Dynamic Financial Contagion Prediction Model Based on Fuzzy Information Granularity SVM

Lin Liu*, Yingfeng Shao**, and Xiaofeng Hui*
* School of Management, Harbin Institute of Technology, Harbin, China
** State Key Laboratory of Nonlinear Mechanics (LNM), Institute of Mechanics, Chinese Academy of Sciences, Beijing, China
lllin1135@sina.com
shaoyf@lnm.imech.ac.cn
xfhui@hit.edu.cn

Abstract—**Contagion time prediction is an important research topic in financial crises. This article put forward a prediction model of contagion time based on fuzzy information granularity SVM. It uses granularity fuzzy and SVM to estimate the bounds of stock index, and further forecast the similarity index. The predicted contagion time from the United States to the United Kingdom, Germany, Frence and China are tested, and compared with the real ones. The empirical analyses comfirm that the model is a feasible method to predict the financial contagion arrival time.**

Keywords—**contagion arrival time; support vector machine; fuzzy information granularit; nonlinear similarity; financial crisis**

I. Introduction

Since the 1990s, the international financial market fluctuated intensified, and the financial crisis broke out frequently. Mexican peso crisis in 1994, Southeast Asia financial crisis in 1997, Russian Flu in 1998, and Brazil crisis in 1999, all destroyed the economy and caused the amazing losses. In the 21st century, the financial crisis expanded furtherly. The global financial crisis in 2007-2009 triggered by the US subprime crisis impacted the whole world. The European debt crisis in 2010 also made people alert. It is of great significance for the healthy development of economy to study financial contagion.

During the last two dacades, many scholars have proved the existence of contagion. Rodriguez[1] took Copula approach to study the Asian crisis and the Mexican crisis, and found evidence of changing dependence. Chiang et al.[2] utilized a dynamic conditional-correlation model to nine Asian markets from 1990 to 2003, and confirmed a contagion effect. Caramazza et al.[3] proved that financial linkages played a significant role in the spread of the Mexican, Asian, and Russian crises. Gallegati[4] investigated contagion during the 2007 US subprime crisis, which results indicated that all sample markets have been affected by the US subprime crisis. Kenourgios et al.[5] studied financial contagion in a multivariate time-varying asymmetric framework, and confirmed a contagion effect from the crisis country to all others.

Time factor is an important to financial contagion. Longstaff[6], Dooley and Hutchison[7] did research on the transmission of the U.S. subprime crisis, based on the division of crisis stages. Madaleno and Pinho[8] pointed out a time delay of the contagion. Kleimeier et al.[9] emphasized time-aligned data in measuring financial contagion. However, there is little research on the contagion arrival time in financial crises.

Bae et al. [10] pointed out that contagion is predictable. Didier et al.[11] also presented that it was impossible to predict whether financial crises or contagion would reappear. In this paper, we bring forward a financial contagion prediction model, which is based on fuzzy information granularity SVM. We exploy the model to forecast the arrival time of contagion from the US to four other countries during the global financial crisis. The experiments show that the model is a feasible one for prediction, and can help to establish early warning in the financial contagion.

II. Methodology

Support Vector Machine (SVM) in this study plays an important role in two steps: LS-SVM helps to forecast for the source, while the fuzzy information granularity SVM, which combines with fuzzy granularity, is used to predict the stock index intervals for the infected countries.

The nonlinear similarity index is the criteria of whether the financial contagion has arrived. Hence, the nonlinear similarity index has to be calculated by the real stock index value and by the predictive one, respectively.

A. *Modeling flow of the dynamic contagion prediction model*

The modeling flow of the proposed method in this paper is shown as Fig. 1.

Step 1 Data normalization: The financial data collected from different market are preprocessed and normalized.

Step 2 Phase space reconstruction: Reconstruct the phase space by the C-C method[12] for the source of contagion. Calculation for the reconstruction parameters.

Step 3 Forecasting for the source: Apply the LS-SVM algorithm based on the phase space to forecast the stock index value for the source.

Step 4 Forecasting for the infected country: Apply the fuzzy information granularity SVM algorithm to forecast two bounds of the stock index for the infected country.

