Information Dynamic Spectrum Predicts Critical Transitions

Kang-Yu Ni and Tsai-Ching Lu

HRL Laboratories, Malibu, CA 90265, USA
{kni,tlu}@hrl.com

Abstract. This paper addresses the need of predicting system instability toward critical transitions occurred in complex systems. A novel information dynamic spectrum framework and a method for automated prediction of system trajectories are proposed. Our framework goes beyond unidirectional diffusion dynamics to investigate heterogeneously networked dynamical systems with transient directional influence dynamics. Our method automatically analyzes the input time series of system instability to predict the instability trajectories toward critical transitions.

Keywords: Transfer entropy, prediction, information dynamics.

1 Introduction

In complex social-technological systems, self-organized emerging interactions provide the benefits of exchanging information and resources effectively, yet at the same time increase the risk and pace of spreading attacks or failures. A small perturbation on a complex system operating in a high-risk unstable region can induce a critical transition that leads to catastrophic failures. In this work, we propose to avoid catastrophic failures by detecting early warnings of critical transitions and predicting the likelihood of system trajectories.

A review on early warning signals for critical transitions, originated in the ecological domain, can be found in [8]. Signals such as increased temporal correlation, skewness, and spatial correlation in population dynamics are used to quantify the phenomena of critical slowing down as early warning indicators of critical transitions. The reviewed methods in [8] have only been applied to homogeneous lattices, not heterogeneously networked dynamical systems. The very recent review in [9] highlights the critical role of heterogeneous network structures in anticipating critical transitions.

There are a few recent works of early warning signals for heterogeneously networked dynamical systems [6,7]. A spectral early warning signals (EWS) theory is developed to detect the approaching of critical transitions and estimate the system structure and network connectivity near critical transitions using the covariance spectrum. Although Spectral EWS quantifies how much entities of a system are moving together, the symmetric nature of covariance spectrum does not permit the analysis of directional influences among entities.

K. Glass et al. (Eds.): COMPLEX 2012, LNICST 126, pp. 267–280, 2013.

There have been attempts in identifying directional influence in complex systems using transfer entropy [10]. In [11], symbolic transfer entropy (STE) is proposed to analyze brain electrical activity data for the detection of the asymmetric dependences, and to identify the hemisphere containing the epileptic focus without observing actual seizure activity. In [4], the transfer entropy matrix is used on financial market data to analyze asymmetrical influence from mature markets to emerging markets. Although global transfer entropy quantified in [4,11] showed promising results in analyzing financial and neurophysiological data, they fail short to quantify local transfer entropy [5] changing in structure and time. Moreover, these papers do not predict system instability.

To detect critical transitions and predict instability trajectories, we propose an information dynamic spectrum framework to quantify directional influences in heterogeneously networked dynamical systems. Our framework is based on a novel Associative Transfer Entropy (ATE) measure which decomposes the pairwise directional influence of transfer entropy to associative states of asymmetric, directional information flows. We transform multivariate time series of complex systems into the spectrum of Transfer Entropy Matrix (TEM) and Associative Transfer Entropy Matrix (ATEM) to capture information dynamics of the system. We develop the novel spectral radius measure of TEM and ATEM to detect early warning signals of source-driven instability, and induce directional influence structure to identify the source and reveal dynamics of directional influences. The nature of convex growth of spectral radius of TEM and ATEM prior to critical transitions enables us to generate the probabilistic light cones of system trajectories using natural logarithmic curve modeling. We demonstrate our methods on (1) early detection and prediction of instability of non-Foster circuit, (2) changes of directional influences in Latin America stock indices during 2008 financial crisis, and (3) asymmetric derivers in Wikipedia editing behaviors.

2 Information Dynamics Spectrum

Given the time series of a system of m elements: $X(t) = [x_1(t), x_2(t), , x_m(t)]^T$, the Transfer Entropy (TE) from source x_j to destination x_i is defined as [10],

$$T_{x_j \to x_i} = \sum_{(x_{i,t+\tau}, x_{i,t}^{(k)}, x_{j,t}^{(\ell)}) \in D} p(x_{i,t+\tau}, x_{i,t}^{(k)}, x_{j,t}^{(\ell)}) \log \frac{p(x_{i,t+\tau} | x_{i,t}^{(k)}, x_{j,t}^{(\ell)})}{p(x_{i,t+\tau} | x_{i,t}^{(k)})}. \quad (1)$$

where D is the set of all possible values of $(x_{i,t+\tau}, x_{i,t}^{(k)}, x_{j,t}^{(\ell)})$. TE quantifies the amount of information transferred from x_j to x_i and is asymmetric. More details can be found in [10].

