Financial Crisis Forecasting via Coupled Market State Analysis

Wei Cao and Longbing Cao, *University of Technology, Sydney*

Coupled market state analysis assumes market observations are governed by a collection of intra- and intercoupled hidden market states.

The impact of a financial crisis is often disruptive from multiple perspectives, including the local economy, daily life, society in general, and globalization. For example, the subprime mortgage crisis that started in the US in 2007 triggered a chain of destructive effects on instruments in the US and global financial markets, before moving on to other sectors.

But effectively detecting a possible financial crisis isn't a trivial task: a financial crisis is a rare and complex problem, and it's hard to determine which observed indicators are more sensitive than others. Moreover, financial markets are always coupled with each other, so any crisis has a strong transfer effect as it moves from one market to another.[1]

Many researchers and practitioners have recognized the need for and challenges of crisis forecasting, with most current approaches enabling prediction through direct observations of indicators. As shown in Figure 1a, for example, a crisis detector (such as a logistic regression classifier) forecasts a crisis for time $t + 1$ based on observations of all market indicators from time $t - k + 1$ to t, where k is the time window size. We argue that such approaches overlook the underlying coupled relationships between markets that can't be directly detected. When these couplings change, they affect direct observations.

To represent the hidden couplings between various markets, we propose a new forecasting framework based on coupled market state analysis (CMSA), where coupled market state (CMS) refers to a set of dynamic hidden states that represent the transition induced by the constant interactions between markets. Our proposal works on the assumption that market indicators are governed by a collection of CMSs, which are better features for capturing a financial crisis.

Accordingly, we insert a coupled state-space model (CSSM) between observations and the crisis detector, as Figure 1b shows, to conduct a CMSA over all markets. Our approach comprises the following steps: learn the CMSs behind the observations using the CSSM, feed the CMSs into a detector as features, and have the detector forecast whether a financial crisis exists. In so doing, we avoid the data vulnerability found in traditional observation-based approaches. Empirical evaluations on real-world financial data demonstrate the superiority of our approach compared to observation-based methods.

