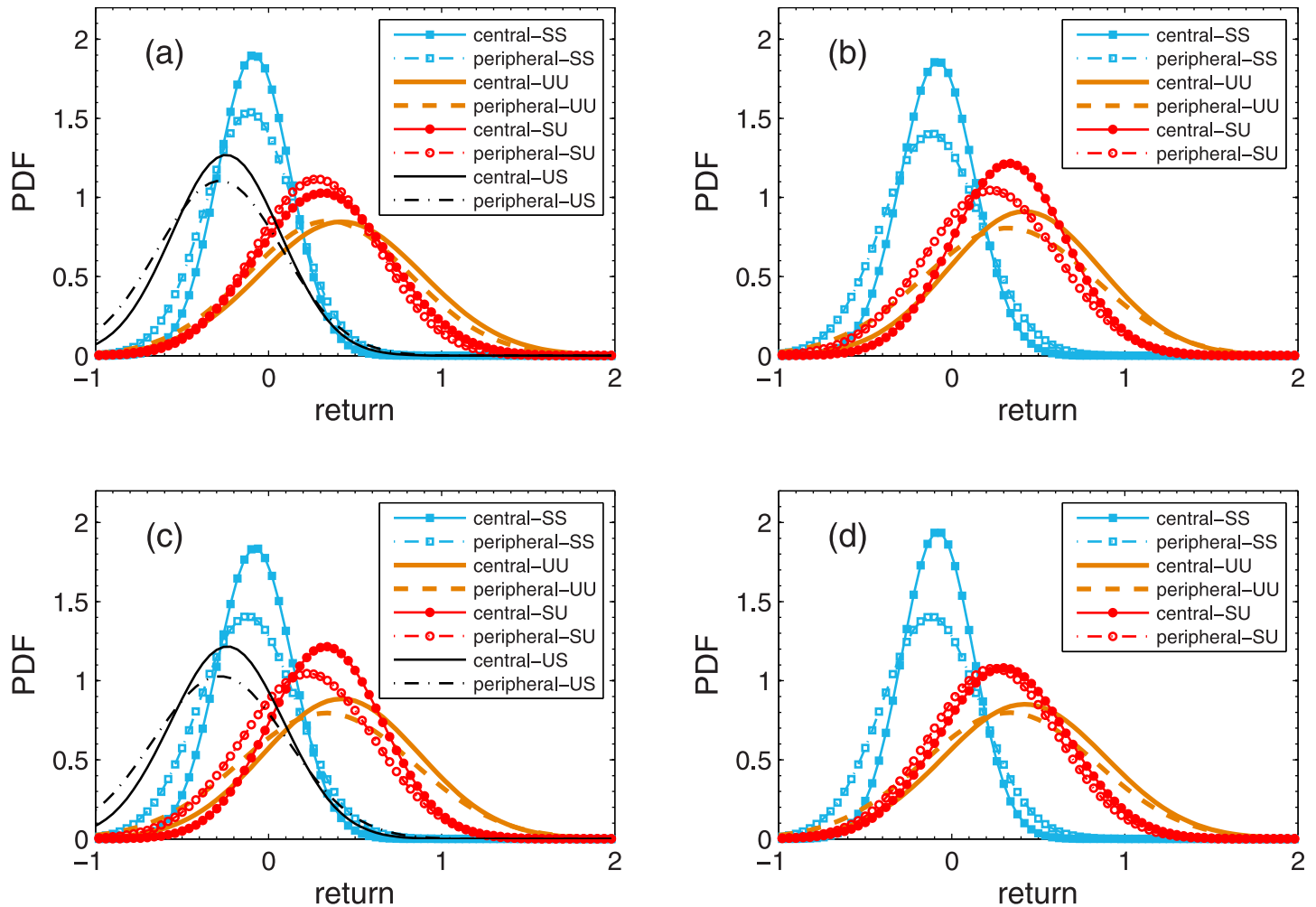
Similar tests and comparisons are carried out using the amplitude criterion, the “AND” criterion and the “OR” criterion in the identification of market conditions for both markets. Using the amplitude criterion and the “AND” criterion, the results are very similar to those using the trading day criterion. For the “OR” criterion, we find that if the market goes through a turnaround, falling in the selection horizon and rallying in the investment horizon, the central portfolios are proven to be more profitable. More comprehensively, extra efforts have been made for other values of thresholds in the identification of market conditions. Despite some statistical distinctions, the conclusion here shows great similarities with the above, which in turn confirm the reliability of our conclusion.



**Fig 7. PDF of individual stock returns for Shanghai A-Share Market.** PDF of individual stock returns under each combination of market conditions based on trading day criterion in the Shanghai A-Share Market are plotted. Stock portfolios are selected with respect to degree  $K$  (a), distance based on degree criterion  $D_{degree}$  (b), distance based on correlation criterion  $D_{correlation}$  (c), and distance based on distance criterion  $D_{distance}$  (d).