2.1 Associative Transfer Entropy

The idea of the proposed associative transfer entropy is to decompose TE by constraining associated states of processes. It is often important to distinguish

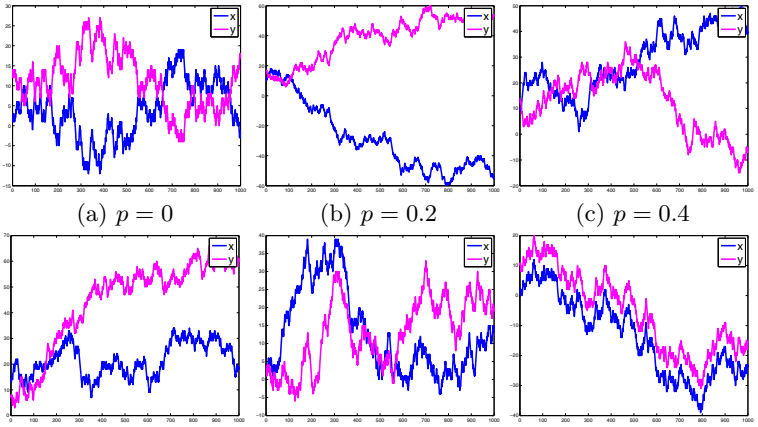
the types of the information flow, in addition to the amount of information flow. We propose a new measure, Associative Transfer Entropy (ATE) at state D_k, for $i \neq j$ defined by:

$$T^{D_k}_{x_j \to x_i} = \sum_{(x_{i,t+\tau}, x^{(k)}_{i,t}, x^{(\ell)}_{j,t}) \in D^{D_k}} p(x_{i,t+\tau}, x^{(k)}_{i,t}, x^{(\ell)}_{j,t}) \log \frac{p(x_{i,t+\tau} | x^{(k)}_{i,t}, x^{(\ell)}_{j,t})}{p(x_{i,t+\tau} | x^{(k)}_{i,t})}, \quad (2)$$

where D_k is a subset of D that represents a certain associated state between x_i and x_j. The purpose of ATE is to capture information transfer between two variables for a particular state association. For the simplest example, we can decompose TE into two influence classes: one positive ATE+ and the other negative ATE−. Therefore ATE in this case is able to distinguish two situations where the amount of information transfer is the same but have opposite effects: one source drives the destination the same direction and the other source drives the destination the opposite direction.

For a simple illustration, we simulate two time series $x(t)$ and $y(t)$ with a binary difference between the current value and the next value, e.g. 1 represents an increment and 0 represents a decrement. The probabilities of increment/decrement are conditioned by the previous increment/decrement. Fig. 1 illustrates a few simulated data with the following setup. Let $\dot{x}(t) = x(t) - x(t-1)$ and $\dot{y}(t) = y(t) - y(t-1)$. We fix conditional probabilities $Pr(\dot{x}(t+1) = a | \dot{x}(t) = b) = 0.5$, $Pr(\dot{y}(t+1) = a | \dot{y}(t) = b) = 0.5$, and $Pr(\dot{x}(t+1) = a | \dot{y}(t) = b) = 0.5$, where a and b are all combinations of 0 and 1. The only bias comes from $Pr(\dot{y}(t+1) = a | \dot{x}(t) = b) = p$. When $p < 0.5$, an increment in x is more likely to cause an decrement in y, and an decrement in x is more likely to cause an increment in y. Fig. 1 (a)-(f) show a few such simulated time series of length 1000 from arbitrary initializations with $p = 0, 0.2, 0.4, 0.6, 1$, respectively.

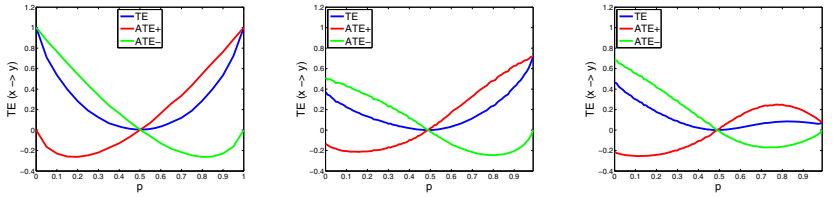
Fig. 2 plots the TE and ATEs as functions of p, for $0 \leq p \leq 1$, averaged over 100 trials for each p. In this binary case, we decompose TE in positive association ATE+ and negative association ATE−, summing over the sets $D_+ = \{(\dot{y}(t+1), \dot{x}(t), \dot{y}(t)) = (0,0,0), (0,0,1), (1,1,0), \text{or} (1,1,1)\}$ and $D_- = \{(\dot{y}(t+1), \dot{x}(t), \dot{y}(t)) = (0,1,0), (0,1,1), (1,0,0), \text{or} (1,0,1)\}$, respectively. ATE+ sums over positive association: $\dot{y}(t+1)$ and $\dot{x}(t)$ are both 0 or both 1. ATE− sums over negative association: one of $\dot{y}(t+1)$ and $\dot{x}(t)$ is 0 and the other is 1. Therefore, as p increases, ATE+ increases and ATE− decreases. One the other hand, TE does not distinguish the types of influence between two time series. One can further carry out the expectation values for ATE+ and ATE− in the case of (a): ATE+ $= p \log_2(p/0.5)$ and ATE− $= (1-p) \log_2((1-p/0.5)$. Near $p = 1$, ATE+ in (c) is much smaller than ATE+ in (b), even though the positive influence from $\dot{x}(t) \to \dot{y}(t+1)$ is high. The explanation of this is that TE/ATE± negate the amount of influence by itself: $\dot{y}(t) \to \dot{y}(t+1)$. The self-influence in (c) is 0.9, much larger than that of (b), which is 0.6. Alternatively, TE/ATE± measures the net influence, which is the external influence subtracted by the internal dynamics.



