Physica A 503 (2018) 1241–1255

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Multivariate financial time series in the light of complex network analysis

Sufang An [a,b], Xiangyun Gao [a,c,d,*], Meihui Jiang [a], Xiaoqi Sun [a]

[a] School of Economics and Management, China University of Geosciences, Beijing, 100083, China
[b] College of Information and Engineering, Hebei GEO University, Shijiazhuang, 050031, China
[c] Key Laboratory of Carrying Capacity Assessment for Resource and Environment, Ministry of Natural Resources, Beijing, 100083, China
[d] Key Laboratory of Strategic Studies, Ministry of Natural Resources, Beijing, 100083, China

HIGHLIGHTS

- We established a complex network for multivariable financial time series.
- A node shows the absolute returns of multi variables and the co-movement among them.
- We studied the dynamics characteristics of transition process in multivariable time series.
- We provided a new perspective for explaining the volatility clustering phenomenon.

ARTICLE INFO

Article history:
Received 18 April 2017
Received in revised form 25 May 2018
Available online 17 August 2018

Keywords:
Complex network
Time series
Dynamics characteristics
Financial market

ABSTRACT

We established a complex network from multivariate financial time series in which one node represents the types of states corresponding to the combination of the fluctuations of the crude oil future prices, the S&P 500 Index, the US Dollar Index, and gold future prices on a given day; one edge denotes the transition time from one node to another; and the weight is the transition frequency between two states. Through analyzing the network's topological structure, we obtain the characteristics of the transitions of these states in financial time series. The results show that nodes' out-strength distribution and betweenness centrality distribution both follow the power-law distribution. A shock to one financial market can be quickly transited to the other three financial markets and a transition probability matrix is proposed to predict the short-term financial market fluctuations. The transition characteristics under volatility clustering of the network that are obtained in this study provide a new perspective to explain financial volatility clustering, which extends the application of complex network theory to financial studies and helps investors understand the financial market.

1. Introduction

Revealing the dynamic behaviors from a univariate or multivariate time series is an outstanding and challenging problem in the research field of time series [1–6]. With the development of complex network theory [7–10], several new methodologies for time series using complex networks have emerged. These methodologies' primary idea is to map a time series into a complex network and later assess the network's structure to understand the dynamic characteristics of time series. There are many methods for the complex network analysis of univariate time series. The first method is the one

* Corresponding author at: School of Economics and Management, China University of Geosciences, Beijing, 100083, China.
E-mail address: gxy5669777@126.com (X. Gao).

https://doi.org/10.1016/j.physa.2018.08.063
0378-4371/

based on the visibility graph, such as a visibility graph method [11], a horizontal visibility graph (HVG) method [12], and a multiscale limited penetrable horizontal visibility graph (MLPHVG) [13]. These models have successfully been applied in many fields, such as exploring the dynamic properties of traffic flow time series in multi-states [14] and the universal and non-universal allometric scaling behaviors of world stock market indices [15]. The second one is a method based on phase space reconstruction [16,17], and it has been applied in different fields, including a reconstruction network for international gasoline regular spot prices series [18]. There are also the other methods for time series, such as a pseudo-periodic time series transition method [19], a coarse graining method applied in characterizing the transmission of autoregressive sub-patterns in financial time series [20] and analyzing the fluctuating behaviors of international crude oil and gasoline prices [21].

In recent years, complex network models of multivariate time series have been proposed, including inferring complex networks [22], the inferring direct directed-information flow from the multivariate nonlinear time series [23], a multilayer network [24], and a non-parametric method for multivariate time series [25]. These models have been used to solve problems in different fields, such as a complex network for researching the correlations between crude oil futures and spot prices time series [26], a complex network for linear regression patterns in the energy market [27], a multivariate weighted complex network analysis for characterizing nonlinear dynamic behaviors in two-phase flows [28], a fluctuation complex network model for investigating the RMB of exchange rates [29], and a complex network perspective on interrelations and evolutional features of international oil trade [30].

The previous studies have already demonstrated the great potential in solving challenging problems for time series, especially in financial markets [18,20,21,26,27,29,30]. However, when a long-period financial time series is divided into several sub-period time series, these sub-period time series transit into each other over time, which can help us analyze the dynamic characteristics of a long-period time series. We have proposed several new approaches by combining complex network theory with univariate [20] and bivariate financial time series [26,27] models. It is necessary to construct the complex network based on multivariate time series for understanding the evolution characteristics of the co-movement among multi-variables.

In this paper, we select four typical financial time series as sample data for studying. Crude oil is a basic raw material for economic development, and most markets directly or indirectly depend on energy markets. An increase in crude oil prices will affect the global economy. In economics, the relationships among crude oil, the US Dollar, the S&P 500 Index and gold have attracted considerable attention from many researchers [31–34], since the US Dollar is the most important currency in the world today. When US monetary policies change, the global economic structure changes, especially the trade in oil, since it is invoiced in US Dollars. The stock market serves as an economic barometer, and the S&P 500 Index that includes many of the largest companies in the world is the best measure of the short-period financial market's direction. Finally, the gold price is a traditional indicator of economic growth or recession and is a hedge against inflation when its value remains stable. The US Dollar is also the currency of the international gold market; therefore, fluctuations in gold prices are closely related to the US Dollar. Based on the above analysis, the study of the relationships among the four variables is an important issue for international investors to understand the financial markets.

Next, we explore the topological structure of the transition network to obtain the characteristics of the transitions. First, we convert the multivariable financial time series into a symbol sequence, where we propose a state as a 4-tuple of symbols corresponding to the absolute returns of four variables in a given day and obtain the state sequence to represent the transitions between the states. Second, we use the types of states as nodes, the transition from one state to another as an edge, the weight of which is the frequency of transition, and we subsequently construct a transition network. Third, we analyze the topological structure of the transition network. (1) We analyze the out-strengths of the nodes to identify the key transition nodes. (2) We discuss the betweenness centralities of the nodes to identify the transition betweenness

abilities of the nodes. (3) We study the average path length, the closeness centrality and the weighted clustering coefficient of the node to discuss the transition speed, and we propose a transition probability matrix to predict the future short-term financial market. (4) We analyze the distribution characteristics of the nodes with higher out-strengths and betweenness centralities. (5) We extract two new networks from the transition network with small and large fluctuations in every financial market and explore characteristics of the transitions in the new transition networks of every market. This approach provides a new perspective to study the volatility clustering effect problem.

2. Materials and methods

2.1. Data

We use the Cushing crude oil future contract 1 (US Dollars per barrel), the US Dollar Index, gold future prices (US Dollars per troy ounce), and the S&P 500 Index daily data as the sample data. Every time series contains 2559 data points from July 17, 2006 to November 11, 2016, which are derived from the U.S. Energy Information Administration (EIA) and finance.yahoo.com. The four financial time series are shown in Fig. 1(a)–(d), and their distribution characteristics are shown in Table 1.

Fig. 1(a)–(d) show that the volatility of the four financial markets is lower before February 2007. The volatility is higher from the second half of 2007 to 2009, since the United States subprime crisis occurred during this period. After that period, the four markets tend be quite stable for the recovery of the global economy. A new oil crisis broke out in November 2014.

Table 1
Distribution characteristics of four financial time series.