doi:10.1371/journal.pone.0169299.g007

We summarize and interpret our results in three aspects. First, if the market rises in the investment horizon, the central portfolios should be the best choice. Specifically, if the market has a drawup trend in both the selection and investment horizons, stocks in central portfolios will more likely rise due to their collective movements in rising, while stocks in peripheral portfolios may be too diversified to make profits. If the market rises in the investment horizon after declining, stocks in central portfolios will more likely suffer losses in the selection horizon, and will more likely rise in the investment horizon after hitting rock bottom. Second, if the market is stable in the investment horizon and has just gone through a drawup or stable trend in the selection horizon, central portfolios are preferred. Since stocks in central portfolios are closely related, their prices move in similar behaviors. After the drawup or stable trend in the selection horizon, stocks in central portfolios tend to maintain the drawup or stable trend in the investment horizon. Returns of stocks in peripheral portfolios are too diversified in the stable investment horizon, some negative returns of individual stocks will more likely be in the portfolio. Finally, if the market falls after a period of stable fluctuations, the peripheral



**Fig 8. PDF of individual stock returns for Shenzhen A-Share Market.** PDF of individual stock returns under each combination of market conditions based on trading day criterion in Shenzhen A-Share Market are plotted. Stock portfolios are selected with respect to degree  $K$  (a), distance based on degree criterion  $D_{degree}$  (b), distance based on correlation criterion  $D_{correlation}$  (c), and distance based on distance criterion  $D_{distance}$  (d).

doi:10.1371/journal.pone.0169299.g008

portfolios are preferred to avoid risks. The diverse characteristic of the peripheral portfolios is a good way to reduce risk and secure capital.

### Empirical test of optimal portfolio strategy

We have compared the performances of the central and the peripheral portfolios under different combinations of market conditions. We will attempt to choose the optimal portfolio strategy and apply the strategy to make real investment based on an empirical test. We use the data from 2000 to 2010 to select the optimal portfolio under each specific combination of market conditions through a training process, and use the selected optimal strategy according to the current market condition to make investment based on the data from 2010 to 2014. The test is performed specifically as follows.

I. Training to find the optimal portfolio strategy. The training process is carried out by using the methods mentioned in the last section to find the optimal portfolios under specific market conditions. A series of MST networks in the selection horizons are constructed to build central and peripheral portfolios, which are then used for investment in the following



horizon. Excess returns of the central and peripheral portfolios in the investment horizons are calculated and tested using a one-way ANOVA under each market condition. Portfolios with significantly higher excess returns under different market conditions will comprise the optimal portfolio strategy.

II. Applying the optimal strategy to investment. Before investment, market conditions in the investment horizon need to be predicted. The trend of the stock market is found to be affected by many factors such as policy [60, 61] and economic environment [62, 63]. The effect of these factors is quite evident in the Chinese stock market [64–66], which makes it possible to use these factors to project the market trend. Since we mainly concentrate on building portfolios and strategies other than trend prediction, we identify the market conditions in the investment horizon using the empirical data. In other words, our strategy performs well when the market condition in the investment horizon has a clear trend. Based on the identified combination of market conditions, the optimal portfolio of the optimal strategy chosen in step I is selected and used for further investment. If the combination of market conditions do not appear, investment will not be made. The length of the selection and investment horizons are 10 months, same as in the previous section.

We calculate the excess returns of our strategy, which is the difference between the average return of the stocks in the optimal portfolio strategy and the random strategy. The random strategy comprises random portfolios, defined as portfolios containing 10% of the total stocks that are randomly selected, under different combinations of market conditions. Since the optimal strategy changes as we use five parameters to select portfolios and four criteria to identify market conditions, the excess returns which are shown in Table 5 differ correspondingly. It can be seen from the table that in most cases higher profits can be obtained by our strategies compared with the random strategy. Specifically, 65% of the returns of our strategies are larger than those of the random strategy in the Shanghai A-Share market, and the proportion is 70% in the Shenzhen A-Share market. Furthermore, when using  $K$  as the parameter to select central or peripheral portfolios, returns of our strategies are always higher than those of the random strategy but the portfolios selected by  $C$  rarely outperform the random portfolios. Since our

Table 5. Excess return of optimal portfolio strategy.

	Shanghai A-Share market				Shenzhen A-Share market			
	trading day criterion	amplitude criterion	“OR” criterion	“AND” criterion	trading day criterion	amplitude criterion	“OR” criterion	“AND” criterion
$K$	<b>0.0169</b>	<b>0.0065</b>	<b>0.0072</b>	<b>0.0156</b>	<b>0.0131</b>	<b>0.0113</b>	<b>0.0190</b>	<b>0.0238</b>
	$\pm 0.0104$	$\pm 0.0068$	$\pm 0.0045$	$\pm 0.0102$	$\pm 0.0119$	$\pm 0.0088$	$\pm 0.0072$	$\pm 0.0076$
$C$	-0.0016	-0.0003	<b>0.0129</b>	-0.0033	-0.0030	-0.0135	-0.0118	-0.0049
	$\pm 0.0096$	$\pm 0.0073$	$\pm 0.0072$	$\pm 0.0083$	$\pm 0.0090$	$\pm 0.0087$	$\pm 0.0117$	$\pm 0.0000$
$D_{degree}$	<b>0.0345</b>	<b>0.0025</b>	<b>0.0072</b>	-0.1290	<b>0.0303</b>	<b>0.0381</b>	-0.0093	<b>0.0529</b>
	$\pm 0.0106$	$\pm 0.0176$	$\pm 0.0160$	$\pm 0.0000$	$\pm 0.0150$	$\pm 0.0102$	$\pm 0.0208$	$\pm 0.0166$
$D_{correlation}$	<b>0.0385</b>	<b>0.0025</b>	<b>0.0072</b>	-0.1290	<b>0.0632</b>	<b>0.0451</b>	<b>0.0542</b>	<b>0.0214</b>
	$\pm 0.0113$	$\pm 0.0176$	$\pm 0.0160$	$\pm 0.0000$	$\pm 0.0120$	$\pm 0.0132$	$\pm 0.0195$	$\pm 0.0293$
$D_{distance}$	<b>0.0416</b>	-0.0005	<b>0.0045</b>	-0.1290	<b>0.0054</b>	-0.0110	<b>0.0286</b>	<b>0.0000</b>
	$\pm 0.0127$	$\pm 0.0153$	$\pm 0.0142$	$\pm 0.0000$	$\pm 0.0103$	$\pm 0.0098$	$\pm 0.0142$	$\pm 0.0089$