(a) $p = 0$ (b) $p = 0.2$ (c) $p = 0.4$

(d) $p = 0.6$ (e) $p = 0.8$ (f) $p = 1$

Fig. 1. Examples of simulated data generated from the following probabilities. The conditional probabilities of $Pr(\dot{x}(t+1)|\dot{x}(t))$, $Pr(\dot{y}(t+1)|\dot{y}(t))$, and $Pr(\dot{x}(t+1)|\dot{y}(t))$ are fixed and unbiased. The only bias is from $x'(t) \to \dot{y}(t+1)$: $Pr(\dot{y}(t+1) = 0|\dot{x}(t) = 0) = p$ and $Pr(\dot{y}(t+1) = 1|\dot{x}(t) = 1) = p$. The influence from x to y is negative when $p < 0.5$, as seen in (a-c). Similarly, the influence from x to y is positive when $p > 0.5$, as seen in (d-f).



(a) $\Pr(\dot{y}(t+1)|\dot{y}(t)) = 0.5$ (b) $\Pr(\dot{y}(t+1)|\dot{y}(t)) = 0.6$ (c) $\Pr(\dot{y}(t+1)|\dot{y}(t)) = 0.9$
$\Pr(\dot{x}(t+1)|\dot{x}(t)) = 0.5$ $\Pr(\dot{x}(t+1)|\dot{x}(t)) = 0.7$ $\Pr(\dot{x}(t+1)|\dot{x}(t)) = 0.8$
$\Pr(\dot{x}(t+1)|\dot{y}(t)) = 0.5$ $\Pr(\dot{x}(t+1)|\dot{y}(t)) = 0.5$ $\Pr(\dot{x}(t+1)|\dot{y}(t)) = 0.5$

Fig. 2. ATE is able to distinguish the types of influence in information transfer. The TE, ATE+ and ATE− curves are functions of $p = Pr(\dot{y}(t+1) = 0|\dot{x}(t) = 0) = Pr(\dot{y}(t+1) = 1|\dot{x}(t) = 1)$.

2.2 Local TE/ATE

We next consider dynamic data, by which we mean the amount of information flow from one node to another is not constant. For dynamic data, we calculate the TE/ATE of time series in a local time window, so that TE/ATE becomes a function of time. For illustration, we consider a two-node network, which produces two time series $x(t)$ and $y(t)$, $t = 1, ..., 1000$. Let $\dot{x}(t) = x(t) - x(t-1)$ and $\dot{y}(t) = y(t) - y(t-1)$. Consider the binary case $n = 2$ for $\dot{x}(t)$ and $\dot{y}(t)$, where $\dot{x}(t) = 1$ represents increment and $\dot{x}(t) = 0$ represents decrement. We then simulate the

data according to $Pr(\dot{y}(t+1)=0|\dot{x}(t)=0)=Pr(\dot{y}(t+1)=1|\dot{x}(t)=1)=p_1$ and $Pr(\dot{y}(t+1)=0|\dot{x}(t)=1)=Pr(\dot{y}(t+1)=1|\dot{x}(t)=0)=p_2$. In the simulated data, the probabilities of associative influence change at $t=300$ and 600. For $1\leq t<300$, $p_1=[0.9,1]$ and $p_2=[0,0.2]$; for $300\leq t<600$, $p_1=[0,0.2]$ and $p_2=[0.9,1]$; and for $600\leq t\leq 1000$, $p_1=[0.9,1]$ and $p_2=[0,0.2]$. Therefore, initially x has a strong positive influence, then a strong negative influence during the middle period, finally a strong positive influence again at the end. For simplicity, we fix $Pr(\dot{x}(t+1)=0|\dot{y}(t)=0)=Pr(\dot{x}(t+1)=1|\dot{y}(t)=1)=p_1$ and $Pr(\dot{x}(t+1)=0|\dot{y}(t)=1)=Pr(\dot{x}(t+1)=1|\dot{y}(t)=0)=p_2$. For larger number of states, for example $n=4$, one way to define the increment and decrement is: 0 and 1 represent decrement by 2 and 1, respectively and 2 and 3 represent increment by 1 and 2, respectively. Similarly, for $n=6$ levels, 0, 1, 2, 3, 4, 5 represent -3, -2, -1, 1, 2, 3, respectively. The top row of Fig. 3 shows the simulated data, and the bottom row shows TE/ATE± as functions of time, which are calculated with sliding window size = 100. We can see that ATE+ and ATE− are able to capture the types of influence besides the amount of influence that change over time. Note that for $n>2$, one can decompose ATE into more than two states.