1541-1672/15/$31.00 © 2015 IEEE
Published by the IEEE Computer Society

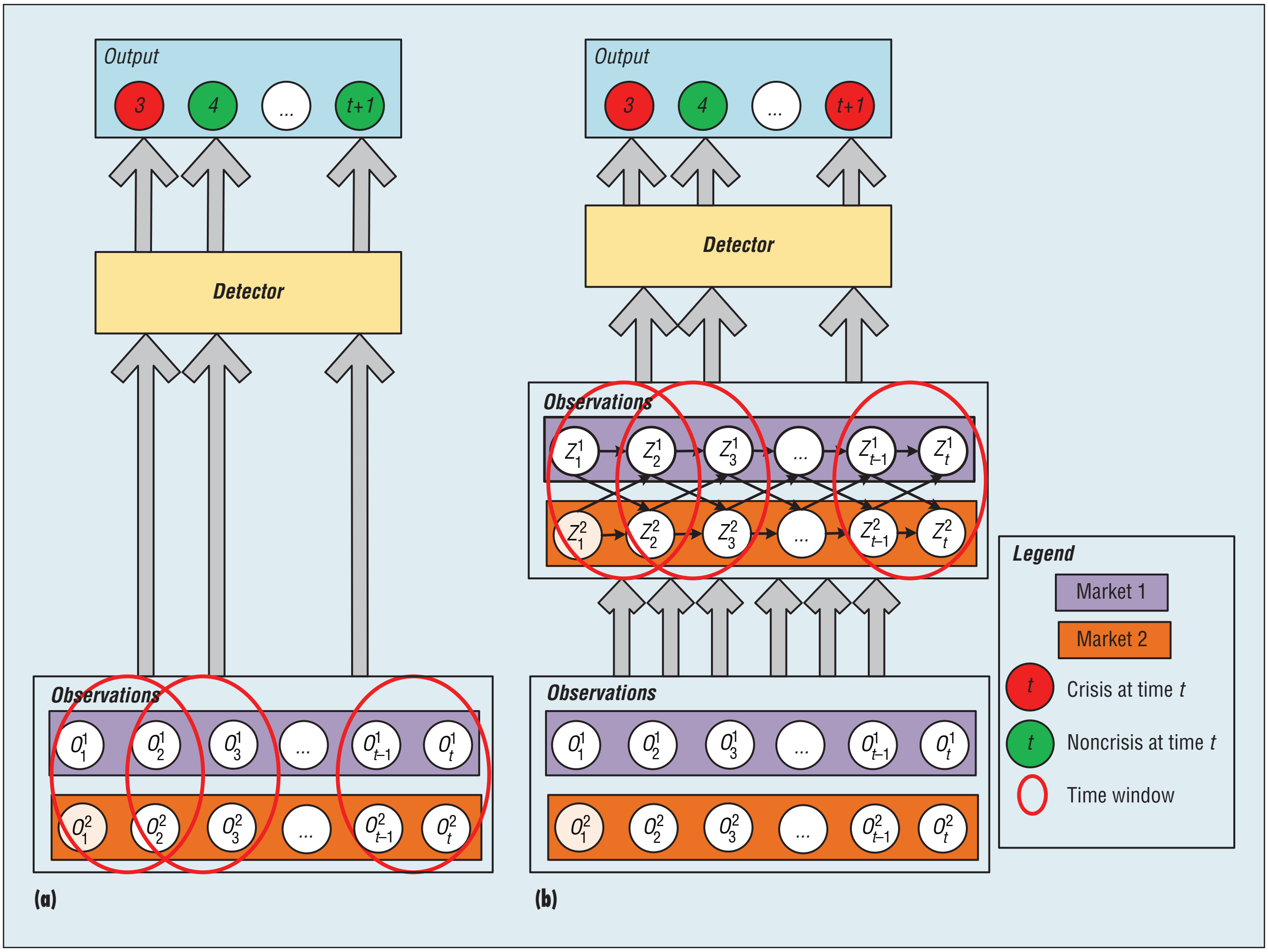

Figure 1. Two types of forecasting: (a) observation based and (b) coupled market state analysis (CMSA) based. The latter forecasts include the underlying coupled relationships between markets that can't be directly detected.

Background

The approaches related to financial crisis forecasting can be roughly categorized into the following three types. The *signal* approach[2] uses financial variables (such as the exchange rate or stock market index) as indicators to identify the difference in economic behaviors between the crisis and normal periods. The main drawback here is that the approach relies on the selection of variables,[3] often resulting in biased results. The second type is *linear time-series analysis*, represented by logistic and probit models[4] that use historical data to infer a future crisis. However, this kind of approach suffers from a linear assumption, whereas most financial crises exhibit nonlinear behaviors.[1] The last type uses *machine learning*, usually through an artificial neural network (ANN).[5] This kind of approach[6] pays more attention to the nonlinear correlations of variables, but many models predict financial crises directly from observations of variables, completely ignoring the underlying complex coupled relations.[7] Consequently, the outcomes might be too sensitive to observations. To avoid such vulnerability, our model uses CMSs behind the observations. Before introducing it, let's first go over our state-space model (SSM) and coupled hidden Markov model (CHMM), which serve as the foundation of this article.

An SSM refers to a class of probabilistic graphical models that describes the probabilistic dependence between the latent variable and the observed measurement.[8] The basic state-space equations are as follows:

$$Y_t = H_t U_t + v_t \quad (1)$$

$$U_t = F_t U_{t-1} + w_t, \quad (2)$$

where Y_t is the observation vector, U_t is the state vector, H_t links the state vector to the observations, F_t is a state transition matrix, and v_t and w_t are the control vectors.

The CHMM was proposed to model multiple processes with coupling relationships.[9] It consists of

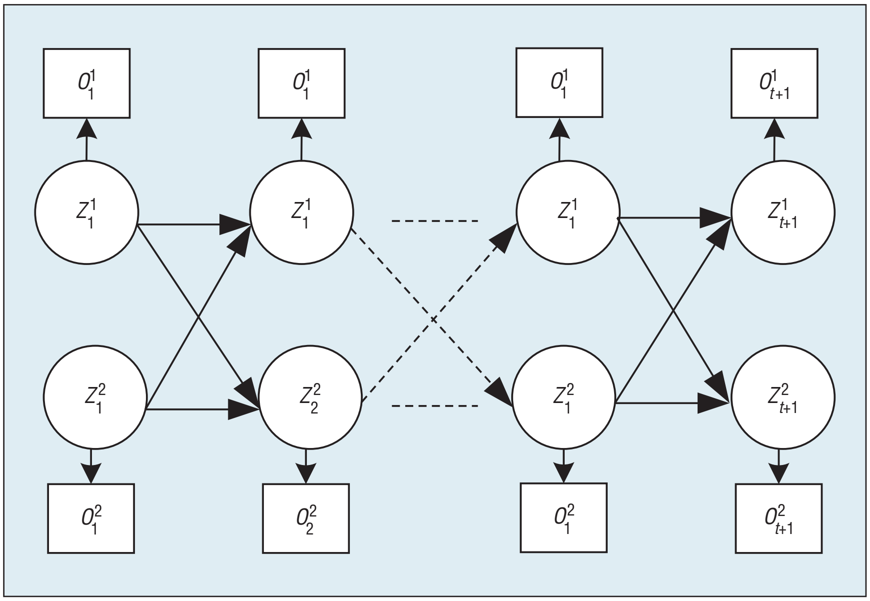

Figure 2. A coupled hidden Markov model (CHMM) with two coupled chains. Here, O_t^1 represents observation (a market indicator) of market 1 at time *t*, and Z_t^1 denotes the corresponding hidden state of market 1 at time *t*.

more than one chain of HMMs representing different processes in which the state of any chain of HMM at time t depends on not only the states of its own chain but also the states of other chains at time $t - 1$—namely, the *interaction* between the processes. Figure 2 is a standard CHMM with two chains.