Excess returns listed are gained by the optimal portfolio strategy for the Shanghai and Shenzhen A-Share markets. The optimal strategy is chosen from all possible categories of portfolios selected with respect to five parameters, namely  $K$ ,  $C$ ,  $D_{degree}$ ,  $D_{correlation}$  and  $D_{distance}$ , based on four criteria, namely trading day criterion, amplitude criterion, “OR” criterion and “AND” criterion. Excess returns in boldface are positive which have returns larger than the random strategy. Figures after  $\pm$  indicate standard error.

doi:10.1371/journal.pone.0169299.t005

**Table 6. Optimal portfolios of the most profitable strategy.**

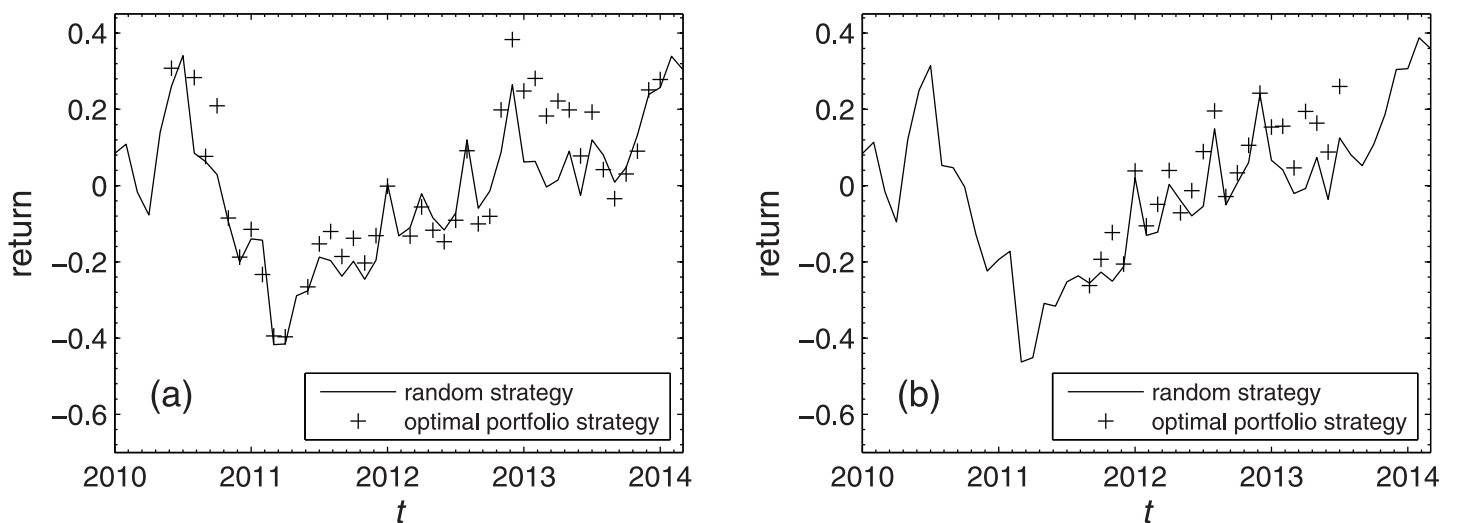
Market condition	Shanghai A-Share market				Shenzhen A-Share market	
	UD	SS	SD	DU	SU	SS
Optimal portfolio	central	central	peripheral	central	central	central

Specific optimal portfolios under particular combinations of market conditions comprise the strategy which has the highest returns. The most profitable strategy for the Shanghai A-Share market uses  $D_{distance}$  to select portfolios and identifies market conditions based on trading day criterion, and the most profitable strategy for the Shenzhen A-Share market uses  $D_{correlation}$  to select portfolios and identifies market conditions based on trading day criterion.

doi:10.1371/journal.pone.0169299.t006

strategy based on the “AND” criterion in the Shanghai A-Share market is rarely used, the excess returns under which are mostly negative. The most profitable strategy for the Shanghai A-share market uses  $D_{distance}$  as the parameter to select portfolios and identifies market conditions based on the trading day criterion, which has an excess return of 0.0416. For the Shenzhen A-Share market, the most profitable strategy uses  $D_{correlation}$  as the parameter to select portfolios and identifies market conditions based on the trading day criterion, which has an excess return of 0.0632.