Name	Maximum	Minimum	Average	Variance	Standard deviation
Crude oil	145.29	26.21	78.37	540.78	23.33
S&P 500 Index	2190.15	676.53	1499.66	144596.43	380.23
US Dollar Index	100.31	71.30	83.33	49.39	7.03
Gold	1888.70	562.00	1183.20	102571.61	320.62

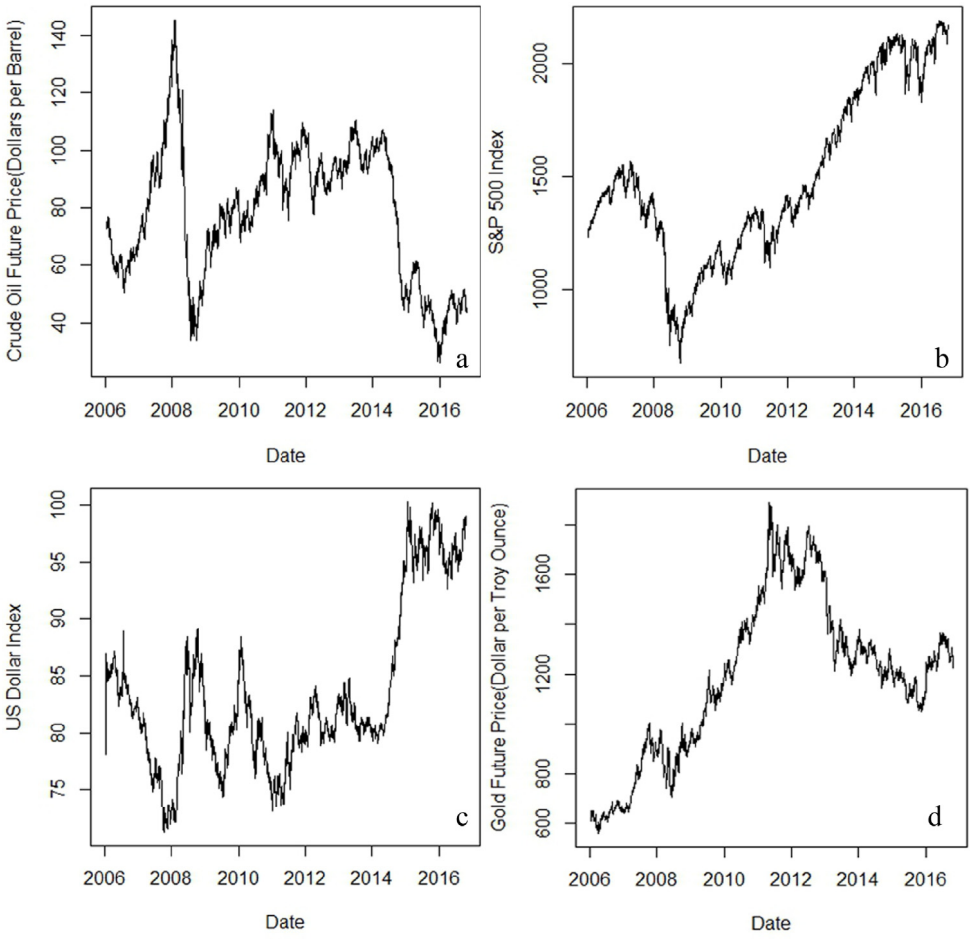

Fig. 1. (a) The time series of crude oil future prices; (b) the time series of the S&P 500 Index; (c) the time series of the US Dollar Index; (d) the time series of gold future prices.

As a result, the volatility is higher during this long-term period, where the WTI future prices crashed to approximately $26 per barrel early in the first quarter of 2016 (Fig. 1(a)).

2.2. Symbolization

The symbolization of a time series is a coarse-graining process for the data, and the obtained symbol sequence can characterize the main features of the time series. One approach of symbolization is to directly incorporate the time-series data into the symbolic sequence according to the characteristics of the database, such as the static transformation or the dynamic transformation [35]. Another approach is to use some transformation strategies, such as the phase pace coarsening graining (a wavelet transformation) [36], to convert the data into a new sequence for generating the symbol sequence.

In this paper, we set a symbol set as $\{R, r, e, d, D\}$ to represent the fluctuation range of the time-series data, where R represents a sharp rise, r represents a rise, e represents a stable one, d denotes a decline and D denotes a sharp decline. For a time series $P(t)(t = 0, 1, \ldots, N)$, we first calculate the absolute return series $\Delta P(t) = P(t) - P(t-1)$, where $P(t)$ is the

Table 2
Probability of five symbols in four markets.

Name	Crude oil (%)	S&P 500 Index (%)	US Dollar Index (%)	Gold (%)
R	19.01	18.44	18.67	19.58
r	31.10	35.84	31.25	31.91
e	0.12	0.04	1.19	0.28
d	30.41	28.21	29.98	30.87
D	19.36	17.47	18.91	17.36

Table 3
Examples of the multivariable transition state.

Time Label	Date	Crude oil	S&P 500 Index	US Dollar Index	Gold	State	State Label
1	2006/7/18	*D*	*r*	*R*	*D*	*DrRD*	$s(1)$
2	2006/7/19	*d*	*R*	*D*	*R*	*dRDR*	$s(2)$
3	2006/7/20	*r*	*d*	*R*	*D*	*rdRD*	$s(3)$
...	...	...	...	...	...	...	...
N	2016/11/11	*D*	*d*	*r*	*D*	*DdrD*	$s(N)$

current price and $P(t-1)$ is the previous one. Thus, we obtain the symbol sequence as follows:

$$f(t) = \begin{cases} R, \Delta P(t) > E \\ r, 0 < \Delta P(t) \leq E \\ e, \Delta P(t) = 0 \\ d, -E < \Delta P(t) < 0 \\ D, \Delta P(t) \leq -E \end{cases} \tag{1}$$

where E represents a fluctuation threshold that is different from the traditional definition used in risk management [37]. It can be described as follows:

$$E = \sum_{i}^{N-1} |\Delta P_i(t)| / N - 1 \tag{2}$$

This method is transparent and easy. We can also increase the symbol set to infinitely approximate the original time series. However, if the number of symbols in the set is too large, there is no significance in the symbolization. Therefore, we need to select an appropriate threshold according to the actual situation, and in general the mathematical expectations can be used as a reference threshold.

In this paper, we choose five symbols to symbolize the financial time series. Therefore, the absolute return series was converted into the symbol sequence $F = \{f(t)\}, f(t) \in \{R, r, e, d, D\}, t = 1, 2, \ldots, N$. The probability of five symbols in four markets is shown in Table 2. The statistical information shows that the number of *R*s and *D*s in every market is small, indicating that there are a few sharp fluctuations in the four markets.

2.3. Constructing the transition network

We propose a state as a 4-tuple of symbols corresponding to the four markets' tendencies on a given day and then obtain the state sequence shown in Table 3. The states evolve into each other over time and can be described as follows:

$$s(1) \to s(2) \to \cdots \to s(i-1) \to s(i) \to \cdots \to s(N) \tag{3}$$

where $s(i)$ is the i-th state, $i = 1, 2, \ldots, N(N = 2558)$ and $s(i-1) \to s(i)$ represents the transition from s $(i-1)$ to $s(i)$.

For example, the transition of the time series shown in Table 3 involves "*DrRD*"→"*dRDR*"→ "*rdRD*" → ⋯ → "*DdrD*", where the state "*DrRD*" indicates that the crude oil and gold future prices decline sharply, the S&P 500 Index increases and the US Dollar Index increases sharply.

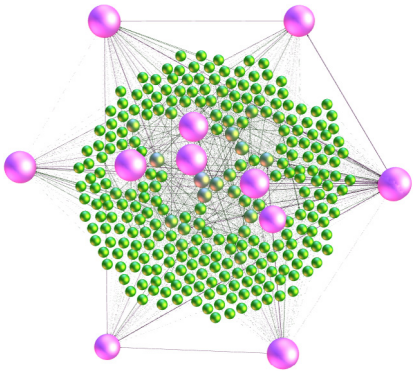

Fig. 2. Transition network of the multivariate financial time series.