Suppose there are C CHMMs, H is the number of states of the Markov chains, and $\{Z_1, Z_2, \ldots, Z_H\}$ is a set of hidden states, where z_t is the hidden state at time t. V is the number of observation symbols, $\{X_1, X_2, \ldots, X_V\}$ is a set of observation symbols, $\{O_1, O_2, \ldots, O_T\}$ is an observation sequence, and o_t is the observation at time t. Given this, we can define the corresponding elements of a CHMM as follows:

Prior probability of initial state

$$\pi = \left\{\pi_j^{(c)}\right\}$$

$$\pi_j^{(c)} = p\left(z_1^{(c)} = Z_j\right), \quad s.t. \sum\nolimits_{j=1}^{H^{(c)}} \pi_j^{(c)} = 1.$$

State transition probability matrix

$$\mathbf{A} = \left\{a_{j'j}^{(c',c)}\right\}$$

$$a_{j'j}^{(c',c)} = p\left(z_{t+1}^{c} = Z_j \mid z_t^{c'} = Z_{j'}\right),$$

$$s.t. \sum\nolimits_{j=1}^{H^{(c)}} a_{j'j}^{(c',c)} = 1.$$

Observation probability matrix

$$b_j^{(c)}(v) = p\left(o_t^{(c)} = X_v \mid z_t = Z_j\right),$$

$$s.t. \sum\nolimits_{v=1}^{V} b_j^{(c)}(v) = 1.$$

Coupling coefficient $R = r_{c',c}\ 1 \le c', c \le C$

$$s.t. \sum\nolimits_{c'=1}^{C} r_{c',c} = 1.$$

A Case Study

The 2008 global financial crisis showed that linkages exist in different financial markets. We can verify these coupling relations quantitatively by using typical market indexes of three major financial markets (commodity market: the price of gold and crude oil futures; equity market: the S&P 500 index and Dow Jones Industrial Average (DJIA); and interest market: the TED and Baa spread), as shown in Figure 3 and Table 1.

The data for Figure 3 and Table 1 spans January 2006 to December 2012, including the 2008 global crisis period. We can see from Figure 3 that the relations between the three markets are complex: the couplings are stable in the noncrisis periods (before late 2007 and after 2009) but fluctuate sharply during the financial crisis interval (from 2007 to 2009). Table 1 uses the Pearson correlations to describe the relations between the indicators in the three markets, and we can see that the coefficients are very significant, namely, that there are strong correlations between the indexes. On the basis of this, we can conclude that there exist some hidden couplings between the three markets that can't observe directly. In addition, the hidden couplings behave differently during and around financial crisis periods.

We can formalize the forecasting problem as follows: $f(\cdot)$ is a function used to capture the complex coupled relationships between market states in different financial markets, and an objective function $g(\cdot)$ is built to forecast the possibilities of crisis and noncrisis. If at time t,

$$g_{t+1}^{D}(c=1) \left| f_{t-k+1}^{t}(CMS) \ge g_{t+1}^{D}(c=0) \right| f_{t-k+1}^{t}(CMS), \tag{3}$$

then time $t + 1$ is a crisis period; otherwise, it's a noncrisis period. Here, $g_{t+1}^{D}(c=1) \mid f_{t-k+1}^{t}(CMS)$ represents the possibility of a crisis by using a detector with the CMSs from time $t - k + 1$ to time t (k denotes the window size). The key task of crisis forecasting, then, is to build a proper model to determine the specific function $f(\cdot)$ and the corresponding objective function $g(\cdot)$.

Modeling Framework

On the basis of the case study and corresponding problem formalization, we propose a CMSA-based

Table 1. Correlations between indicators in three types of markets.