Table 6 lists the optimal portfolios of the most profitable strategy under different combinations of market conditions for both markets. One can see that the central portfolios are chosen as the optimal portfolios under market conditions of UD, SS and DU, and the peripheral portfolios are chosen under market conditions of SD for the Shanghai A-share market. For the Shenzhen A-Share market, the central portfolios are chosen as the optimal portfolios under market conditions of SU and SS. The average returns of each individual stock in every investment horizon of the most profitable strategy and the random strategy are plotted in Fig 9, where the crosses are the average returns gained by the most profitable strategy in each investment horizon and the black solid line shows the average returns gained by the random strategy. In 65.85% of the investment horizons, the average returns gained by the most profitable strategy are larger than those gained by the random strategy in the Shanghai A-Share market, and the proportion is 91.30% in the Shenzhen A-Share market.



**Fig 9. Average return of the most profitable strategy.** Average returns of the most profitable strategy (cross) and random strategy (black solid line) for the Shanghai A-Share market (a) and the Shenzhen A-Share market (b). This strategy is described in Table 6.

doi:10.1371/journal.pone.0169299.g009

**Table 7. Optimal portfolios of the strategy with the largest probability of gaining more profits.**

Market condition	Shanghai A-Share market			Shenzhen A-Share market	
	UD	SS	DU	SU	SS
Optimal portfolio	central	central	central	central	central

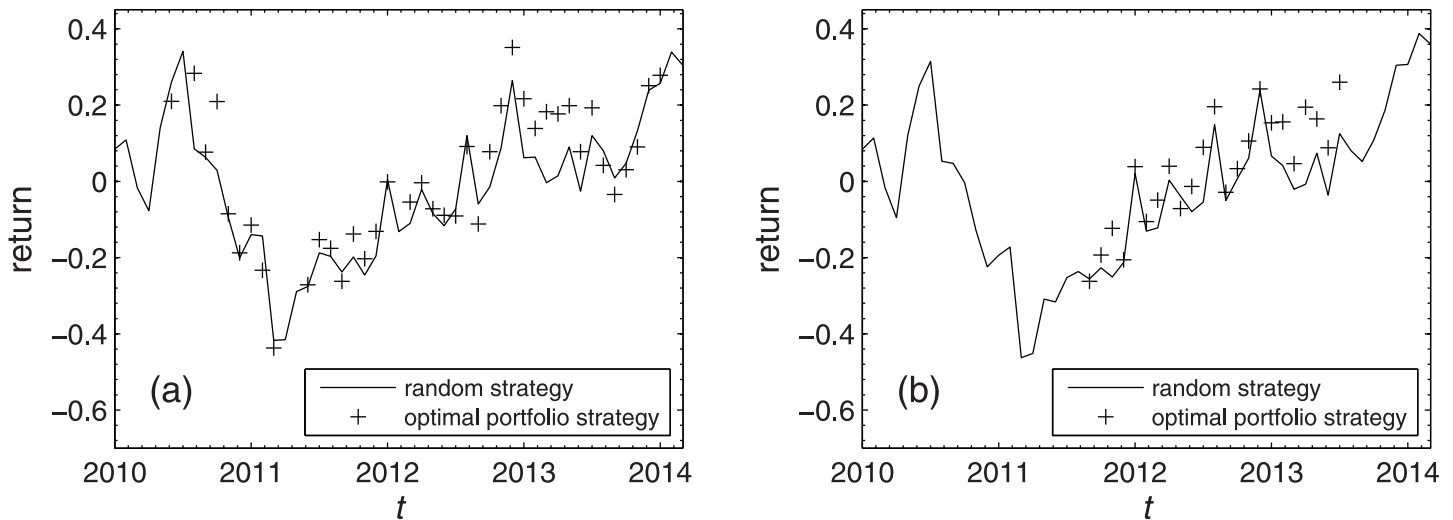
Specific optimal portfolios under particular combinations of market conditions comprise the strategy that has the largest probability of gaining more profits than the random strategy. This strategy uses  $D_{correlation}$  as the parameter to select portfolios and identifies market conditions based on trading day criterion for both markets.

doi:10.1371/journal.pone.0169299.t007

We further find the strategy that has the largest probability of gaining more profits than the random strategy, i.e., it outperforms the random strategy in most of investment horizons. This strategy uses  $D_{correlation}$  as the parameter to select portfolios and identifies market conditions based on the trading day criterion for both markets. The optimal portfolios under different combinations of market conditions are listed in Table 7. One can see that the central portfolios are chosen as the optimal portfolios under market conditions of UD, SS and DU for the Shanghai A-share market. For the Shenzhen A-Share market, this strategy is also the most profitable strategy with the central portfolios chosen as the optimal portfolios under market conditions of SU and SS. Similar to the above, the average returns of each individual stock in every investment horizon of this strategy and the random strategy are plotted in Fig 10. In 70% of the investment horizons, the average returns gained by our strategy are larger than the random strategy in the Shanghai A-Share market, and the proportion is 91.30% in the Shenzhen A-Share market.