We propose a transition network based on complex network theory, which can be defined as an ordered pair $G_\beta = (V_\beta, E_\beta)$ as follows:

$$G_\beta = \left(V_\beta, E_\beta\right) = \begin{bmatrix} \beta_{1,1} & \beta_{1,2} \cdots & \beta_{1,n} \\ \vdots & \ddots & \vdots \\ \beta_{n,1} & \beta_{n,2} \cdots & \beta_{n,n} \end{bmatrix} \tag{4}$$

where V_β is a set of nodes, E_β is a set of edges, and every element in the matrix represents an adjacent relationship between two nodes. We define the types of states as nodes and the transition from one state to another as an edge with a weight that is the frequency of the transition between the two states (weight is the element of the matrix). Therefore, we obtain a transition network (TN) with 280 nodes and 2425 edges as shown in Fig. 2.

3. Results

Traditional researchers have been interested in the relationships between multivariable financial time series and rarely explain transitions based on complex network theory. In this paper, we construct a TN to characterize the transition process of four financial time series, and we explore the topological structure of the TN to obtain the characteristics of the transition process.

3.1. Key transition nodes

When a node represents the absolute returns of four variables in the daily financial market, the out-strength of a node represents the frequency of the transition of this node to its neighbors, which are the absolute returns of the four variables for the next day. We identify the key transition nodes as the nodes with greater out-strength, which indicate that the frequency of the transition of a key transition node is larger. The out-strength of the node can be described as follows:

$$S_i = \sum_{j}^{K_i} w_{ij} \tag{5}$$

where w_{ij} is the weight for node i to node j, and K_i is the number of neighbors of node i.

The distribution characteristics of the out-strength is shown in Table 4 in descending order, where S is the out-strength, P(S) is the probability of the out-strength and CP(S) is the cumulative probability of the out-strength. In this table, "*rrdr*" is the node with the maximum value of the out-strength, thereby demonstrating that the crude oil future prices, the S&P 500 Index and gold future prices rise, while the Dollar Index declines. This table also indicates that "*rrdr*" is a key transition node with the greatest out-strength. Furthermore, the out-strength of a node follows a power-law distribution (Fig. 3(a)). This finding indicates that fewer nodes have greater out-strength, and the nodes with greater out-strength are larger in Fig. 2. The results show that 25% of nodes account for 60.22% of the total out-strength, as shown in Fig. 3(b), thereby indicating that there are few nodes with larger frequency of transition.

Moreover, to further examine the key transition nodes, we choose 25% of nodes to form a transition network composed of key transition nodes, as shown in Fig. 4, where the out-strengths of these nodes are greater than 12 in the TN, and there are 870 edges. In Fig. 4, the out-strength of a node represents the frequency of transition of this node corresponding to its new neighbors in the transition network composed of key transition nodes. The results show that the out-strength of "*rrdr*" in the TN is 60, and the out-strength of this node in Fig. 4 (blue and largest node) is 42. This finding indicates that the frequency

Table 4
Top 5 out-strength distribution characteristics of the transition network.

Rank	Node	S	P(S) (%)	CP(S) (%)
1	*rrdr*	60	2.32	2.32
2	*drrd*	52	2.00	4.32
3	*rrrd*	45	1.73	6.05
4	*rrrr*	45	1.73	7.78
5	*ddrd*	44	1.69	9.47

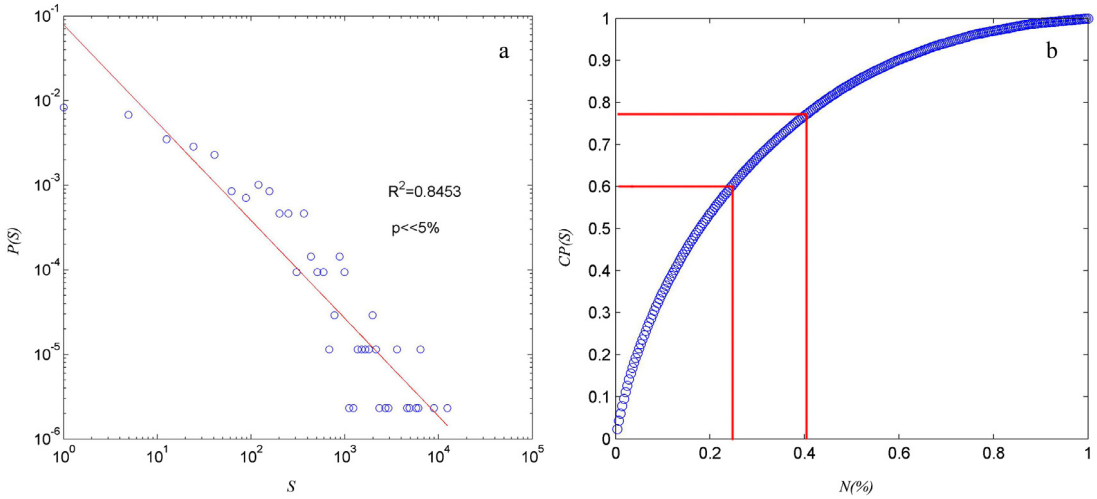

Fig. 3. (a) Double-logarithmic plot between the out-strength of node and its probability (the goodness of fit $R^2 = 0.8453$, and the P value $p \ll 5\%$ which passes the significance test); (b) the cumulative probability of the out-strength of the node (N is the Normalized Rank in Table 4).

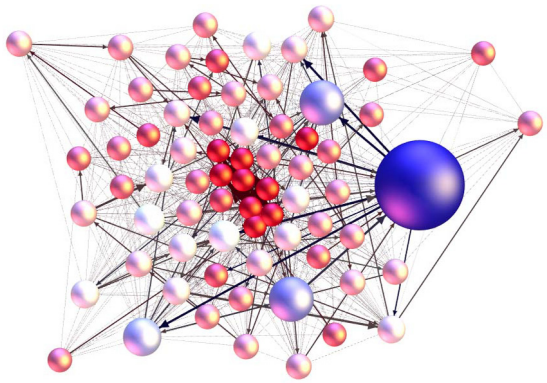

Fig. 4. Transition network composed of key transition nodes.

of transition of "*rrdr*" corresponding to the key transition nodes is 42, and the frequency to the non-key transition nodes in the TN is 18.

3.2. Transition betweenness

The betweenness centrality of a node can identify the transition betweenness ability of a node in the TN. A node with higher betweenness centrality provides a greater transition betweenness ability, thereby indicating that the node as a betweenness plays an important mediating role in the transition. In this paper, the betweenness centrality of a node is equal to the number of shortest paths from all nodes to all others that pass through that node in a complex network. The

Table 5
Top 5 betweenness centrality distribution characteristics of the transition network.

Rank	Node	BC	P(BC) (%)	CP(BC) (%)
1	*drrd*	0.069	3.69	3.69
2	*rrdr*	0.064	3.38	7.07
3	*ddrd*	0.063	3.35	10.42
4	*DDRD*	0.058	3.08	13.50
5	*rrrr*	0.057	3.04	16.54