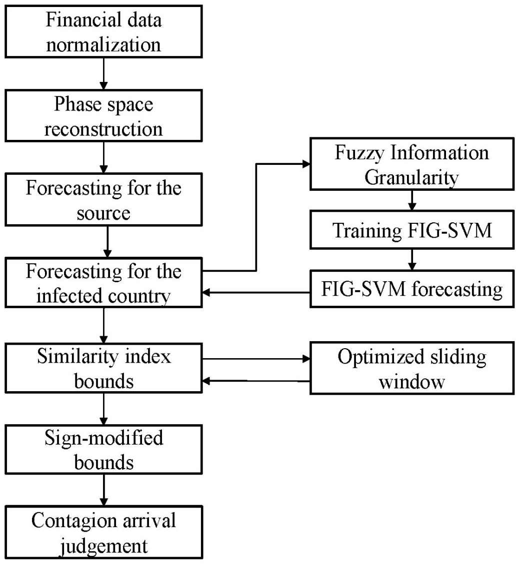


Figure 1. Modeling flow of the dynamic contagion prediction model

Step 5 Similarity index interval: Choose the optimized sliding window. Calculate the similar index interval between the source and the infected country.

Step 6 Sign-modified predictive similarity index interval: Sign the predictive similarity index according to the trend of the stock index of the source.

Step 7 Contagion arrival judgement: Judge the contagion arrival time on the basis of the Sign-modified similarity index.

B. Forecasting for the Source

First, apply the C-C method to calculate the parameters of phase space reconstruction. For nonlinear time series $x=\{x_i \mid i=1,2,\cdots,N\}$, estimate the delay time window τ_w, the optimal delay time τ_d, and the embedding dimension m. The phase space reconstructed can be expressed as

$$X=\left\{X_i \mid X_i=\left[x_i,x_{i+\tau},\cdots,x_{i+(m-1)\tau}\right]^T,\ i=1,2,\cdots,N_1\right\} \quad (1)$$

where $N_1=N-(m-1)\tau$. N_1 represents the number of embedded points in the phase space.

Then, we set that W represents the length of the sliding window. The phase space in the sliding window should be

$$X=\begin{bmatrix} x_i & x_{i+1} & \cdots & x_{i-1+n} \\ x_{i+\tau} & x_{i+1+\tau} & \cdots & x_{i-1+n+\tau} \\ \vdots & \vdots & \vdots & \vdots \\ x_{i+(m-1)\tau} & x_{i+1+(m-1)\tau} & \cdots & x_{i-1+W} \end{bmatrix} \quad (2)$$

where $n=W-(m-1)\tau$ is the number of embedded points in a sliding window.

SVM is a machine learning method to solve data classification, nonlinear pattern recognition and function estimation. LS-SVM [13], which results in a set of linear equations instead of a quadratic programming problem, is an improvement of the traditional SVM. LS-SVM can be applied to high dimension problems.

The basic algorithm to forecast the stock index for the source is stated as follows.

Input: nonlinear time series $x=\{x_i \mid i=1,2,\cdots,N\}$, τ, m

Output: $x'=\{x'_{i-1+W} \mid i=1,2,\cdots,N+1-W\}$

Algorithm: Predict_the_source(x, τ, m, W)

(1) for $i=1:N-W+1$

Set the training sample

Input: $X=\begin{bmatrix} x_i & x_{i+1} & \cdots & x_{i-2+n} \\ x_{i+\tau} & x_{i+1+\tau} & \cdots & x_{i-2+n+\tau} \\ \vdots & \vdots & \vdots & \vdots \\ x_{i+(m-2)\tau} & x_{i+1+(m-2)\tau} & \cdots & x_{i-2+W-\tau} \end{bmatrix}$

Output: $Y=\begin{pmatrix} x_{i+(m-1)\tau} & x_{i+1+(m-1)\tau} & \cdots & x_{i-2+W} \end{pmatrix}^T$