### Discussion and conclusion

In this paper, we propose a new dynamic portfolio strategy based on the time-varying structures of MST networks for the Shanghai and Shenzhen A-Share markets. The strategy first select central and peripheral portfolios in the selection horizon using five topological parameters, and uses the selected portfolios for investment in the investment horizon. Nine



**Fig 10. Average return of the strategy with the largest probability of gaining more profits.** Average returns of the strategy (cross) that has the largest probability of gaining more profits than the random strategy (black solid line) in the Shanghai A-Share market (a) and the Shenzhen A-Share market (b). This strategy is described in Table 7.

doi:10.1371/journal.pone.0169299.g010

combinations of market conditions have been considered when comparing the excess returns of the central and the peripheral portfolios, which are identified either by the ratio of the number of trading days with rising index to the total number of trading days, or the ratio of the sum of the amplitudes of the trading days with rising index to the sum of the amplitudes of the total trading days. By picking out the portfolios with larger excess returns under different combinations of market conditions, the optimal portfolios under specific market conditions could be found: (i) If the market is likely to have a drawup trend in the following investment horizon, central portfolios should be the best choice, while the peripheral portfolios usually perform worse for excessive diversification. (ii) If the market is in a relatively stable state in the investment horizon, central portfolios are preferred unless the market just went through a drawdown trend in the selection horizon. (iii) If the market is likely to have a drawdown trend in the investment horizon and the market is stable in the selection horizon, peripheral portfolios should be chosen to reduce risks.

Empirical tests have also been carried out and verified the efficiency of our optimal portfolio strategy. We have used the data from 2000 to 2010 to select the optimal portfolio under each specific combination of market conditions through a training process. The selected optimal strategy is selected according to the current market condition to make investment based on the data from 2010 to 2014. By calculating the excess returns of the optimal portfolio strategies, our strategies are found to outperform the random strategy in most cases. Among all possible optimal portfolio strategies based on different parameters to select portfolios and different criteria to identify market conditions, 65% of our optimal portfolio strategies outperform the random strategy for the Shanghai A-Share market and 70% for the Shenzhen A-Share market. Using degree  $K$  as the parameter to select central or peripheral portfolios, returns of our strategies are always higher than those of the random strategy. The excess returns of the most profitable strategies in the Shanghai and Shenzhen A-Share markets are 0.0416 and 0.0632 respectively. The strategy that has the largest probability of gaining more profits outperforms the random strategy in 70% of the investment horizons for the Shanghai A-Share market, and 91.30% for the Shenzhen A-Share market.

## Supporting Information

**S1 File. Data sources.** The data sources and the specific URLs and steps to download the raw data used in the paper. A minimal data set is available on <http://pan.baidu.com/s/1kUXfj3D>. (PDF)

## Acknowledgments

We are grateful to Prof. Wei-Xing Zhou for helpful comments and suggestions. This work was partially supported by the National Natural Science Foundation (Nos. 10905023, 71131007, 71371165 and 11501199), Fok Ying Tong Education Foundation Grant 132013, Ningbo Natural Science Foundation (No. 2015A610160), and the Fundamental Research Funds for the Central Universities (2015).