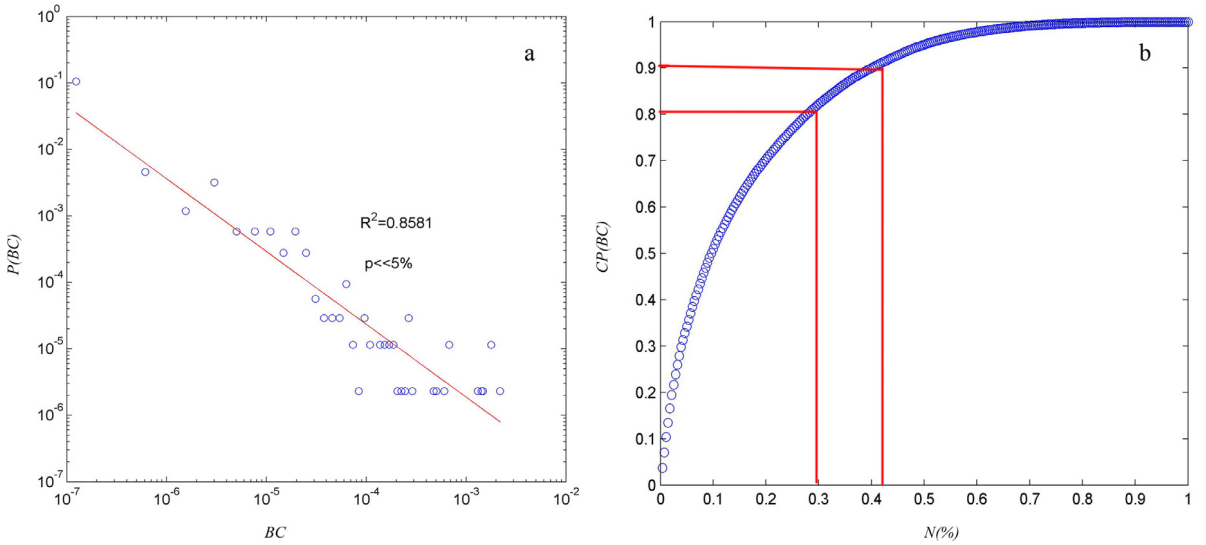

Fig. 5. (a) Double-logarithmic plot between the betweenness centrality of the node and its probability (we define 0.001 as the interval extension of the BC to allocate the BC into different intervals to analyze the relationship between the BC and the P(BC), the goodness of fit $R^2 = 0.8581$, and the P value $p \ll 5\%$ which passes the significance test); (b) the cumulative probability of the out-strength of the node (N is the Normalized Rank in Table 5).

betweenness centrality can be defined as follows:

$$BC_i = \sum_{j}^{n} \sum_{k}^{n} \frac{g_{jk}(i)}{g_{jk}} / n^2 - 3n + 2, j \neq k \neq i, j < k \tag{6}$$

where $g_{jk}(i)$ is the number of shortest paths from node j to node k that cross node i. g_{jk} is the total number of shortest paths from node j to node k.

The distribution characteristics of betweenness centrality are shown in Table 5 in descending order, where BC is the normalized betweenness centrality, P(BC) is the probability of betweenness centrality, and CP(BC) is the cumulative probability of betweenness centrality. In Table 5, the node "*drrd*" with the maximum value of betweenness centrality denotes that crude oil and gold future prices decline, while there are increases in the S&P 500 Index and the US Dollar Index. The node "*drrd*" as a betweenness plays the most important mediating role in the transition. Meanwhile, the betweenness centrality of a node in the TN follows a power-law distribution, as shown in Fig. 5(a), thereby indicating that few absolute returns of the four variables as betweennesses play an important mediating role in the transition. Furthermore, there are 28.21% nodes accounting for 80.48% the total betweenness centrality value, as shown in Fig. 5(b).

Moreover, to further examine the nodes with higher transition betweenness ability, we choose these 28.21% nodes in the TN to construct a new transition network with 895 edges, as shown in Fig. 6, where the betweenness centralities of these nodes are greater than or equal to 0.007. There are different betweenness centralities of the nodes in different transition networks. For example, the betweenness centrality of "*drrd*" in the new network is 0.061 and the betweenness centrality of "*drrd*" in the TN is 0.069, as shown in Fig. 2. The result shows that "*drrd*" is the node with the greatest betweenness centrality in the TN or the new transition network shown Fig. 6 (the red and largest node).

3.3. Transition speed and preference

(1) Transition speed

We measure the transition speed with the average path length and closeness centrality of the node, where the average path length is defined as the average number of steps along the shortest paths for all possible pairs of transition network

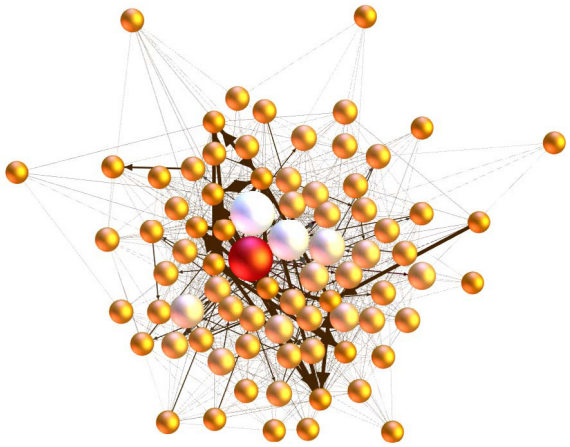

Fig. 6. Transition network composed of nodes with higher levels of betweenness centrality . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 6
Top 5 closeness centrality distribution characteristics of the transition network.

Rank	Node	CC	P(CC) (%)
1	*drrd*	2.02	0.25
2	*rrdr*	2.05	0.25
3	*ddrd*	2.09	0.26
4	*rrrr*	2.11	0.26
5	*rrrd*	2.15	0.27

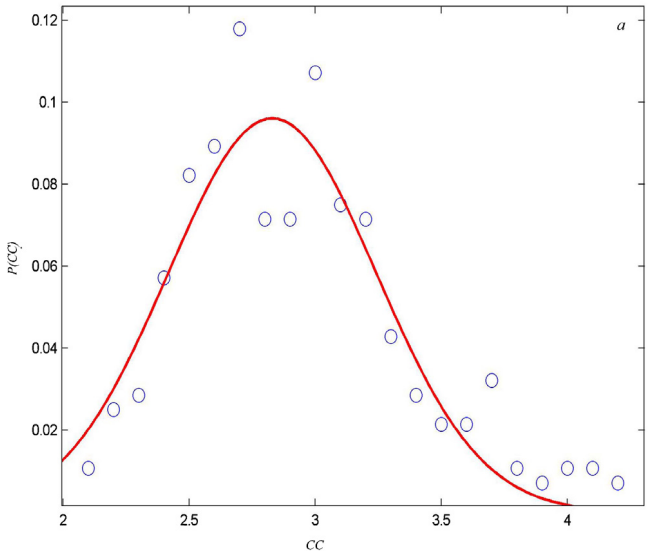

Fig. 7. Plot between the closeness centrality of the node and its probability (we define 0.1 as the interval extension of the CC to allocate the CC into different intervals to analyze the relationship between the CC and the P(CC)).

nodes, and the closeness centrality of a node represents the average length of the shortest paths between the nodes and all the other possible nodes.

The average path length of the TN is 2.87, and the diameter is 6. The distribution characteristics of the closeness centrality are shown in Table 6 in ascending order, where CC is the closeness centrality of a node, P(CC) is the probability of the closeness centrality. In Table 6, "*drrd*" with the minimum value of the closeness centrality demonstrates that the crude oil and gold future prices decline, and the S&P 500 Index and the Dollar Index rise. This finding indicates that "*drrd*" can easily evolve

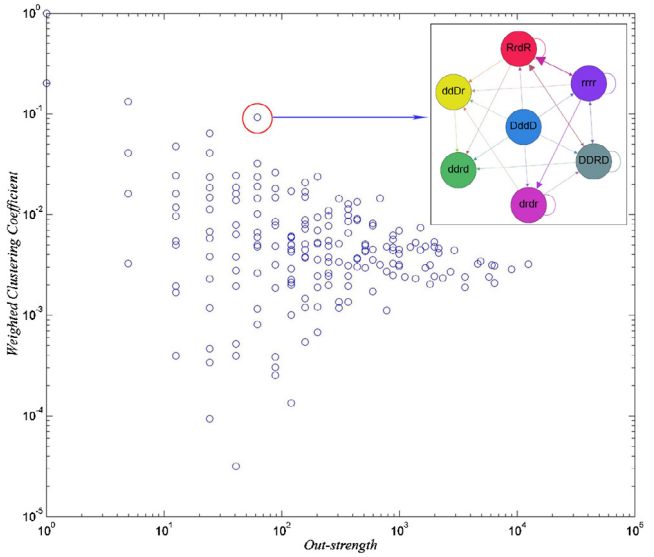

Fig. 8. Double-logarithmic plot between the out-strength and weighted clustering coefficient of the node (Pearson correlation is −0.351, and the P value $p \ll 5\%$ which passes the significance test) . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

into other nodes in the TN. Furthermore, the distribution of the closeness centrality of a node is shown in Fig. 7. The result shows the values of closeness centrality mainly distribute between 2.6 and 3.2.