Indicators	Metrics	S&P 500	Dow Jones Industrial Average (DJIA)	Gold	Oil futures	TED	Baa
S&P 500	Pearson correlation sig. (2-tailed)	1	.976** .000	.010 .843	.425** .000	–.092 .081	–.739** .000
DJIA	Pearson correlation sig. (2-tailed)	.976** .000	1	.194** .000	.517** .000	–.100 .057	–.635** .000
Gold	Pearson correlation sig. (2-tailed)	.010 .843	.194** .000	1	.470** .000	–.383** .000	.248** .000
Oil futures	Pearson correlation sig. (2-tailed)	.425** .000	.517** .000	.470** .000	1	.033 .534	–.108* .041
TED	Pearson correlation sig. (2-tailed)	–.092 .081	–.100 .057	–.383** .000	.033 .534	1	.384** .000
Baa	Pearson correlation sig. (2-tailed)	–.739** .000	–.635** .000	.248** .000	–.108* .041	.384** .000	1

* Correlation is significant at the 0.05 level (2-tailed); ** Correlation is significant at the 0.01 level (2-tailed).

financial crisis forecasting framework as shown in Figure 4a. It comprises three major steps: CMSA, mapping, and forecasting.

Coupled Market State Analysis

CMS refers to the hidden states from multiple markets with inter- and intracouplings. Suppose there are I markets $\{M_1, M_2, \ldots, M_I\}$, and a market M_i undertakes J market states $\{\mathbb{MS}_{i1}, \mathbb{MS}_{i2}, \ldots, \mathbb{MS}_{iJ}\}$. A market state behavior feature matrix $FM(\mathbb{MS})$ is then defined as follows:

$$FM(\mathrm{MS}) = \begin{pmatrix} \mathrm{MS}_{11} & \mathrm{MS}_{12} & \cdots & \mathrm{MS}_{1J} \\ \mathrm{MS}_{21} & \mathrm{MS}_{22} & \cdots & \mathrm{MS}_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{MS}_{I1} & \mathrm{MS}_{I2} & \cdots & \mathrm{MS}_{IJ} \end{pmatrix}.$$

The intracouplings within each market state are the relationships within one row of this matrix, whereas how the states interact between different markets is embodied through the columns of $FM(\mathbb{MS})$, indicated as intercouplings.

Definition 1. CMS refers to market states $\mathbb{MS}_{i1j1}$ and $\mathbb{MS}_{i2j2}$, which are coupled in terms of intra- and intercoupling:

$$f_{Intra}(\mathrm{MS}) = \left\{\mathrm{MS}_{i_1j_1} \odot \mathrm{MS}_{i_2j_2} | i_1 = i_2, 1 \le i_1, i_2 \le I\right\} \quad (4)$$

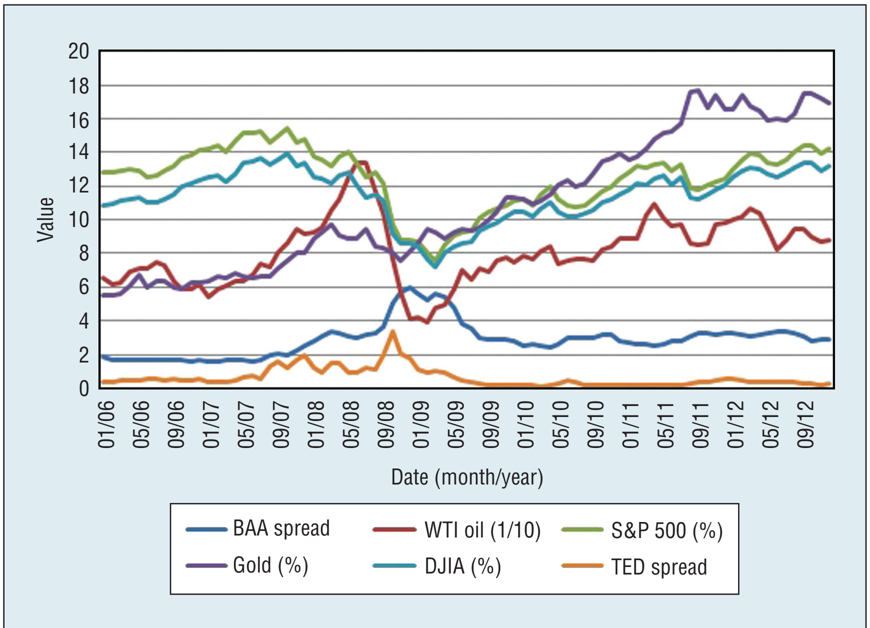

Figure 3. Indexes series in three markets. The relations between the three markets are complex: the couplings are stable in the noncrisis periods (before late 2007 and after 2009) but fluctuate sharply during the financial crisis interval (from 2007 to 2009).

$$f_{Inter}(\mathrm{MS}) = \left\{\mathrm{MS}_{i_1j_1} \odot \mathrm{MS}_{i_2j_2} | i_1 \ne i_2, 1 \le i_1, i_2 \le I\right\}, \quad (5)$$

where $\odot$ means the interactions of $\mathrm{MS}_{i_1j_1}$ and $\mathrm{MS}_{i_2j_2}$.