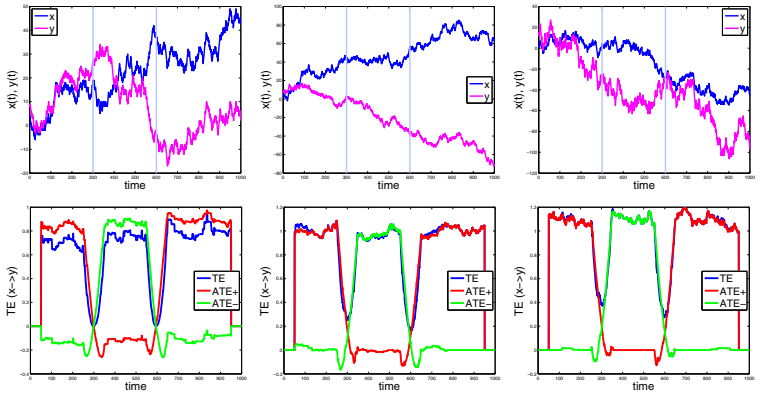


Fig. 3. TE/ATE as functions of time in the dynamic case. $Pr(\dot{x}(t+1)|\dot{x}(t),\dot{y}(t))$ is fixed and unbiased for all cases and $Pr(\dot{y}(t+1)|\dot{y}(t),\dot{x}(t))$ is switched at time $t=300$ and 600, indicated by the blue vertical lines. The number of increment and decrement levels are $n=2,4,6$ for (a), (b), and (c), respectively.

2.3 Spectral Radius of TEM/ATEM

The ATE Matrix (ATEM) T^{D_k} of a system of m elements is an $m\times m$ matrix with ij^{th} entry $(T^{D_k})_{ij}=T^{D_k}_{x_i\to x_j}$. Similarly, the ij^{th} entry of the $m\times m$ TE Matrix (TEM) is $D_{x_i\to x_j}$. TEM has been used in [4] to reveal the asymmetric influences from mature markets to emerging markets. For dynamic data, we calculate pairwise local TE/ATEs to form TEM/ATEM. Since TE/ATE is directional, TEM/ATEM is non-symmetric. To identify the amount of information transfer in the system, we calculate the spectral radius of the TEM/ATEM, the largest absolute eigenvalue of the matrix.

Since the TEM matrix is nonsymmetric, its eigenvalues are complex-valued. We use the spectral radius of the TEM matrix to measure the total amount of information flow of the entire network. Fig. 4 shows spectral radius of TEM as a function of time for each simulated data. The data is simulated according to the following pitchfork bifurcation equation:

$$x_t = \tanh(ct - 10)x - x^3 + \alpha \triangle x + \sigma d\omega, \tag{3}$$

where $\triangle$ represents graph Laplacian. Fig. 4 shows a few canonical structures of graph: (a) chain, (b) directed chain, and directed binary trees, (c) downward and (d) upward. The adjacency matrix A and Laplacian matrix L of (d) are:

$$A = \begin{bmatrix} 0&1&1&0&0&0&0 \\ 0&0&0&1&1&0&0 \\ 0&0&0&0&0&1&1 \\ 0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0 \end{bmatrix} \text{ and } L = \begin{bmatrix} 2&-1&-1&0&0&0&0 \\ 0&2&0&-1&-1&0&0 \\ 0&0&2&0&0&-1&-1 \\ 0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0 \end{bmatrix}, \tag{4}$$

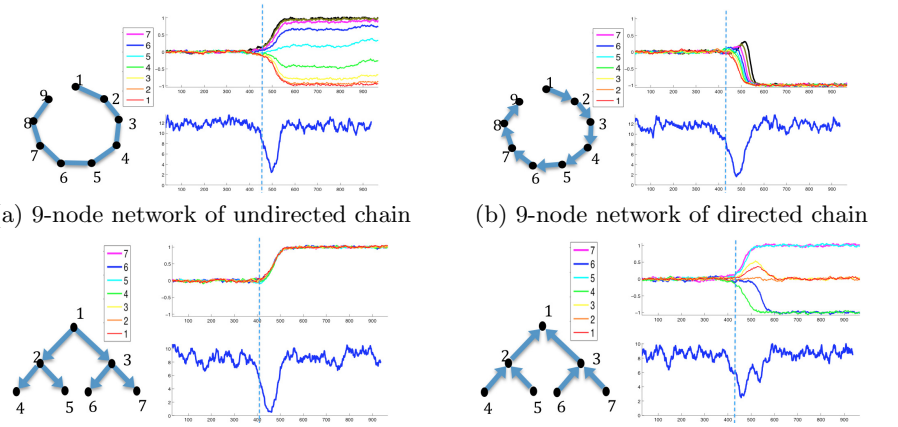
respectively. In the adjacency matrix, $a_{ij} = 1$ means there is a connection from node$j$ to node i, and $a_{ij} = 0$ means there is not a connection from node j to node i. The Laplacian matrix $L = D_{\text{in}} - A$, where D_{in} is the in-degree matrix whose diagonal entry d_{ii} is the sum of i^{th} row of A, the number of connections going into node i. Fig. 4 plots the simulated time series and their spectral radius of the TEM. We observe that before transitioning or bifurcation, the spectral radius decreases rapidly, which provides early indication of system transitioning. The spectral radius of TEM in these cases drops to the lowest point during transitioning, due to the strong internal dynamics within each node, and TE measures the pair-wise influence dynamics between nodes.