(2) Set the forecasting sample

Input: $X=\begin{pmatrix} x_{i-1+n} & x_{i-1+n+\tau} & \cdots & x_{i-1+W+\tau} \end{pmatrix}^T$

Output: $Y=x'_{i-1+W}$

(3) Take the training sample input and output into

$$y(x)=\sum_{k=1}^{n-1}\alpha_k K(x_k,x)+b \quad (3)$$

to gain the support vector α and the bias b, where $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_{n-1})^T$ denotes the Lagrange multiplier vector, and $K(x_k,x)=\exp\left(\frac{-\|x-x_k\|^2}{\sigma^2}\right)$ is a $(n-1)\times(n-1)$ kernel matrix;

(4) Take the forecasting sample input $\begin{pmatrix} x_{i-1+n} & x_{i-1+n+\tau} & \cdots & x_{i-1+W+\tau} \end{pmatrix}^T$ into the regression function (3), to calculate the output x'_{i-1+W};

end for

(5) Return $x'=\{x'_{i-1+W} \mid i=1,2,\cdots,N+1-W\}$

end

The result x'_{i-1+W} is the prediction of the stock index for the source.

C. Forecasting for the Infected Country

The aim of information granularity[14-15] is to split the object into its basic operational chunks. In information granularity method, the problem is divided into a series of granules. An information granule[16] is viewed as an collection of objects drawn together by the criteria of indistinguishability, similarity or functionality.

The common fuzzy granules are in the forms of triangular fuzzy, trapezoid fuzzy, Gaussian fuzzy and so on. In this paper, we employ the triangular fuzzy granule. The membership function of a triangular fuzzy granule can be expressed as

$$A(x,a_1,m_1,a_2)=\begin{cases}0, x<a_1\\ \dfrac{x-a_1}{m_1-a_1}, a_1\le x\le m_1\\ \dfrac{a_2-x}{a_2-m_1}, m_1<x\le a_2\\ 0, x>a_2\end{cases} \tag{4}$$

where a_1 and a_2 are the upper and lower bounds, and m_1 is the median of fuzzy granule.

The algorithm to predict the stock index upper and lower bounds for an infected country by FIG-SVM is stated as follows:

***Input**:* $x=\{x_i|\ i=1,2,\cdots,N\}$, the fuzzy set length f

***Output**:* $x^{up}=\{x^{up}_{i+W}|i=1,2,\cdots,N-W\}$

***Algorithm**:* Predict_the_infected_upper(x, f, W)

(1) for $i=1:N-W$
Set the training sample
Input: $X=1:\dfrac{W}{f}$
Output: $Y=\{x_j|\mathrm{j}=i, i+1, \cdots, i-1+W\}$

(2) Set the forecasting sample
Input: $X=\dfrac{W}{f}+1$
Output: $Y=x^{up}_{i+W}$

(3) Take the training sample input and output into Formula (4) to obtain the uper bound for every training granule.

(4) Use the training input and the upper bound to train the regression function

$$y(x)=(w^Tx)+b \tag{5}$$

where w is the weight, and b is the bias;

(5) Take the forecasting sample input $X=\dfrac{W}{f}+1$ into the regression function (5), and calculate the output $Y=x^{up}_{i+W}$;

end for

(6) Return $x^{up}=\{x^{up}_{i+W}|i=1,2,\cdots,N-W\}$

end

The returned result x^{up}_{i+W} is the predictive stock-index upper bound of the infected country. In a similar way, we calculate the predictive lower bound of the infected country x^{low}_{i+W}. The predictive stock index $x''_{i+W}\in[x^{low}_{i+W}, x^{up}_{i+W}]$.

FIG-SVM for time series is a prediction based on information granules, hence the prediction result is a granule. In other words, the prediction is an interval with upper and lower bounds.

D. Similarity Index Interval

The dynamic similarity theory of nonlinear time series is based on the following ideas: in science and engineering, we can study systems through the characteristic time evolution of observable properties. For different dynamical systems, or different states of a single system, by analyzing the similarities of the observable time series, we can obtain the variable characteristics in different systems or different stages of one system.