Definition 2. CMSA builds the objective function $g(\cdot)$ under the condition that the hidden market states are coupled with each other by coupling function $f(\cdot)$ and satisfy the following conditions:

$$f(\cdot) ::= \{f_{Intra}(\mathbb{MS}), f_{Inter}(\mathbb{MS})\} \quad (6)$$

$$argmax_c g(c) \mid f(\cdot), \quad (7)$$

where $c \in \{0,1\}$, where 0 represents the noncrisis set, and 1 denotes the crisis set.

Definition 3. CSSM refers to a type of graphic model that not only

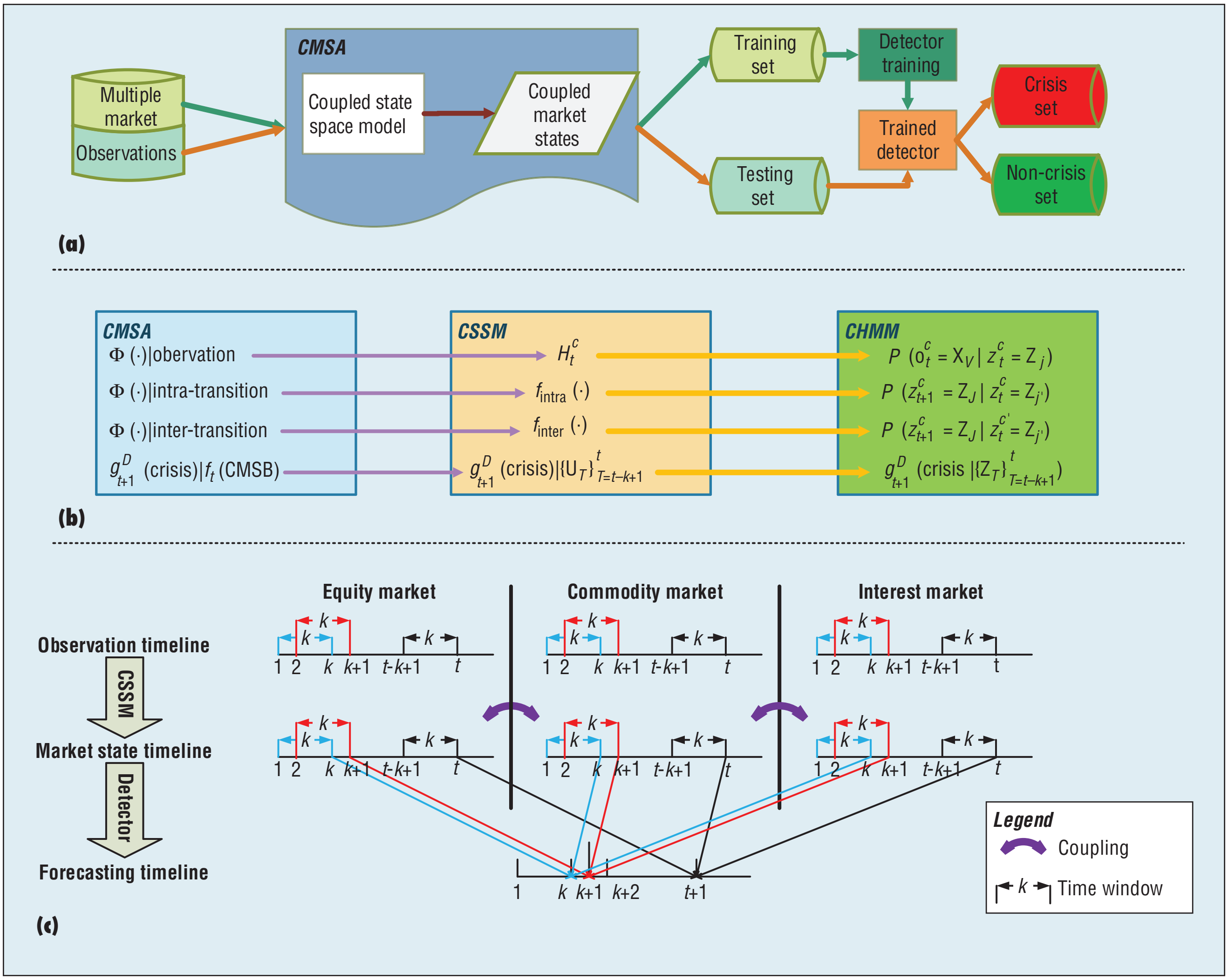

Figure 4. CMSA overview. (a) Our proposed financial crisis forecasting framework, (b) the mapping process, and (c) the forecasting process.