To efficiently estimate transfer entropy of continuous data, the method in [11] uses the symbolization technique for permutation entropy [1]. This method to estimate transfer entropy is robust and computationally fast. We adapt the symbolization technique to calculate ATE. Specifically, the continuous-valued time sequence $\{x(t)\}_{t=1}^{N}$ is symbolized by first ordering the values of $\{x(t), x(t+1), x(t+m)\}$ with $1, 2, , m$ and denoting $\hat{x}(t) =$ associated permutation of order m the symbol of $x(t)$ at t. Then we estimate the ATE of $\{x(t)\}_{t=1}^{N}$ by calculating ATE of $\{\hat{x}(t)\}_{t=1}^{N}$.

3 Numerical Results of Information Dynamic Spectrum

3.1 ATE Early Indication of Instability of non-Foster Circuit Data

Fig. 5 shows an ATE analysis of a non-Foster network [12]. The circuit is initially operated in the stable region, where there is no oscillation. Then a small perturbation is added. It is unknown whether the circuit will become oscillatory



(a) 9-node network of undirected chain (b) 9-node network of directed chain

(c) 7-node network of downward binary tree (d) 7-node network of upward binary tree

Fig. 4. Simulated pitchfork bifurcations with difference canonical graph structures. The spectral radius of TEM decreases before bifurcation, which provides an early indication for phase transition.

(unstable) or stay stable. We perform an ATE analysis to detect if the circuit will become synchronized (unstable). The top plot of Fig. 5 is the circuit in voltage over time. The bottom plot shows the TE/ATE± curves over time. The curves are obtained from absolute sum of spectrum of TEM/ATEM±, respectively. We found that the spectral radius of TEM/ATEM± also show similar outcome, but since TEM is the sum of ATEM+ and ATEM−, the ATE+ curve will never cross the TE curve. Thus, the absolute sum of the spectrum is more informative. As shown in the plot, the ATE+ curve crosses over the TE curve near 800, indicating the increasing in synchronization. In addition, the ATE+ curve reaches peak right before full synchronization, while the ATE− curve flattens because there is no negative association.

3.2 TEMs Infer Directional Influences in Latin America Stock Indices

We use TEMs of different periods to analyze the dynamics of the stocks indices around a critical event and discover the directional structures in different periods. Fig. 6 shows an analysis of a 9-node network in which each node represents a Latin America stock market index. The indices are detailed in Table 1. The top row of Fig. 6 shows the TEMs before, during, and after the 2008 October Crash from left to right. The red color corresponds to a large value, while blue corresponds to a small value. During crash, the total amount of information transfer decreases. After the crash, the TEM came back to TEM before the crash. Visualization of the network structure shows that Panama is strongly influenced by Columbia and Brazil before and after the crash, yet during the crash Panama is primarily driven by Mexico and Venezuela.

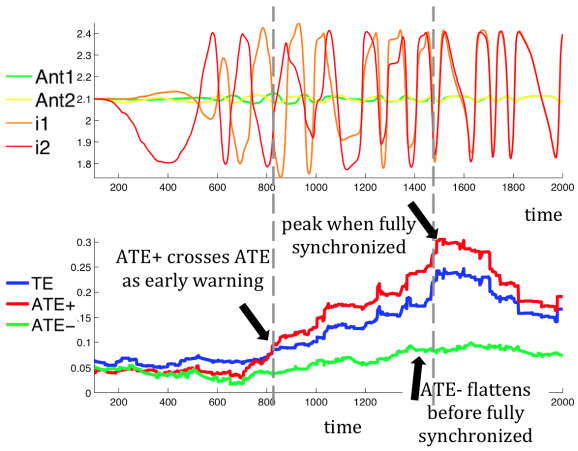


Fig. 5. The system takes multivariate time series of observed system behavior as inputs, and outputs the warnings and trends toward critical transitions with sources and paths of propagations