The Real Similarity Index

For two different stock index time series $x=\{x_i|i=1,2,\cdots,N\}$, $y=\{y_i|i=1,2,\cdots,N\}$, reconstruct phase space for the two countries by τ and m of the source, and get two phase spaces

$$X=\{X_j|X_j=[x_{i-1+j}, x_{i-1+j+\tau}, \cdots, x_{i-1+j+(m-1)\tau}]^T, j=1,2,\cdots,n\},$$

$$Y=\{Y_j|Y_j=[y_{i-1+j}, y_{i-1+j+\tau}, \cdots, y_{i-1+j+(m-1)\tau}]^T, j=1,2,\cdots,n\},$$

where $n=W-(m-1)\tau$

The cross-correlation sum of X, Y is

$$C_{XY}=\frac{1}{n_1\times n_2}\sum_{j=1}^{n_1}\sum_{k=1}^{n_2}\Theta\left(\varepsilon-\|X_j-Y_k\|\right) \tag{6}$$

where n_1, n_2 are the numbers of points in phase space X, Y respectively, and ε is a threshold. $\|\cdot\|$ denotes Euclidean distance of two embedded points. $\Theta(\cdot)$ is Heaviside function.

The self-correlation sum of X is

$$C_{XX}=\frac{1}{n_1^2}\sum_{j=1}^{n_1}\sum_{k=1}^{n_1}\Theta\left(\varepsilon-\|X_j-X_k\|\right) \tag{7}$$

Calculate the similarity index[17] of the two countries

$$\gamma_i=\frac{C_{XY}}{\sqrt{C_{XX}\times C_{YY}}} \tag{8}$$

The Predictive Similarity Index Interval

For the sliding window $x=\{x_j|j=i, i+1, \cdots, i-1+W\}$, $y=\{y_j|j=i, i+1, \cdots, i-1+W\}$, we can calculate the predictive similarity index upper bound by replacing x_{i-1+W} with the predictive value x'_{i-1+W} and y_{i-1+W} with the predictive upper bound y^{up}_{i-1+W}. Other steps are the same with the algorithm of the real similarity index.

And correspondingly, we calculate the predictive similarity index lower bound by replacing x_{i-1+W} with the predictive value x'_{i-1+W} and y_{i-1+W} with the predictive lower bound y^{low}_{i-1+W}.

Thus we gain one real similarity index curve and two predictive similarity index curves for each infected country to the source.

E. Sign-modified Similarity Index

To distinguish the financial crisis and common prosperity between two countries, we give the nonlinear similarity index γ a sign. Here we assign γ plus when financial crisis appears, that is when financial time series declines. On the contrary, we assign γ minus when common prosperity come out, that is when financial time series rises.

Thus, the similarity index γ ranges from -1 to 1.

III. Empirical Experiment

A. Data Preprocessing

In this paper, the sample period is from Jan 4th, 2007 to May 30th, 2012. The daily closing prices of five markets are used, including DJI of the United States, FTSE100 of the United Kingdom, DAX of Germany, FCHI of French, and SSI of China. Data were taken from Bloomberg.

Since the opening dates of different countries are not the same with each other, we adjust the time series by the piecewise linear interpolation method. After interpolation, there are N_1=1406 observations. We give the descriptive statistics after piecewise linear interpolation in Table I.

TABLE I.
Descriptive Statistics of five Stock Market Indices

	DJI	FTSE100	DAX	FCHI	SSI
Obs.	1406	1406	1406	1406	1406
Max	14164.53	6732.40	8105.69	6168.15	6092.06
Min	6547.05	3512.10	3666.41	2519.29	1706.70
Mean	11310.91	5515.12	6320.65	4082.06	3045.82
Std	1684.77	713.03	991.53	941.28	885.40

All the data in the time series are preprocessed and normalized to [0,1].

B. Empirical Results

As it is known to all that the US is the source of the globle financial crisis, we use the US's parameters to reconstruct the phase space. The sampling interval τ_s=1. we can obtain the optimal delay time τ_d=30 and the delay time window τ_w=9. Therefore, the embedding dimension m=5. We get N_0=1370 embedded points for each market.