describes the latent variable and the observed measurement but also captures the coupled relationships between latent variables across different SSMs. Suppose there are C SSMs; the basic state-space equations are then

$$Y_t^c = H_t U_t^c + v_t \tag{8}$$

$$U_t^c = F\left(f_{intra}\left(U_{t-1}^c\right), f_{inter}\left(U_{t-1}^{c'}\right)\right), \tag{9}$$

where $U_t = [U_t^1, U_t^2, \ldots, U_t^C]$. Here, $(1 \le c', c \le C) \wedge (c \ne c')$, F is a transition function, and $f_{intra}(\cdot)$ is built to capture the intra relations in the same SSM, whereas $f_{inter}(\cdot)$ encodes the couplings between different SSMs.

We use CHMM as a concrete implementation of CSSM to learn coupled states between markets because CHMM is a probabilistic state-space model to capture the nonlinear coupling relationship in multiple processes and the transitional effect from one hidden state to another.[10]

Modeling Process

In CHMM, we use one Markov chain to represent one financial market, so in this article, we select one indicator for each market that has higher correlations with other markets:

$$CI_{i_1 j_1} = \sum_{i_2=1}^{m} \sum_{j_2=1}^{n} corr\left(I_{i_1 j_1}, I_{i_2 j_2}\right). \tag{10}$$

The indicator correlation $CI_{i_1 j_1}$ refers to the correlations of indicator $I_{i_1 j_1}$ with indicators in other markets $I_{i_2 j_2}$, where m is the number of markets, and each market owns n indicators $(i_1 \ne i_2) \wedge (1 \le i_1, i_2 \le m) \wedge (1 \le j_1, j_2 \le n)$. Here, $corr(\cdot)$ is the Pearson correlation coefficient of the two indicators.

There are two mapping processes: one from CMSA to CSSM, namely, from a formalized issue to an abstracting model, and the other from CSSM to CHMM, which resolves the issue with a specific toolCHMM.

As we mentioned earlier, there are three market state sequences: Φ(E) encloses the equity market state sequence, and Φ(C) and Φ(I) represent commodity and interest market state sequences, separately. $\{Z_1, Z_2, \ldots, Z_H\}$ is a set of hidden states, where z_t is the hidden state at time t. $\{X_1, X_2, \ldots, X_V\}$ is a set of observation symbols, $O = \{O_1, O_2, \ldots, O_T\}$ is an observation sequence, and o_t is the observation at time t. Figure 4b illustrates the specific mapping relations.

Table 2. Accuracy of various approaches. Boldface indicates the best performance with each k.

Approach	Accuracy		
	$k = 2$	$k = 3$	$k = 4$
Signal-crisis	0.5604		
Logistic-crisis	0.6926	0.6953	0.7137
Artificial neural network (ANN)-crisis	0.7471	0.7617	0.7294
Coupled hidden Markov model (CHMM)-logistic	0.8132	0.8320	0.7647
CHMM-ANN	0.8016	0.8281	0.8157

Forecasting Process

Figure 4c shows the general framework of the proposed forecasting process. For each observation interval $O_{t-k+1:t}$ (k is the time window), the first step is to train the CHMM using the k observations ($O_{t-k+1:t}$) in the three markets to obtain corresponding market states $Z_{t-k+1:t}$. Then, on the basis of the coupled states, the trained detector gives the probabilities of crisis $P^D_{(t+1)}(c=1 \mid Z^t_{T(T=t-k+1)})$ and non-crisis $P^D_{(t+1)}(c=0 \mid \{Z_T\}^t_{T=t-k+1})$. After comparing the two probabilities, we can find whether time $t + 1$ is in a financial crisis set.

Evaluation and Discussion

We selected one indicator from each market according to Equation 10: the DJIA, the crude oil price futures, and the Baa spread. The dataset includes weekly closing prices from January 1990 to December 2010, obtained from the Federal Bank of St. Louis (http://research.stlouisfed.org); we decoded the prices into returns as symbols that can be calculated by $RI_t = (PI_t - PI_{t-1}/PI_{t-1} \times 100\%)$, where RI_t and PI_t are, respectively, the return and closing price at time t.