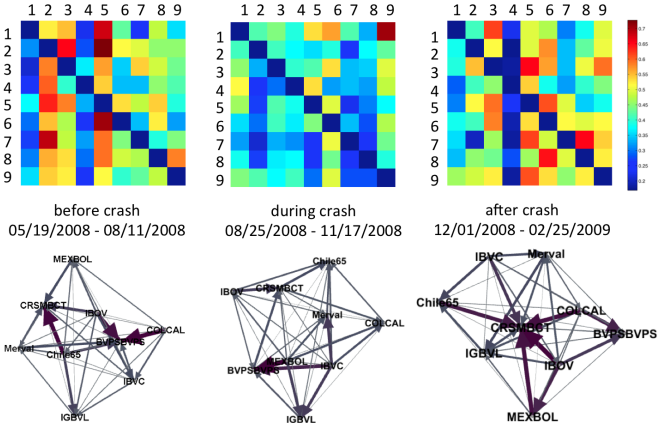


Fig. 6. TEMs infer directional influences in Latin America stock indices before, during, and after the 2008 October Crash

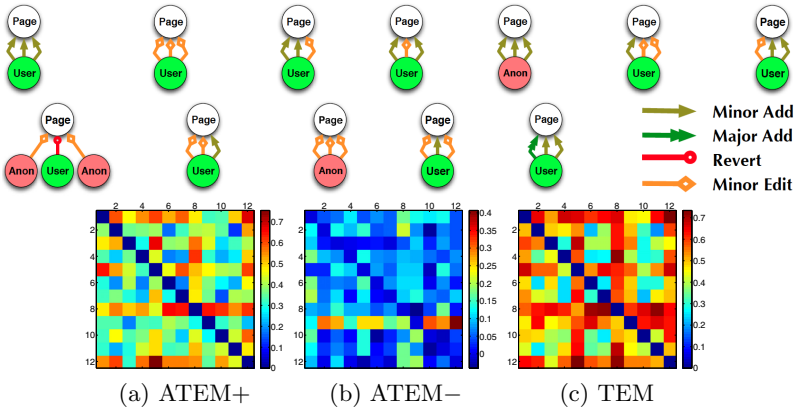
3.3 ATE Reveals Wikipedia Motifs That Drive the Changes

The dynamics of editor behaviors of Wikipedia's content is explored in [3] using temporal motifs, which are temporal bipartite graphs with multiple node and edge types for users and revisions. The first two rows or Fig. 7 shows 12 most frequent motifs. For example, in the second row, the first motif from the left, shows a minor edit of a Wikipedia page by an anonymous author, followed by a revert from a registered user, then followed by a minor edit from an anonymous author. The bottom row from left to right are global ATEM+, ATEM−, and TEM 2001 to 2011. We see that by using ATEM+ and ATEM−, we can identify

Table 1. Latin American stock market indices

1	2	3	4	5	6	7	8	9
BVPSBVPS	Chile65	COLCAP	CRSMBCT	IBOV	IBVC	IGBVL	Merval	MEXBOL
Panama	Chile	Columbia	Costa Rica	Brazil	Venzuela	Peru	Argentina	Mexico

important motifs that drive the changes in other motifs. In particular, ATEM+ shows that motifs 1, 8 and 12 have the most positive influence on other motifs, observing that "major add", "revert", and consecutive minor add by registered users encourage Wikipedia's content growth. ATEM− shows that motif 9 has the most negative influence on other motifs. TEM shows the asymmetric influence among the motifs but the contrast is not as strong as ATEM+ and ATEM−. Fig. 8 shows the TE/ATE± curves, spectral radii of local TEM/ATEM±, in blue/red/green, respectively. There is a significant increase in the ATE+ curve near 55, as an early indication of rapid growth in Wikipedia's contents.



(a) ATEM+ (b) ATEM− (c) TEM

Fig. 7. Most frequent motifs of Wikipedia from left to right and first row to second row. (a) ATEM+ shows motifs 1, 8 and 12 have the most positive influence on other motifs. (b) ATEM− shows motif 9 has the most negative influence on other motifs. (c) TEM shows asymmetric influence among the motifs but the contrast is not as strong as ATEM+ and ATEM−.

4 Probabilistic Cones for Trajectories Prediction

The heart of the prediction method is the model-based probabilistic light cone prediction of instability trajectories using the spectral radius (TE/ATE±) of associative transfer entropy matrix, TEM/ATEM±. A MCMC method to predict citation growth based on preferential attach models is described in [2]. However, critical transitions is not discussed. Our method differs in model-based probabilistic light cone generation using natural logarithmic curve modeling. Given

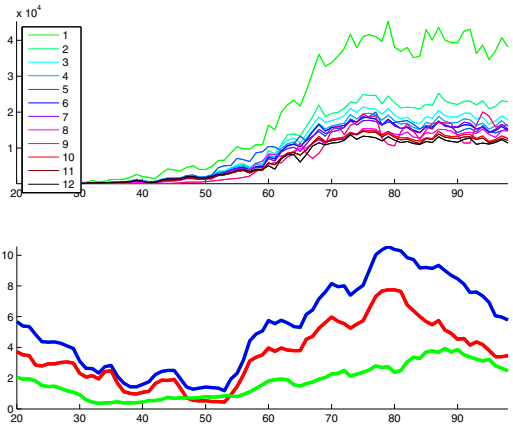


Fig. 8. Information dynamics of Wikipedia motifs. Top: Wikipedia motif occurrence in time. Bottom: TE/ATE± curves in blue/red/green, respectively. There is a significant increase in the ATE+ curve near 55, as an early indication of rapid growth in Wikipedia's contents.