## Author Contributions

**Conceptualization:** FR SPL.

**Data curation:** YNL.

**Formal analysis:** FR YNL.

**Investigation:** FR YNL.

**Methodology:** FR SPL.

**Software:** YNL.

**Supervision:** FR.

**Writing – original draft:** YNL FR.

**Writing – review & editing:** SPL XFJ LXZ TQ.

## References

1. Markowitz H. Portfolio selection. *J Finance*. 1952; 7:77–91. doi: [10.2307/2975974](https://doi.org/10.2307/2975974)
2. Fernández A, Gómez S. Portfolio selection using neural networks. *Comput Oper Res*. 2007; 34:1177–1191. doi: [10.1016/j.cor.2005.06.017](https://doi.org/10.1016/j.cor.2005.06.017)
3. Ko PC, Lin PC. Resource allocation neural network in portfolio selection. *Expert Syst Appl*. 2008; 35:330–337. doi: [10.1016/j.eswa.2007.07.031](https://doi.org/10.1016/j.eswa.2007.07.031)
4. Nazemi A, Abbasi B, Omid F. Solving portfolio selection models with uncertain returns using an artificial neural network scheme. *Appl Intell*. 2015; 42:609–621. doi: [10.1007/s10489-014-0616-z](https://doi.org/10.1007/s10489-014-0616-z)
5. Chen Y, Mabu S, Hirasawa K. A model of portfolio optimization using time adapting genetic network programming. *Comput Oper Res*. 2010; 37:1697–1707. doi: [10.1016/j.cor.2009.12.003](https://doi.org/10.1016/j.cor.2009.12.003)
6. Chen Y, Hirasawa K. A portfolio selection model using genetic relation algorithm and genetic network programming. *IEEJ T Electr Electr*. 2011; 6:403–413. doi: [10.1002/tee.20676](https://doi.org/10.1002/tee.20676)
7. Crama Y, Schyns M. Simulated annealing for complex portfolio. *Eur J Oper Res*. 2003; 150:546–571. doi: [10.1016/S0377-2217\(02\)00784-1](https://doi.org/10.1016/S0377-2217(02)00784-1)
8. Daly J, Crane M, Ruskin HJ. Random matrix theory filters in portfolio optimisation: A stability and risk assessment. *Physica A*. 2008; 387:4248–4260. doi: [10.1016/j.physa.2008.02.045](https://doi.org/10.1016/j.physa.2008.02.045)
9. Dai YH, Xie WJ, Jiang ZQ, Jiang GJ, Zhou WX. Correlation structure and principal components in global crude oil market. *Emp Econ*. 2016; 12:95–107.
10. Onnela JP, Chakraborti A, Kaski K, Kertesz J, Kanto A. Dynamics of market correlations: Taxonomy and portfolio analysis. *Phys Rev E*. 2003; 68:056110. doi: [10.1103/PhysRevE.68.056110](https://doi.org/10.1103/PhysRevE.68.056110)
11. Pozzi F, Di Matteo T, Aste T. Spread of risk across financial markets: better to invest in the peripheries. *Sci Rep*. 2013; 3:1665. doi: [10.1038/srep01665](https://doi.org/10.1038/srep01665) PMID: [23588852](https://pubmed.ncbi.nlm.nih.gov/23588852/)
12. Nanda SR, Mahanty B, Tiwari MK. Clustering Indian stock market data for portfolio management. *Expert Syst Appl*. 2010; 37:8793–8798. doi: [10.1016/j.eswa.2010.06.026](https://doi.org/10.1016/j.eswa.2010.06.026)
13. Liao SH, Chou SY. Data mining investigation of co-movements on the Taiwan and China stock markets for future investment portfolio. *Expert Syst Appl*. 2013; 40:1542–1554. doi: [10.1016/j.eswa.2012.08.075](https://doi.org/10.1016/j.eswa.2012.08.075)
14. Pai GAV, Michel T. Evolutionary optimization of constrained k-means clustered assets for diversification in small portfolios. *IEEE Trans Evolut Comput*. 2009; 13:1030–1053. doi: [10.1109/TEVC.2009.2014360](https://doi.org/10.1109/TEVC.2009.2014360)
15. Nanda SJ, Panda G. A survey on nature inspired metaheuristic algorithms for partitioning clustering. *Swarm Evolut Comput*. 2014; 16:1–18. doi: [10.1016/j.swevo.2013.11.003](https://doi.org/10.1016/j.swevo.2013.11.003)
16. Mantegna RN. Hierarchical structure in financial markets. *Eur Phys J B*. 1999; 11:193–197. doi: [10.1007/s100510050929](https://doi.org/10.1007/s100510050929)
17. Onnela JP, Kaski K, Kertész J. Clustering and information in correlation based financial networks. *Eur Phys J B*. 2004; 38:353–362. doi: [10.1140/epjb/e2004-00128-7](https://doi.org/10.1140/epjb/e2004-00128-7)
18. Tumminello M, Aste T, Di Matteo T, Mantegna RN. A tool for filtering information in complex systems. *Proc Natl Acad Sci USA*. 2005; 102:10421–10426. doi: [10.1073/pnas.0500298102](https://doi.org/10.1073/pnas.0500298102) PMID: [16027373](https://pubmed.ncbi.nlm.nih.gov/16027373/)
19. Tumminello M, Di Matteo T, Aste T, Mantegna RN. Correlation based networks of equity returns sampled at different time horizons. *Eur Phys J B*. 2007; 55:209–217. doi: [10.1140/epjb/e2006-00414-4](https://doi.org/10.1140/epjb/e2006-00414-4)
20. Brida JG, Rizzo WA. Dynamics and structure of the main Italian companies. *Int J Mod Phys C*. 2007; 11:1783–1793. doi: [10.1142/S0129183107011741](https://doi.org/10.1142/S0129183107011741)
21. Garas A, Argyrakis P. Filtering of complex systems using overlapping tree networks. *EPL(Europhys Lett)*. 2009; 86:28005. doi: [10.1209/0295-5075/86/28005](https://doi.org/10.1209/0295-5075/86/28005)
22. Aste T, Shaw W, Di Matteo T. Correlation structure and dynamics in volatile markets. *New J Phys*. 2010; 12:085009. doi: [10.1088/1367-2630/12/8/085009](https://doi.org/10.1088/1367-2630/12/8/085009)