The minimum and maximum values of the closeness centrality are 2.02 and 4.19, respectively, thereby indicating that the transition speed is fast. The result means that when one variable fluctuates in four financial markets, people often have no time to feel the effects of today's fluctuation over the following days that may have already caused financial market turmoil.

(2) Weighted clustering coefficient of a node in TN.

The nodes in a real-world network tend to create tightly knit groups characterized by a higher density of ties. A node with a greater weighted clustering coefficient indicates that the neighbors of the node have closer ties. The weighted clustering coefficient of the node [38] can be described as follows:

$$C_i^w = \frac{1}{S_i\,(k_i - 1)} \sum_{j,k} \frac{w_{ij} + w_{ik}}{2} a_{ij} a_{jk} a_{ik} \tag{7}$$

where k_i is the degree of node i, w_{ik} is the weight from node i to node k, $a_{ij}a_{jk}a_{ik} = 1$ denotes that node i, node j and node k can form a triangle and $a_{ij}a_{jk}a_{ik} = 0$ denotes that these three nodes cannot form a triangle.

We analyze the relationship between the out-strength and the weighted clustering coefficient of the node, as shown in Fig. 8. This analysis indicates that there is no strong correlation between them. However, there are several interesting nodes that have greater out-strengths and weighted clustering coefficients, and they indicate that the node whose neighbors have closer ties have higher transition frequencies corresponding to its neighbors. For example, the out-strength of "*DddD*" (red circle in Fig. 8.) is 6, and the weighted clustering coefficient of that node is 0.36. This finding indicates that the node transits to six neighbors with closer ties, as shown in the small plot of Fig. 8.

(3) Transition preference

In the TN, the nodes transit into each other over time. Therefore, it is necessary to analyze the transition preferences between nodes, which creates a basis for short-term forecasting in financial markets [20,29]. The transition probability between nodes that shows the transition preference is defined as follows:

$$p_{i\to j} = w_{ij}/S_i \tag{8}$$

where $p_{i\to j}$ denotes the transition probability of the current node to the next node in the TN. To explain this concept, some nodes in the TN are shown as an example in Fig. 9, and we represent the transition pattern as the source node → target node (transition probability).

If the source node is "*drdD*", there are 12 target ones, and one of the target nodes is "*rRrd*" with the greatest transition probability of 15.38%. The sum of the transition probabilities of the source node "*drdD*" corresponding to its neighbors is 100%. This finding indicates that the target node with the greatest transition probability is a better choose for investors when the fluctuation of four variables is the source node. Meanwhile, the node "*rrrr*" transits to itself. The fluctuations of the four variables with slight rises indicate that the four financial markets during this period occupy a stable state.

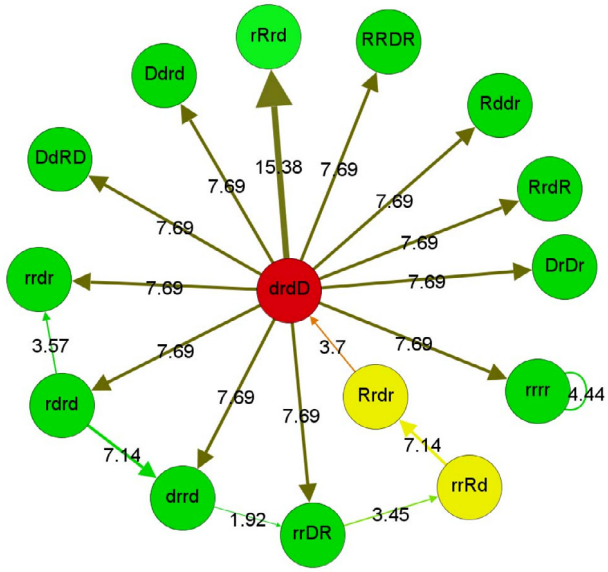

Fig. 9. Example of a transition route in a transition network.

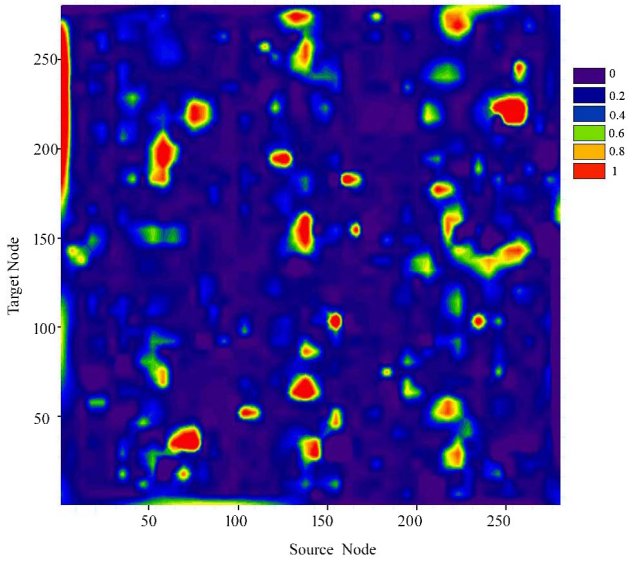

Fig. 10. Distribution characteristics of the transition probability between the source node and the target node.

Furthermore, we obtain the following route in Fig. 9: *drdD* → *rrDR*(7.69%) → *rrRd*(3.45%) → *Rrdr*(7.14%) → *drdD*(3.7%). It shows that the transition route is a closed loop, thus indicating that the nodes in the loop only can be transited in one direction, and if any node disappears, the transition in the loop will be paralyzed.

We can obtain a transition probability matrix of the TN where the element in the matrix represents the transition probability from the source node to the target one, and the matrix is represented in the Appendix. In this matrix, the sum of the columns is equal to 100%, which represents the summed transition probability of the same source node. Only 17 diagonal elements are not zero, thus indicating that these nodes transit to themselves. This means that most nodes transit to others, and the four financial markets occupy a complex state for a long time.

The distribution characteristics of the transition probability between the source and target nodes are shown in Fig. 10. The results show the transition probabilities between nodes. We find that the transitions between nodes are not random, and the nodes with higher transition probabilities concentrate in some small areas. This finding means that there are transition preferences in the transition process.

Table 7
Maximum and minimum absolute returns between "*rrdr*" and the four variables.

Name		Maximum	Minimum
Crude oil & *rrdr*	Crude oil	16.37	−14.31
	rrdr	1.23	0.01
S&P 500 Index & *rrdr*	S&P 500 Index	104.13	−106.85
	rrdr	11.22	0.28
US Dollar Index & *rrdr*	US Dollar Index	8.41	−8.92
	rrdr	−0.01	−0.33
Gold & *rrdr*	Gold	70.1	−140.4
	rrdr	10.1	0.1

Table 8
Transition distribution characteristics of the transition network under volatility clustering.

Name	Node	Edge	AS	AWCC	ABC
Crude oil (large change)	129	389	3.19	0.06	0.02
Crude oil (small change)	139	700	5.55	0.12	0.01
S&P 500 Index (large change)	120	364	3.11	0.05	0.02
S&P 500 Index (small change)	140	799	6.34	0.14	0.01
US Dollar Index (large change)	120	385	3.36	0.06	0.02
US Dollar Index (small change)	142	705	5.50	0.12	0.01
Gold (large change)	117	330	2.98	0.04	0.02
Gold (small change)	145	703	5.30	0.12	0.01

In fact, when the source node with a higher transition probability is the one with smaller out-strength, the transition between two nodes may be meaningless. For example, the transition pattern is *RRer* → *Drrd* (100%) and the out-strength of "*RRer*" is 1, indicating that the transition between these two nodes should be ignored. Therefore, we suggest that the investors use the transition matrix according to the actual situation and the out-strength of the node.