We apply the LS-SVM regression to forecast the next day stock index of the source. The prediction value is closely related to the sliding window length W. The relationship of W and the number of the embedded points in the window is shown in Table II.

TABLE II.
Relationship of sliding window length and number of points

W	40	50	60	70	80
n	4	14	24	34	44

In this experiment, W=60 is the most appropriate parameter for the sliding window length. Therefore, there are n=24 embedded points in every sliding window. We forecast for the source. The result is 1347 prediction value for the US.

Similarly, we forecast for the infected country. The results are a pair of upper and lower bounds. In a financial crisis, the linkage between the source and the infected country is a result of contagion. As FIG-SVM forecasts data for a changing interval, it reflects more accurately than the traditional SVM for the infected country. Define the fuzzy granule length f=5. Taking FIG-SVM prediction for the UK, Germany, French and China.

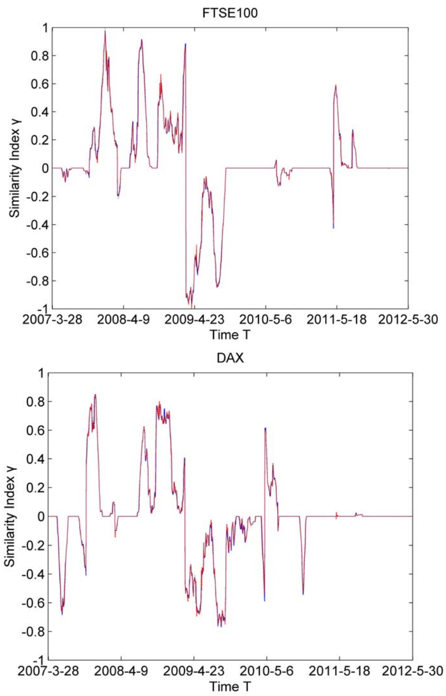
The threshold ε[18] is defined as 30% of the cumulative neighborhood distribution of the US in the same time period. Calculate the predictive similarity indices and the real ones for each infected country. There are 1347 real similarity indices and 1347 pairs of predictive ones, which means 1347 predictive similarity index intervals.

The sign of similarity index γ is decided by the trend of DJI. In the experimental period, the extreme points after smoothing are shown in Table III.

TABLE III. Date of Extreme points of the US

Maximum	Minimum
2007-10-09	2008-03-07
2008-05-05	2009-03-09
2010-04-26	2010-07-02
2011-04-29	2011-05-17
2012-04-02	

According to the points in Table III, we devide the sample into 10 periods. Assign the similarity index γ plus when DJI declines, γ minus when DJI rises. The similarity indices, including one real similarity index and two predictive ones from 2007-2012 for the four infected countries, are shown as Fig. 2.



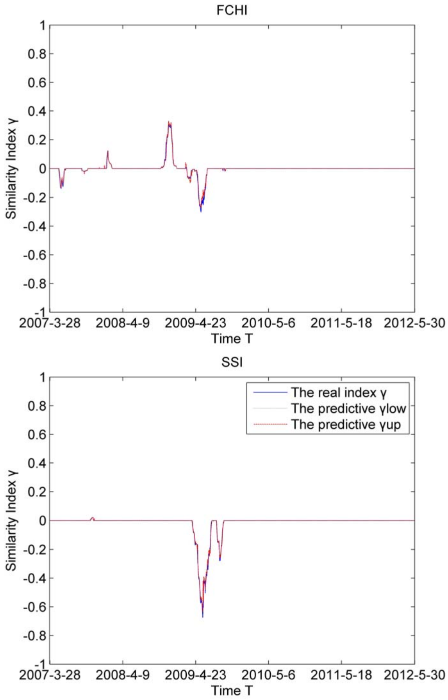


Figure 2. Real and predictive similarity indices to the US

We focus on the positive similarity index which represents the financial crisis periods. When the threshold γ_0 selects different values, the contagion arrival time of the UK and China is shown in Table IV. γ is the real similarity index, while γ_{pre} is the predictive similarity index.