23. Tumminello M, Lillo F, Mantegna RN. Correlation, hierarchies, and networks in financial markets. *J Econ Behav Org.* 2010; 75:40–58. doi: [10.1016/j.jebo.2010.01.004](https://doi.org/10.1016/j.jebo.2010.01.004)
24. Kwapień J, Drożdż S. Physical approach to complex systems. *Phys Rep.* 2012; 515:115–226. doi: [10.1016/j.physrep.2012.01.007](https://doi.org/10.1016/j.physrep.2012.01.007)
25. Yang CX, Zhu XH, Li Q, Chen YH, Deng QQ. Research on the evolution of stock correlation based on maximal spanning trees. *Physica A.* 2014; 415:1–18. doi: [10.1016/j.physa.2014.07.069](https://doi.org/10.1016/j.physa.2014.07.069)
26. West DB. *Introduction to Graph Theory.* Englewood Cliffs NJ: Prentice-Hall; 1996.
27. Onnela JP, Chakraborti A, Kaska K, Kertész J. Dynamic asset tree and portfolio analysis. *Eur Phys J B.* 2002; 30:285–288. doi: [10.1140/epjb/e2002-00380-9](https://doi.org/10.1140/epjb/e2002-00380-9)
28. Song DM, Tumminello M, Zhou WX, Mantegna RN. Evolution of worldwide stock markets, correlation structure, and correlation based graphs. *Phys Rev E.* 2011; 84:026108. doi: [10.1103/PhysRevE.84.026108](https://doi.org/10.1103/PhysRevE.84.026108)
29. Jiang XF, Chen TT, Zheng B. Structure of local interactions in complex financial dynamics. *Sci Rep.* 2014; 4:5321. doi: [10.1038/srep05321](https://doi.org/10.1038/srep05321) PMID: [24936906](https://pubmed.ncbi.nlm.nih.gov/24936906/)
30. Boginski V, Butenko B, Shirokikh O, Trukhanov S, Lafuente JG. A network-based data mining approach to portfolio selection via weighted clique relaxations. *Ann Oper Res.* 2014; 216:23–24. doi: [10.1007/s10479-013-1395-3](https://doi.org/10.1007/s10479-013-1395-3)
31. Choudhury S, Ghosh S, Bhattacharya A, Fernandes KJ, Tiwari MK. A real time clustering and SVM based price-volatility prediction for optimal trading strategy. *Neurocomputing.* 2014; 131:419–426. doi: [10.1016/j.neucom.2013.10.002](https://doi.org/10.1016/j.neucom.2013.10.002)
32. Ross GJ. Dynamic multifactor clustering of financial networks. *Phys Rev E.* 2014; 89:022809. doi: [10.1103/PhysRevE.89.022809](https://doi.org/10.1103/PhysRevE.89.022809)
33. Fenn DJ, Porter MA, Williams S, McDonald M, Johnson NF, Jones NS. Temporal evolution of financial-market correlations. *Phys Rev E.* 2011; 84:026109. doi: [10.1103/PhysRevE.84.026109](https://doi.org/10.1103/PhysRevE.84.026109)
34. Drożdż S, Grümmer F, Górski A, Ruf F, Speth J. Dynamics of competition between collectivity and noise in the stock market. *Physica A.* 2000; 287:440–449. doi: [10.1016/S0378-4371\(00\)00383-6](https://doi.org/10.1016/S0378-4371(00)00383-6)
35. Podobnik B, Wang D, Horvatic D, Grosse I, Stanley HE. Time-lag cross-correlations in collective phenomena. *EPL (Europhys Lett).* 2010; 90:68001. doi: [10.1209/0295-5075/90/68001](https://doi.org/10.1209/0295-5075/90/68001)
36. Kenett D, Tumminello M, Madi A, Gur-Gershgoren G, Mantegna R, Ben-Jacob E. Dominating clasp of the financial sector revealed by partial correlation analysis of the stock market. *PLoS One.* 2010; 5: e15032. doi: [10.1371/journal.pone.0015032](https://doi.org/10.1371/journal.pone.0015032) PMID: [21188140](https://pubmed.ncbi.nlm.nih.gov/21188140/)
37. Kenett DY, Raddant M, Lux T, Ben-Jacob E. Evolvement of uniformity and volatility in the stressed global financial village. *PLoS One.* 2012; 7:e31144. doi: [10.1371/journal.pone.0031144](https://doi.org/10.1371/journal.pone.0031144) PMID: [22347444](https://pubmed.ncbi.nlm.nih.gov/22347444/)
38. Ren F, Zhou WX. Dynamic Evolution of Cross-Correlations in the Chinese Stock Market. *PLoS One.* 2014; 9:e97711. doi: [10.1371/journal.pone.0097711](https://doi.org/10.1371/journal.pone.0097711) PMID: [24867071](https://pubmed.ncbi.nlm.nih.gov/24867071/)
39. Case B, Yang YW, Yildirim Y. Dynamic correlations among asset classes: REIT and stock returns. *J Real Estate Finance Econ.* 2012; 44:298–318. doi: [10.1007/s11146-010-9239-2](https://doi.org/10.1007/s11146-010-9239-2)
40. Kotkatvuori-Örnberg J, Nikkinen J, Äijö J. Stock market correlations during the financial crisis of 2008–2009: Evidence from 50 equity markets. *Int Rev Financial Anal.* 2013; 28:70–78. doi: [10.1016/j.irfa.2013.01.009](https://doi.org/10.1016/j.irfa.2013.01.009)
41. Miralles-Marcelo JL, Miralles-Quirós MM, Miralles-Quirós JL. Improving international diversification benefits for US investors. *N Am J Econ Finance.* 2015; 32:64–76. doi: [10.1016/j.najef.2015.01.005](https://doi.org/10.1016/j.najef.2015.01.005)
42. Sun XL, Liu ZX. Optimal portfolio strategy with cross-correlation matrix composed by DCCA coefficients: Evidence from the Chinese stock market. *Physica A.* 2016; 444:667–679. doi: [10.1016/j.physa.2015.10.065](https://doi.org/10.1016/j.physa.2015.10.065)
43. Podobnik B, Stanley HE. Detrended cross-correlation analysis: A new method for analyzing two nonstationary time series. *Phys Rev Lett.* 2008; 100:084102. doi: [10.1103/PhysRevLett.100.084102](https://doi.org/10.1103/PhysRevLett.100.084102) PMID: [18352624](https://pubmed.ncbi.nlm.nih.gov/18352624/)
44. Zhou WX. Multifractal detrended cross-correlation analysis for two nonstationary signals. *Phys Rev E.* 2008; 77:066211. doi: [10.1103/PhysRevE.77.066211](https://doi.org/10.1103/PhysRevE.77.066211)
45. Jiang ZQ, Zhou WX. Multifractal detrending moving-average cross-correlation analysis. *Phys Rev E.* 2011; 84:016106. doi: [10.1103/PhysRevE.84.016106](https://doi.org/10.1103/PhysRevE.84.016106)
46. Suganya NC, Pai GAV. Wavelet evolutionary network for complex-constrained portfolio rebalancing. *Int J Syst Sci.* 2012; 43:1367–1385. doi: [10.1080/00207721.2011.601351](https://doi.org/10.1080/00207721.2011.601351)
47. DeMiguel V, Garlappi L, Uppal R. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy. *Rev Financial Stud.* 2009; 53:1915–1953. doi: [10.1093/rfs/hhm075](https://doi.org/10.1093/rfs/hhm075)