We divided the data into two parts: a training set from January 1990 to December 2006, and a testing set from January 2007 to December 2010. According to the National Bureau of Economic Research (NBER) Business Cycle Dating Committee (www.nber.org/cycles.html), there are two crisis periods in the training dataset (July 1990 to March 1991, led by the Gulf War, and March 2001 to November 2001, triggered by the dot-com bubble), and one crisis period in the testing dataset (December 2007 to June 2009, caused by the subprime mortgage crisis). Because indicators in different markets could appear on different trading days, we deleted those days on which some market data is missing and only chose the days with trading data from all markets. Here are the methods we used:

- *Signal-crisis*. The basic idea of this method is that variables behave differently in a financial crisis period when compared with a normal period.[2] We used it as a baseline method.
- *Logistic-crisis*. We used this approach with indicators from the three different markets; the parameters can be obtained through maximum likelihood estimation.
- *ANN-crisis*. We used a back-propagation algorithm[5] with indicators from the three different markets to train the model.

We based our technical performance evaluation on the following metrics:

$$Accuracy = \frac{TN + TP}{TP + FP + FN + TN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN},$$

where TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative, respectively. We treat the financial crisis cases as the positive class here.

We also compared the technical performance of our approach against other approaches on the testing data with a different window size k, calculating accuracy, precision, and recall. Table 2 and Figure 5 show the results.

Table 2 shows the accuracy performance of the five approaches over the whole testing period. We can see that the baseline method of signal doesn't achieve a good performance; this is because signal relies on the selection of indicators and pays no attention to hidden couplings between indicators. For a similar reason, the logistic and ANN approaches don't perform very well either. Note that ANN outperforms the logistic approach, primarily because the latter is under a linear assumption, but the financial crisis reveals nonlinear characteristics. Our CHMM-based approaches have better performance than the comparative methods with all window sizes. For instance, the CHMM-logistic has about 14 percent improvement over the logistic approach when the time window equals 3, and CHMM-ANN has a roughly 9 percent gain over the ANN method when the time window equals 4. The main reason can be interpreted as follows: unlike those methods that predict financial crisis directly from data, our approach uses a framework of CMSA to learn the hidden CMSs over different markets, which removes