3.4. Nodes with higher out-strengths and betweenness centralities

From Tables 4 and 5, we find that the node "*rrdr*" has the greatest out-strength and the second highest value of betweenness centrality. This finding shows that crude oil future prices, the S&P 500 Index and gold future prices rise and US Dollar Index declines. This result shows that "*rrdr*" is a key transition node and plays the second most important mediating role as a betweenness for transitions.

We also explore the characteristics of "*rrdr*" by calculating the maximum and minimum absolute returns between "*rrdr*" and the four variables, as shown in Table 7. The results show that the absolute returns of crude oil future prices belong to [−14.31,16.37] and those of "*rrdr*" against crude oil are [0.01,1.23]. Therefore, "*rrdr*" represents that the absolute returns of crude oil are positive and fluctuate only slightly.

We plot the distribution between "*rrdr*" and the four variables in Fig. 11(a)–(h), where the left plots (blue lines) denote the absolute returns of the four variables and the right plots (red lines) denote the absolute returns against these variables when the symbol set of absolute returns is "*rrdr*". The results show that the node is distributed across the fluctuation of crude oil, gold and S&P 500 Index with a slight rise and that of US Dollar Index with a slight decline. The fact that there is only one node for August 03, 2007 to May 13, 2009 indicates that the fluctuations of the four variables are large changes and "*rrdr*" does not heavily affect the transition process during the United States sub-prime crisis period.

3.5. Transition characteristics under volatility clustering

Financial time series often exhibit volatility clustering properties. Large or small changes in financial market prices tend to cluster together, resulting in persist amplitudes of prices changes [39]. Many scholars and investors have proposed stochastic models in the field of volatility clustering in financial markets [40–42]. In this paper, we present a new way to explain this phenomenon: transition characteristics under volatility clustering.

When the absolute value of absolute returns of crude oil future prices is greater, it indicates that the fluctuation of crude oil prices has undergone a large change. We can extract a crude oil network undergoing a large change from the TN (Fig. 12(a)), where the crude oil column of Table 3 is only equal to *R* or *D*. This reflects the transition process of fluctuations on the whole-term time series when the crude oil prices have undergone large changes. We analyze the distribution characteristics of the crude oil network in Table 8, where the network includes 129 nodes, 389 edges, an average out-strength of 3.19, an average weighted clustering coefficient of 0.06 and an average betweenness centrality of 0.02. The results show that 46.07% of the nodes in the TN form the crude oil network, and there are many edges.

When fluctuation of the crude oil price undergoes a small change, we can also extract a crude oil network undergoing a small change from the TN shown in Fig. 12(b), where the crude oil column of Table 3 belongs to $\{r, e, d\}$. The distribution characteristics are shown on the second line of Table 8, where there are 139 nodes and 700 edges. The result shows that both

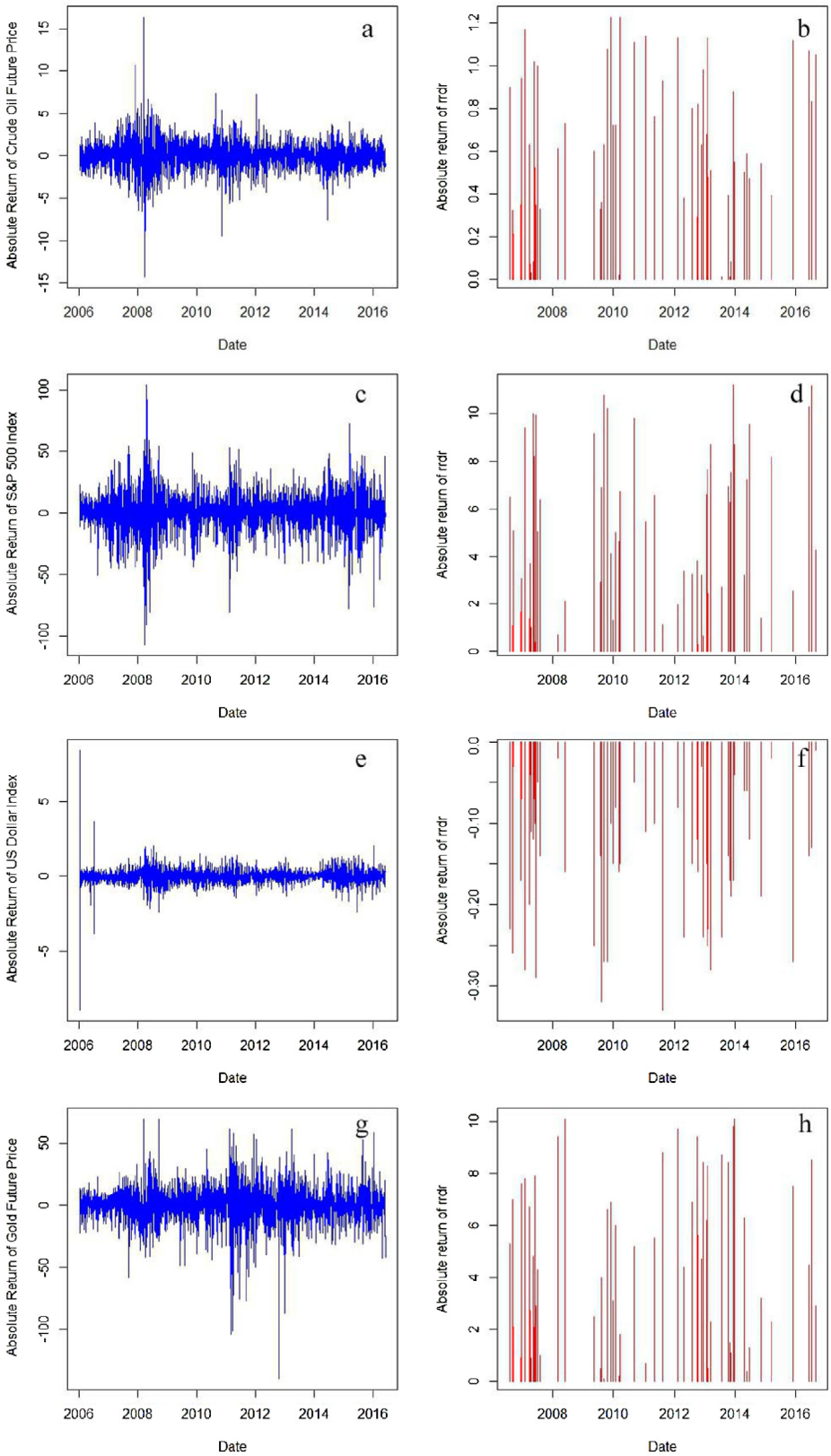

Fig. 11. (a) Absolute returns of crude oil future prices; (b) absolute returns of "*rrdr*" against crude oil future prices; (c) absolute returns of the S&P 500 Index; (d) absolute returns of "*rrdr*" against the S&P 500 Index; (e) absolute returns of the US Dollar Index; (f) absolute returns of "*rrdr*" against the US Dollar Index; (g) absolute returns of gold future prices; (h) absolute returns of "*rrdr*" against gold future prices.