48. Strogatz SH. Exploring complex networks. *Nature*. 2001; 410:268–276. doi: [10.1038/35065725](https://doi.org/10.1038/35065725) PMID: [11258382](https://pubmed.ncbi.nlm.nih.gov/11258382/)
49. Krings G, Karsai M, Bernhardsson S, Blondel VD, Saramäki J. Effects of time window size and placement on the structure of an aggregated communication network. *EPJ Data Science*. 2012; 1:1–16. doi: [10.1140/epjds4](https://doi.org/10.1140/epjds4)
50. Ledoit O, Wolf M. A well-conditioned estimator for large-dimensional covariance matrices. *J Multivar Anal*. 2004; 88:365–411. doi: [10.1016/S0047-259X\(03\)00096-4](https://doi.org/10.1016/S0047-259X(03)00096-4)
51. Mardia KV, Kent JT, Bibby JM. *Multivariate Analysis*. San Diego: Academic Press; 1979.
52. Kim HJ, Lee Y, Kahng B. Weighted scale-free network in financial correlations. *J Phys Soc Jpn*. 2002; 71:2133–2136. doi: [10.1143/JPSJ.71.2133](https://doi.org/10.1143/JPSJ.71.2133)
53. Lee KE, Lee JW, Hong BH. Complex networks in a stock market. *Comput Phys Commun*. 2007; 177:186. doi: [10.1016/j.cpc.2007.02.047](https://doi.org/10.1016/j.cpc.2007.02.047)
54. Micciche S, Bonanno G, Lillo F, Mantegna RN. Degree stability of a minimum spanning tree of price returns and volatility. *Physica A*. 2003; 324:66–73. doi: [10.1016/S0378-4371\(03\)00002-5](https://doi.org/10.1016/S0378-4371(03)00002-5)
55. Sharpe WF. The Sharpe Ratio. *J Portf Manag*. 1994; 21:49–58. doi: [10.3905/jpm.1994.409501](https://doi.org/10.3905/jpm.1994.409501)
56. Jagannathan R, Ma T. Risk reduction in large portfolios: why imposing the wrong constraints helps. *J Finance*. 2003; 54:1651–1683. doi: [10.1111/1540-6261.00580](https://doi.org/10.1111/1540-6261.00580)
57. Scholz H. Refinements to the Sharpe ratio: Comparing alternatives for bear markets. *J Asset Manag*. 2007; 7:347–357. doi: [10.1057/palgrave.jam.2250040](https://doi.org/10.1057/palgrave.jam.2250040)
58. Newman MEJ. Scientific collaboration networks. I. Network construction and fundamental results. *Phys Rev E*. 2001; 64:016131. doi: [10.1103/PhysRevE.64.016131](https://doi.org/10.1103/PhysRevE.64.016131)
59. Clauset A, Shalizi CR, Newman MEJ. Power-law distributions in empirical data. *SIAM Rev*. 2009; 51:661–703. doi: [10.1137/070710111](https://doi.org/10.1137/070710111)
60. Thorbecke W. On stock market returns and monetary policy. *J Finance*. 1997; 52:635–654. doi: [10.1111/j.1540-6261.1997.tb04816.x](https://doi.org/10.1111/j.1540-6261.1997.tb04816.x)
61. Bomfim A. Pre-announcement effects, news effects, and volatility: Monetary policy and the stock market. *J Bank Finance*. 2003; 27:133–151. doi: [10.1016/S0378-4266\(01\)00211-4](https://doi.org/10.1016/S0378-4266(01)00211-4)
62. Ferreira M, Gam P. Does sovereign debt ratings news spill over to international stock markets?. *J Bank Finance*. 2007; 31:3162–3182. doi: [10.1016/j.jbankfin.2006.12.006](https://doi.org/10.1016/j.jbankfin.2006.12.006)
63. Kasman S, Vardar G, Tunç G. The impact of interest rate and exchange rate volatility on banks' stock returns and volatility: Evidence from Turkey. *Econ Model*. 2011; 28:1328–1334. doi: [10.1016/j.econmod.2011.01.015](https://doi.org/10.1016/j.econmod.2011.01.015)
64. Tang Y, Luo Y, Xiong J, Zhao F, Zhang YC. Impact of monetary policy changes on the Chinese monetary and stock markets. *Physica A*. 2013; 392:4435–4449. doi: [10.1016/j.physa.2013.05.023](https://doi.org/10.1016/j.physa.2013.05.023)
65. Guo GF, Zheng ZF. Effect of Chinese stock market to the volatility of international energy-based on empirical econometric model. *Chin Ind Econ (in Chinese)*. 2011; 6:26–35.
66. Zhang B, Feng SX, Li XD, Wang HJ. Exchange rates and stock prices interactions in China: An empirical studies after 2005 exchange rate reform. *Econ Res J (in Chinese)*. 2008; 9:70–81.