ATE+/TE cross-over signature as the early warning signals of critical transition, we set our goal to predict the likelihood of system trajectories. We first observe that ATE+ and TE are maximized, at the same time ATE− flattened, prior to the system switching to alternative stable regimes. We therefore apply model-based statistical forecasting to estimate the likelihood of TE/ATE± trajectories. Our prediction method will continuously output the update of probabilistic cones as the system progresses.

We choose natural logarithm curves to model the growth rate of instability trajectory as information transfer grows toward maximization prior to critical transitions. We apply a moving time window over the observed spectral radius (TE/ATE±) time series to derive the unknown coefficients and constants for natural logarithm curves. For a given prediction time point, we generate the probabilistic light cone based on 95% confidence intervals of predicted instability trajectories with fitted natural logarithm curves.

4.1 Model-Based Forecasting

We propose a model-based statistical forecasting to estimate the TE/ATE trajectories. An advantage of analyzing the TE/ATE curves is that TE/ATE removes the spikes of the raw data. We have observed that the TE curve can be well approximated by the natural logarithm with an unknown coefficient a and a constant c:

$$g(t) = a \ln(t) + c. \tag{5}$$

Our observations of the TE/ATE curve in different types of data show that the curve approaches its maximum gradually, instead of a spiking and sudden

approach. The latter, convex increase, is very common, as in Schaffer's autocorrelation, variances, and other signals. We believe that TE/ATE serves as a better a prediction function because the increase is more concave.

To estimate the coefficient a, instead of fitting the TE/ATE curves deterministically with the logarithmic function in (5), we estimate the rate of change with various discrete time steps. This generates a prediction cone. Taking the derivative of (5), we have $\hat{g}'(t) = \frac{a}{t}$. We obtain the following discretized version for the unknown a :

$$\frac{\triangle g}{\triangle t} = a\frac{1}{t}. \tag{6}$$

For a fixed timestep $\triangle t$ in a window $[T_{\text{start}}, T_{\text{end}}]$, we use the least squares to solve the unknown a. First, we write the following matrix equation:

$$\begin{bmatrix} \frac{1}{t_1+\triangle t/2} \\ \frac{1}{t_2+\triangle t/2} \\ \vdots \\ \frac{1}{t_k+\triangle t/2} \end{bmatrix} a = \begin{bmatrix} \frac{g(t_1+\triangle t)-g(t_1)}{\triangle t} \\ \frac{g(t_2+\triangle t)-g(t_2)}{\triangle t} \\ \vdots \\ \frac{g(t_k+\triangle t)-g(t_k)}{\triangle t} \end{bmatrix}, \tag{7}$$

where $t_1, , t_k+\triangle t \in [T_{\text{start}}, T_{\text{end}}]$. The approximation of $a_\triangle t$ for the fixed timestep $\triangle t$ is then obtained by

$$a_{\triangle t} = \text{argmin}_a \left\| \begin{bmatrix} \frac{1}{t_1+\triangle t/2} \\ \frac{1}{t_2+\triangle t/2} \\ \vdots \\ \frac{1}{t_k+\triangle t/2} \end{bmatrix} a - \begin{bmatrix} \frac{g(t_1-\triangle t)-g(t_1)}{\triangle t} \\ \frac{g(t_2-\triangle t)-g(t_2)}{\triangle t} \\ \vdots \\ \frac{g(t_k-\triangle t)-g(t_k)}{\triangle t} \end{bmatrix} \right\|_2, \tag{8}$$

The constant corresponding to this timestep is then $c_{\triangle t} = g(T_{\text{end}}) - a_{\triangle t} \ln(T_{\text{end}})$. Therefore, the predicted value for the future time $T_{\text{end}} + t_d$ is

$$F_{\triangle t}(t_d) = a_{\triangle t} \ln(T_{\text{end}} + t_d) + c_{\triangle t}. \tag{9}$$

4.2 Probabilistic Light Cone and Error Estimation

Now that we have established a method to estimate the constants c and a, we can generate a probabilistic light cone at each time as we vary the δt to obtain multiple estimations of c and a. Therefore, a probabilistic light cone at a given time consists of a collections of natural logarithm curves starting from that point.