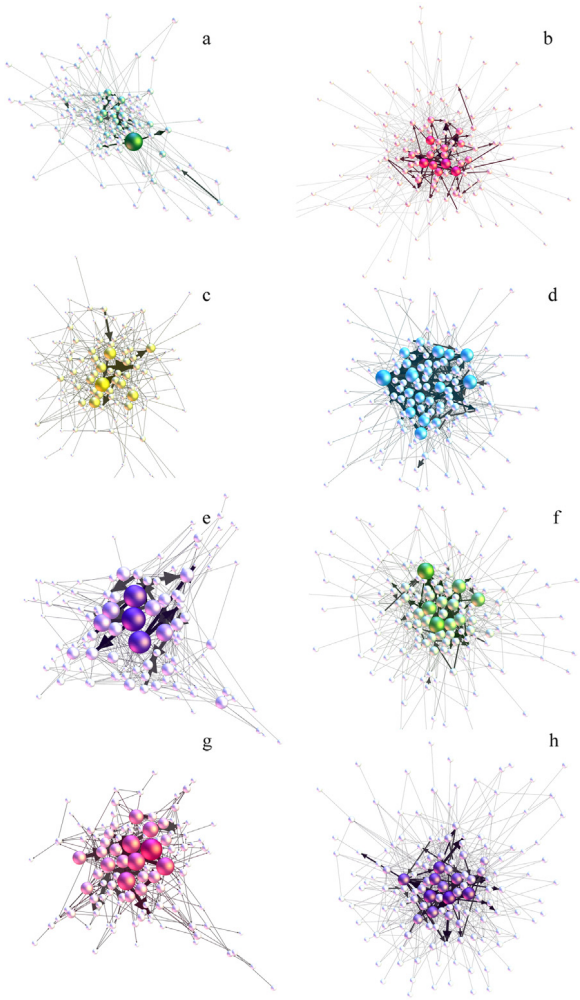

Fig. 12. (a) Crude oil network undergoing a large change; (b) crude oil network undergoing a small change; (c) S&P 500 Index network undergoing a large change; (d) S&P 500 Index network undergoing a small change; (e) US Dollar Index network undergoing a large change; (f) US Dollar Index network undergoing a small change; (g) gold network undergoing a large change; (h) gold network undergoing a small change.

the average out-strength and the average weighted clustering coefficient in Fig. 12(b) are greater than those in Fig. 12(a), and the average betweenness centrality in Fig. 12(b) is smaller than that in Fig. 12(a). According to the network extraction method, we can obtain six other transition networks from the TN shown in Fig. 12(c)–(h) when the fluctuations of the other three variables are large or small changes. The results are shown in Table 8.

Moreover, to further examine the nodes of the networks (Fig. 12(a)–(h)), we propose the transition probability between the transition networks undergoing large and small changes based on the fluctuations of every financial market, which is defined as

$$Tran_{a\to b} = \sum_{i\in C_a, j\in C_b} w_{ij} / \sum_{a=1}^{2} \sum_{i\in C_a, j\in C_b} w_{ij} \tag{9}$$

where node i and node j belong to the transition network undergoing a large or small change C_a and C_b, respectively. If $C_a = C_b$, $Tran_{a\to b}$ represents the transition probability of the nodes transiting to the other nodes within the same network. If $C_a \neq C_b$, $Tran_{a\to b}$ represents the transition probability of nodes transiting to the other nodes of the other transition network.

Fig. 13(a)–(d) show the transition probabilities between transition networks undergoing large and small changes based on the fluctuations of every financial market. When we discuss the nodes of transition networks undergoing large changes (the left plots of Fig. 12), the transition probability of the nodes transiting to the other nodes within the same network is smaller. However, the transition probability of these nodes transiting to the other nodes of other networks undergoing small changes is higher. The opposite occurs when the nodes of the transition network undergo small changes (the right

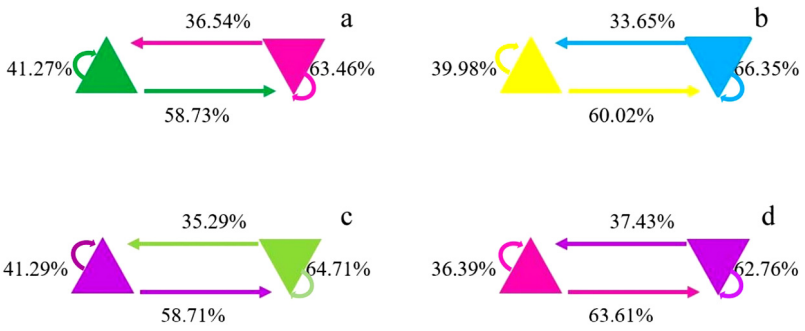

Fig. 13. (a) Transition probability between crude oil networks; (b) transition probability between S&P 500 Index networks; (c) transition probability between US Dollar Index networks; (d) transition probability between gold networks.

plots of Fig. 12). For example, the transition probability of the nodes in Fig. 12(a) transiting to the other nodes within the same network is 41.27% and that of the nodes transiting to the other nodes of Fig. 12(b) is 58.73%. Meanwhile, the transition probability of the nodes in Fig. 12(b) transiting to the other nodes within the same network is 63.46%, and that of the nodes transiting to the other nodes of Fig. 12(a) is 36.54%.

4. Discussion and conclusions

In this paper, we analyze the evolution of the co-movement among multivariate financial time series in light of complex network analysis. We construct a transition network to explore the characteristics of the transitions in financial system in which one node represents the types of states corresponding to the combination of the fluctuations of the crude oil future prices, the S&P 500 Index, the US Dollar Index, and gold future prices on a given day, one edge denotes the transition time from one node to another, and the weight is the frequency of transition between two states.

(1) A node composed of four symbolized variables of the financial market on a given day indicates the absolute returns of the four variables and the co-movement among them. The identification of the node as the key transition node can help us to understand the node with higher frequency of the transition to its neighbors. And the identification of the nodes with greater betweenness centrality can help us understand the nodes as betweennesses in the transition process. The results show that nodes' out-strength distributions and betweenness centrality distributions both follow the power-law distribution. It indicates that a few nodes as betweennesses play an important mediating role in transitions and a few nodes with greater out-strengths are the key transition nodes.
(2) A probability matrix is proposed to predict the short-term financial market fluctuations, where the source node transits into one or two target nodes with higher probabilities. Meanwhile, the identification of the node with a greater out-strength and higher weighted clustering coefficient can help us to understand the node with a higher frequency of transition corresponding to its neighbors with closer ties.
(3) The node "*rrdr*" indicates that crude oil future prices, the S&P 500 Index and gold future prices rise and the Dollar Index declines. Moreover, the node "*rrdr*" with the greatest out-strength indicates that it is a key transition node. And this node with the second greatest betweenness centrality indicates that it plays the second most important mediating role as a betweenness in the transition process. Furthermore, here is only one "*rrdr*" for August 03, 2007 to May 13, 2009, which indicates that the fluctuations of the four variables are large and "*rrdr*" does not heavily affect the transition process during the United States sub-prime crisis period.
(4) To explain volatility clustering, we construct the transition networks undergoing large and small changes based on the fluctuation of every financial market. When we discuss the nodes of transition networks undergoing large changes, the transition probability of the nodes transiting to the other nodes within the same network is smaller. However, the transition probability of these nodes transiting to the other nodes of other networks undergoing small changes is higher. The opposite occurs when the nodes of the transition network undergo small changes.

In our study, we select four typical variables in the financial system to explore the characteristics of the transitions in the system. In fact, there are complex co-movements among various variables in the financial system. Future study will consider more variables to study the characteristics of the transitions in the financial system.

Acknowledgments

This research is supported by grants from the Humanities and Social Sciences Planning Funds, China project under the Ministry of Education of the PRC (Grant No. 17YJCZH047), the Beijing Natural Science Foundation, China (Grant No. 9174041), the fund from the Key Laboratory of Carrying Capacity Assessment for Resource and Environment, Ministry of Natural

Resources, China (Grant No. CCA2017.11) and the Fundamental Research Funds for the Central Universities, China (Grant No. 2-9-2015-303).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.physa.2018.08.063.