TABLE IV. CONTAGION ARRIVAL TIME UNDER DIFFERENT THRESHOLDS

		FTSE100		SSI	
		γ	γ_{pre}	γ	γ_{pre}
1	0.1	2007-10-10	2007-10-10		
	0.8	2007-12-24	2007-12-24		
2	0.1	2008-05-05	2008-05-05		
	0.8	2008-06-26	2008-06-27		
	0.1	2008-10-06	2008-10-07		
	0.8	2009-02-23	2009-02-24		
3	0.1				
	0.8				
4	0.1	2011-05-18	2011-05-18		
	0.8				
5	0.1				
	0.8				

When the threshold γ_0 selects different values, the prediction accuracy P is shown in Table V. γ corresponds to the number of dates of the real similarity index $\geq\gamma_0$, while γ_{pre} corresponds to the number of dates of the predictive similarity index $\geq\gamma_0$, which selects the earlier of γ_{low} and γ_{up}.

TABLE V. PREDICTION PERFORMANCE OF THE PREDICTIVE SIMILARITY INDEX MODEL

γ_0		FTSE100	DAX	FCHI	SSI
0.05	γ	333	264	40	0
	γ_{pre}	328	263	39	0
	P	98.50%	99.62%	97.5%	
0.1	γ	303	233	27	0
	γ_{pre}	302	230	24	0
	P	99.70%	98.71%	88.89%	
0.2	γ	236	185	18	0
	γ_{pre}	232	183	19	0
	P	98.31%	98.92%	94.44%	
0.3	γ	161	152	7	0
	γ_{pre}	159	149	8	0
	P	98.76%	98.03%	85.71%	
0.5	γ	90	111	0	0
	γ_{pre}	91	108	0	0
	P	98.89%	97.30%		
0.8	γ	29	7	0	0
	γ_{pre}	29	7	0	0
	P	100%	100%		

C. *Analysis and Discussion*

It can be seen from Fig. 2 that the financial markets of the UK, Germany and Frence performed obviously positive linkage with that of the US, while Chinese market has a positive linkage with that of the US but is not distinct. Financial contagion[19] is defined as a significant increase in cross-market linkages after a shock to other countries. Therefore, it implies the former three countries had strong financial contagion with the US, and the contagion to China is much weaker.

There are five decline of the US DJI in the sample. Fig. 2 also shows that in the first decline from 2007-10-09 to 2008-03-07, the contagion impacts all the four countries, and the impacted to the UK and Germany is severe. In the second decline from 2008-05-05 to 2009-03-09, the severe contagion extends to Frence, besides the UK and Germany. In the third decline, the contagion influences Germany, and in the fourth decline, it influences the UK. The fifth decline does not affect any country, so it is not a contagion.

Table V shows that the contagion to the UK is stronger than to Germany, than to Frence, than to China. P near to 1 means that the predictive similarity index interval is close to the real similarity index, and further means the prediction is relatively accurate.

We can see from Fig. 2 and Table V that the main contagion from the US to other countries is from October 2007 to March 2009. This is also the major outbreak period of the global financial crisis triggered by the US subprime crisis. The contagion from the US to China was in the end of 2007, and much shorter than other three countries. It is consistent with other researchers'[20, 21] opinions.

IV. Conclusion

The prediction of contagion arrival time in financial crisis is important and necessary. The traditional researches did not develop the arrival time prediction. In this paper, we construct a new model combined SVM and fuzzy information granularity. Using the nonlinear similarity index as the measurement criteria, this method can dynamically forecast the arrival of contagion. Empirical experiment based on the globle financial crisis triggered by the US is carried out test the model. The experimental results and analysis show, the prediction model works well in forecast the arrival time of contagion. It can also be used to describe the contagion degree. The model proves that it is feasible to do contagion prediction. However, it still has some localization: the prediction period is short. It should be improved by following work.

References

[1] J. C. Rodriguez, "Measuring financial contagion: A copula approach," *Journal of Empirical Finance,* vol. 14, pp. 401–423, June 2007.