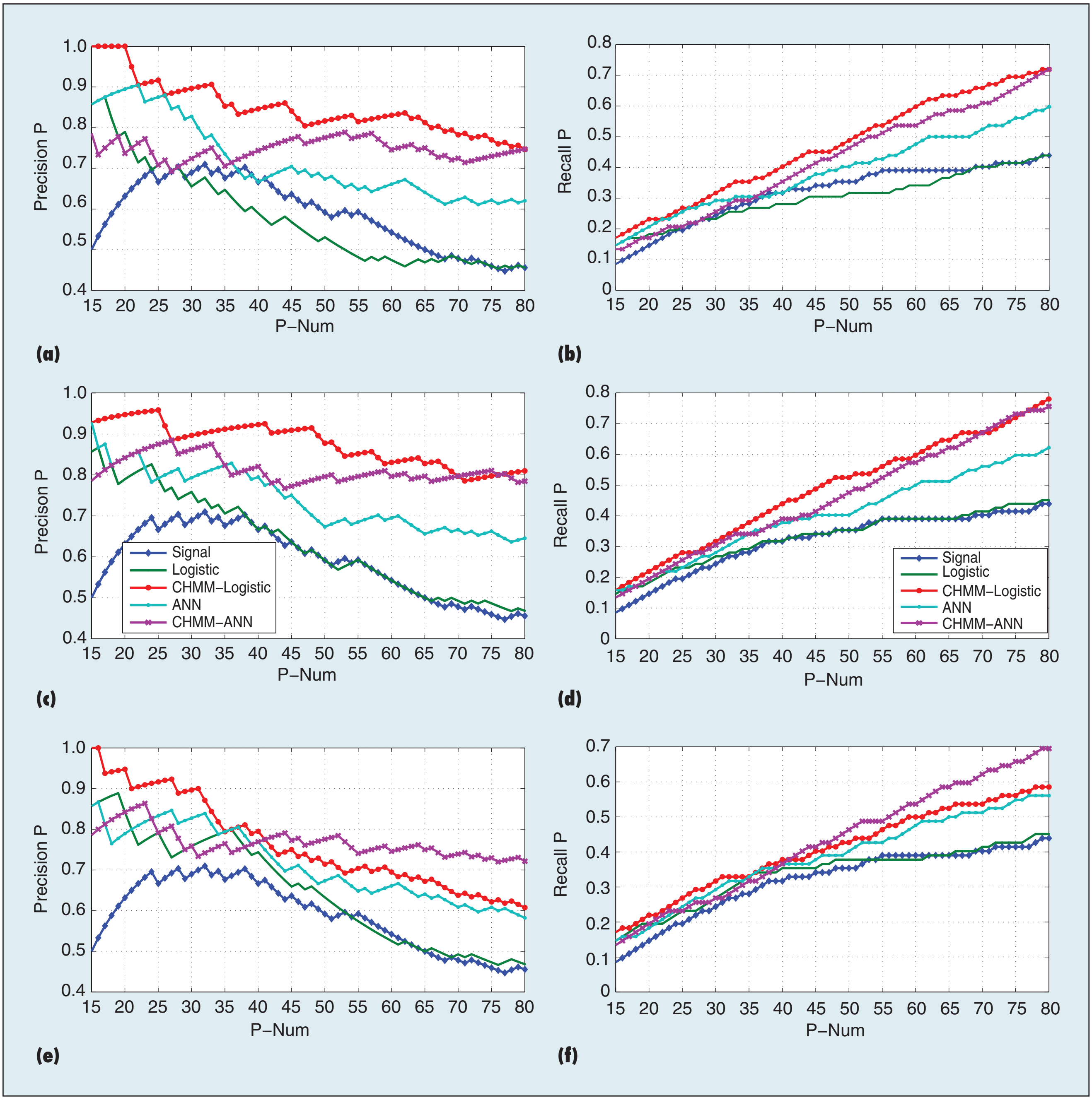

Figure 5. Technical performance of various approaches on predicting a crisis: (a) precision (*k* = 2), (b) recall (*k* = 2), (c) precision (*k* = 3), (d) recall (*k* = 3), (e) precision (*k* = 4), and (f) recall (*k* = 4), where *k* denotes window size. The CHMM-based approaches outperform the other approaches at any window size.

the vulnerabilities in the data; in addition, CHMM has been demonstrated to be a useful model to characterize the CMSs.

Figure 5 shows the technical performance of precision and recall by setting three different window sizes, where the horizontal axis (P-Num) stands for the number of predicted trading weeks in the financial crisis, and the vertical axis represents the values of the technical measures. We can see that the CHMM-based approaches outperform the other approaches at any window size. For instance, in Figure 5c, when P-Num = 45, the precision of CHMM-logistic is 90 percent, which is 40 percent better than logistic. In addition, recall represents the probability that a crisis is retrieved. Figures 5b, 5d, and 5f show the CHMM-based approach achieves higher recall than observation-based models with any P-Nums.

References

1. F.A. Longstaff, "The Subprime Credit Crisis and Contagion in Financial Markets," *J. Financial Economics,* vol. 97, no. 3, 2010, pp. 436–450.
2. D. Peng and C. Bajona, "China's Vulnerability to Currency Crisis: A KLR Signals Approach," *China Economic Rev.*, vol. 19, no. 2, 2008, pp. 138–151.
3. L. Yu et al., "A Multiscale Neural Network Learning Paradigm for Financial Crisis Forecasting," *Neurocomputing*, vol. 73, no. 4, 2010, pp. 716–725.
4. E.K. Laitinen and T. Laitinen, "Bankruptcy Prediction: Application of the Taylor's Expansion in Logistic Regression," *Int'l Rev. Financial Analysis,* vol. 9, no. 4, 2001, pp. 327–349.
5. E.N. Ozkan-Gunay and M. Ozkan, "Prediction of Bank Failures in Emerging Financial Markets: An ANN Approach," *J. Risk Finance,* vol. 8, no. 5, 2007, pp. 465–480.
6. T.Y. Kim et al., "Usefulness of Artificial Neural Networks for Early Warning System of Economic Crisis," *Expert Systems with Applications,* vol. 26, no. 4, 2004, pp. 583–590.
7. L. Cao, "Coupling Learning of Complex Interactions," to appear in *J. Information Processing and Management*, 2015.
8. Z. Chen and E.N. Brown, "State Space Model," *Scholarpedia*, vol. 8, no. 3, 2013, article no. 30868.
9. N.M. Oliver, B. Rosario, and A.P. Pentland, "A Bayesian Computer Vision System for Modeling Human Interactions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, 2000, pp. 831–843.
10. S. Zhong and J. Ghosh, *A New Formulation of Coupled Hidden Markov Models*, tech. report, Dept. Electronic and Computer Eng., Univ. Texas at Austin, 2001; www.shi-zhong.com/research.

THE AUTHORS

Wei Cao is a PhD candidate in the Advanced Analytics Institute at the University of Technology, Sydney. Her research interests include data mining, machine learning, and behavior analysis in finance. Cao has a PhD in information technology from the University of Technology, Sydney. Contact her at caoweivivi@gmail.com.

Longbing Cao is a professor in the AAI department at the University of Technology, Sydney. His research interests include data mining, machine learning and their applications, behavior informatics, open complex intelligent systems, multiagent technology, and agent mining. Cao has one PhD in intelligent sciences from the Chinese Academy of Sciences and another in computing sciences from the University of Technology, Sydney. He's a senior member of IEEE. Contact him at longbing.cao@uts.edu.au.