References

[1] Y. Zou, M. Small, Z. Liu, J. Kurths, Complex network approach to the statistical features of the sunspot series, New J. Phys. 16 (2013) 499–502.
[2] B. Podobnik, H.E. Stanley, Detrended cross-correlation analysis: a new method for analyzing two nonstationary time series, Phys. Rev. Lett. 100 (2008) 084102.
[3] Z. Gao, Q. Cai, Y. Yang, N. Dong, S. Zhang, Visibility graph from adaptive optimal-Kernel time-frequency representation for classification of Epileptiform EEG, Int. J. Neural Syst. (2017) 1750005.
[4] N. Marwan, J. Kurths, P. Saparin, Generalised recurrence plot analysis for spatial data, Phys. Lett. A 360 (2007) 545–551.
[5] M. Wang, L. Tian, R. Du, Research on the interaction patterns among the global crude oil import dependency countries: A complex network approach, Appl. Energy 180 (2016) 779–791.
[6] N. Apergis, B.T. Ewing, J.E. Payne, H. Lund, M.J. Kaiser, A time series analysis of oil production, rig count and crude oil price: evidence from six U.S. oil producing regions, Energy 97 (2016) 339–349.
[7] A. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
[8] M.E.J. Newman, D.J. Watts, Renormalization group analysis of the small-world network model, Phys. Lett. A 263 (1999) 341–346.
[9] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440–442.
[10] Z. Gao, M. Small, J. Kurths, Complex network analysis of time series, Europhys. Lett. 116 (2016) 50001.
[11] L. Lacasa, On the degree distribution of horizontal visibility graphs associated to markov processes and dynamical systems: diagrammatic and variational approaches, Nonlinearity 27 (2014) 2063–2093.
[12] B. Luque, L. Lacasa, F. Ballesteros, J. Luque, Horizontal visibility graphs: exact results for random time series, Phys. Rev. E: Stat. Nonlinear Soft Matter Phys. 80 (2009) 046103.
[13] Z. Gao, Q. Cai, Y. Yang, W. Dang, S. Zhang, Multiscale limited penetrable horizontal visibility graph for analyzing nonlinear time series, Sci. Rep. 6 (2016) 35622.
[14] J. Tang, F. Liu, W. Zhang, S. Zhang, Y. Wang, Exploring dynamic property of traffic flow time series in multi-states based on complex networks: Phase space reconstruction versus visibility graph, Physica A 450 (2016) 635–648.
[15] M.C. Qian, Z.Q. Jiang, W.X. Zhou, Universal and nonuniversal allometric scaling behaviors in the visibility graphs of world stock market indices, J. Phys. A Math. Theoret. 43 (2009) 161–165.
[16] X. Xu, J. Zhang, M. Small, Superfamily phenomena and motifs of networks induced from time series, Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 19601–19605.
[17] Z. Gao, N. Jin, W. Wang, Y. Lai, Motif distributions in phase-space networks for characterizing experimental two-phase flow patterns with chaotic features, Phys. Rev. E: Stat. Nonlinear Soft Matter Phys. 82 (2010) 016210.
[18] M. Wang, L. Tian, K.A. Dawson, J.O. Indekeu, H.E. Stanley, C. Tsallis, From time series to complex networks: the phase space coarse graining, Physica A 461 (2016) 456–468.
[19] J. Zhang, M. Small, Complex network from pseudoperiodic time series: topology versus dynamics, Phys. Rev. Lett. 96 (2006) 238701.
[20] X. Gao, H. An, F. Wei, X. Huang, H. Li, W. Zhong, Characteristics of the transmission of autoregressive sub-patterns in financial time series, Sci. Rep. 4 (2014) 6290.
[21] M. Wang, Y. Chen, L. Tian, S. Jiang, Z. Tian, R. Du, Fluctuation behavior analysis of international crude oil and gasoline price based on complex network perspective, Appl. Energy 2016 (2016) 109–127.
[22] M.A. Kramer, U.T. Eden, S.S. Cash, E.D. Kolaczyk, Network inference with confidence from multivariate time series, Phys. Rev. E: Stat. Nonlinear Soft Matter Phys. 79 (2009) 061916.
[23] M. Jachan, K. Henschel, J. Nawrath, A. Schad, J. Timmer, B.R. Schelter, Inferring direct directed-information flow from multivariate nonlinear time series, Phys. Rev. E: Stat. Nonlinear Soft Matter Phys. 80 (2009) 011138.
[24] T. Nakamura, T. Tanizawa, M. Small, Constructing networks from a dynamical system perspective for multivariate nonlinear time series, Phys. Rev. E: Stat. Nonlinear Soft Matter Phys. 93 (2016) 032323.
[25] L. Lacasa, V. Nicosia, V. Latora, Network structure of multivariate time series, Sci. Rep. 5 (2015) 15508.
[26] X. Huang, H. An, X. Gao, X. Hao, P. Liu, Multiresolution transmission of the correlation modes between bivariate time series based on complex network theory, Physica A 428 (2015) 493–506.
[27] X. Gao, H. An, W. Fang, X. Huang, H. Li, W. Zhong, Y. Ding, Transmission of linear regression patterns between time series: from relationship in time series to complex networks, Phys. Rev. E: Stat. Nonlinear Soft Matter Phys. 90 (2014) 012818.
[28] Z. Gao, P. Fang, M. Ding, N. Jin, Multivariate weighted complex network analysis for characterizing nonlinear dynamic behavior in two-phase flow, Exp. Therm. Fluid Sci. 60 (2015) 157–164.
[29] C. Yao, J. Lin, X. Zheng, X. Liu, The study of RMB exchange rate complex networks based on fluctuation mode, Physica A 436 (2015) 359–376.
[30] R. Du, Y. Wang, G. Dong, L. Tian, Y. Liu, M. Wang, G. Fang, A complex network perspective on interrelations and evolution features of international oil trade, 2002–2013, Appl. Energy 196 (2017) 142–151.
[31] F. Novotný, The link between the brent crude oil price and the US dollar exchange rate, Prague Econ. Pap. 21 (2012) 220–232.
[32] R. Bondia, S. Ghosh, K. Kanjilal, International crude oil prices and the stock prices of clean energy and technology companies: evidence from non-linear cointegration tests with unknown structural breaks, Energy 101 (2016) 558–565.
[33] N.B. Behmiri, J.R.P. Mansu, Crude oil conservation policy hypothesis in OECD countries: a multivariate panel granger causality test, Energy 43 (2012) 213–260.
[34] F. Lu, H. Qiao, S. Wang, K.K. Lai, Y. Li, Time-varying coefficient vector autoregressions model based on dynamic correlation with an application to crude oil and stock markets, Environ. Res. 152 (2017) 351–359.
[35] C.S. Daw, C.E.A. Finney, E.R. Tracy, A review of symbolic analysis of experimental data, Rev. Sci. Instrum. 74 (2003) 915–930.
[36] A. Ray, Symbolic dynamic analysis of complex systems for anomaly detection, Signal Process. 84 (2004) 1115–1130.
[37] T. Bollerslev, A.J. Patton, R. Quaedvlieg, Exploiting the errors: A simple approach for improved volatility forecasting, Creates Res. Pap. 192 (2015) 1–18.
[38] J. Ponton, P. Wei, D. Sun, Weighted clustering coefficient maximization for air transportation networks, in: Control Conference, Vol. 415, 2013, pp. 866–871.
[39] B. Mandelbrot, The variation of certain speculative prices, J. Bus. 36 (1963) 394–394.
[40] J.J. Tseng, S.P. Li, Quantifying volatility clustering in financial time series, Int. Rev. Financ. Anal. 23 (2012) 11–19.
[41] J.J. Tseng, S.P. Li, Asset returns and volatility clustering in financial time series, Physica A 390 (2011) 1300–1314.
[42] H. Niu, J. Wang, Volatility clustering and long memory of financial time series and financial price model, Digit. Signal Process. 23 (2013) 489–498.