[2] T. C. Chiang, B. N. Jeon, and H. Li, "Dynamic correlation analysis of financial contagion: Evidence from Asian markets," *Journal of International Money and Finance,* vol. 26, pp. 1206–1228, November 2007.

[3] F. Caramazza, L. Ricci, and R. Salgado, "International financial contagion in currency crises," *Journal of International Money and Finance,* vol. 23, pp. 51–70, February 2004.

[4] M. Gallegati, "A wavelet-based approach to test for financial market contagion," *Computational Statistics & Data Analysis,* vol. 56, pp. 3491–3497, November 2012.

[5] D. Kenourgios, A. Samitas, and N. Paltalidis, "Financial crises and stock market contagion in a multivariate time-varying asymmetric framework," *Journal of International Financial Markets, Institutions and Money,* vol. 21, pp. 92–106, February 2011.

[6] F. A. Longstaff, "The subprime credit crisis and contagion in financial markets," *Journal of Financial Economics,* vol. 97, pp. 436–450, September 2010.

[7] M. Dooley and M. Hutchison, "Transmission of the US subprime crisis to emerging markets: Evidence on the decoupling–recoupling hypothesis," *Journal of International Money and Finance,* vol. 28, pp. 1331–1349, December 2009.

[8] M. Madaleno and C. Pinho, "International stock market indices comovements: a new look," *International Journal of Finance & Economics,* vol. 17, pp. 89–102, January 2012.

[9] S. Kleimeier, T. Lehnert, and W. F. Verschoor, "Measuring Financial Contagion Using Time - Aligned Data: The Importance of the Speed of Transmission of Shocks," *Oxford Bulletin of Economics and Statistics,* vol. 70, pp. 493–508, August 2008.

[10] K. H. Bae, G. A. Karolyi, and R. M. Stulz, "A new approach to measuring financial contagion," *Review of Financial studies,* vol. 16, pp. 717–763, 2003.

[11] T. Didier, P. Mauro, and S. L. Schmukler, "Vanishing financial contagion?," *Journal of Policy Modeling,* vol. 30, pp. 775–791, September 2008.

[12] H. S. Kim, R. Eykholt, andJ. Salas, "Nonlinear dynamics, delay times, and embedding windows," *Physica D: Nonlinear Phenomena,* vol. 127, pp. 48–60, March 1999.

[13] J. A. Suykens, J. Vandewalle, and B. De Moor, "Optimal control by least squares support vector machines," *Neural Networks,* vol. 14, pp. 23–35, January 2001.

[14] L. A. Zadeh, "The role of fuzzy logic in the management of uncertainty in expert systems," *Fuzzy sets and Systems,* vol. 11, pp. 197–198, 1983.

[15] L. A. Zadeh, "Fuzzy logic= computing with words," *Fuzzy Systems, IEEE Transactions on,* vol. 4, pp. 103–111, May 1996.

[16] Y. Q. Zhang, "Granular Neural Network," in *Encyclopedia of Complexity and Systems Science*. Springer, 2009, pp. 4402-4411.

[17] T. Schreiber and A. Schmitz, "Classification of time series data with nonlinear similarity measures," *Physical review letters,* vol. 79, pp. 1475–1478, August 1997.

[18] M. L. V. Quyen, J. Martinerie, V. Navarro, P. Boon, M. D'Havé, C. Adam, B. Renault, F. Varela, and M. Baulac, "Anticipation of epileptic seizures from standard EEG recordings," *The Lancet,* vol. 357, pp. 183–88, January 2001.

[19] K. J. Forbes, R. Rigobon, "No Contagion, Only Interdependence: Measuring Stock Markets Comovements," *The Journal of Finance* vol. 5, pp. 2223–2261, October 2002.

[20] S. S. Kang and N. N. Song, "In the United States, Britain and other countries to address the financial crisis, fiscal and monetary policy comparison," *Wuhan Finance,* vol. 11, pp. 22–24, August 2010.

[21] F. X. Huang, L. Kou and D. L. Yang, "Before and after the financial crisis in Anglo-American stock market volatility spillovers between comparative," *Journal of Applied Statistics and Management,* vol. 4, pp. 751–754, September 2012.