To estimate the error of this prediction for the immediate next timestep, we calculate the following:

$$\text{error}_{\triangle t}(i) = g(t_i + \triangle t) - [a_{\triangle t} \ln(t_i + \triangle t) + c(i)], \tag{10}$$

where $c(i) = g(t_{i-1}) - a_{\triangle t} \ln(t_{i-1})$.

Let E be the collection of all $\text{error}_{\triangle t}$ over all timesteps $\triangle t$ and let $\sigma =$ standard deviation of E. Let G be the collection of all predicted value $G_\triangle t(t_d)$ over all

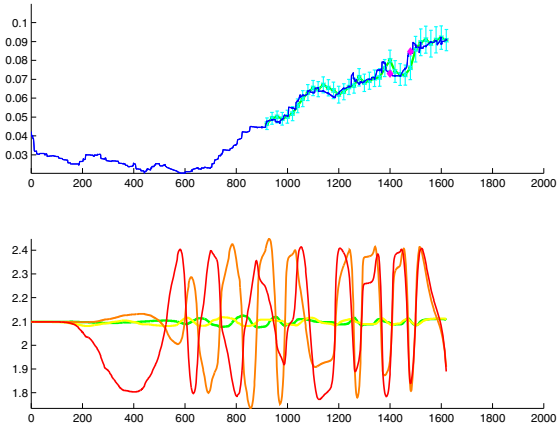


Fig. 9. Dark blue: actual trajectory. Green: predicted trajectory. Light blue: 95% confidence interval of predicted trajectory. Magenta: actual trajectory points outside the 95% confidence interval.

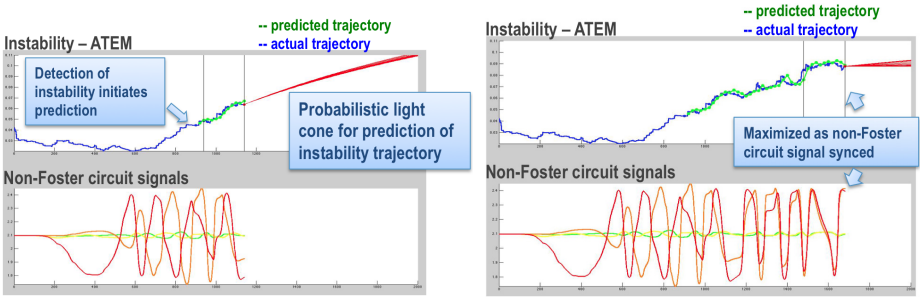


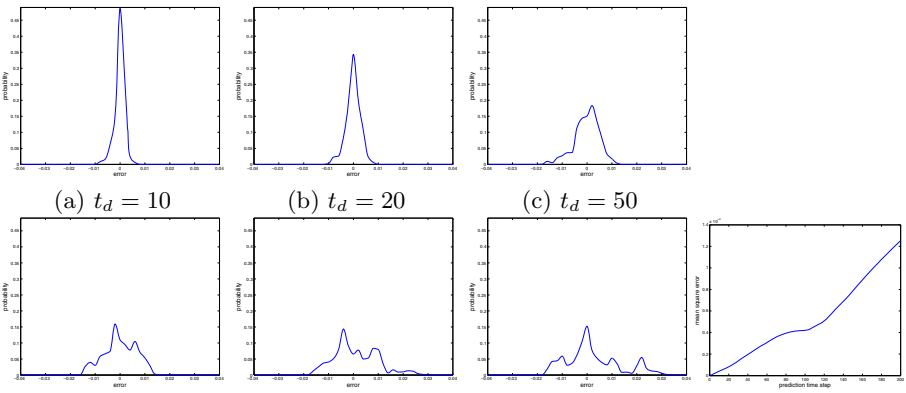
Fig. 10. Red: prediction light cone. Blue: actual TE trajectory. Green: predicted value.

timesteps $\triangle t$ and μ =mean(G). Therefore, the 95% confidence interval for the future time $T_{end} + t_d$ is

$$CI = [\mu - 1.96\frac{\sigma}{\sqrt{N}}, \mu + 1.96\frac{\sigma}{\sqrt{N}}], \tag{11}$$

where N is the size of E.

Fig. 9 shows the 95% confidence interval in light blue. Two out of 36 points of the actual trajectory were outside the 95% confidence interval, right before the non-Foster circuits are fully synched around time $t = 1500$. Fig. 10 shows two snapshots of the light cone produced according to the method described above and the predicted trajectory.



(a) $t_d = 10$ (b) $t_d = 20$ (c) $t_d = 50$

(d) $t_d = 100$ (e) $t_d = 150$ (f) $t_d = 200$ (g) prediction MSE

Fig. 11. Error distribution for each prediction leap time step. As expected, a smaller prediction leap time step t_d gives more accurate prediction, and as t_d increases, the error distribution flattens out. (g) MSE of prediction values increases as prediction timestep increases.

To see how far ahead in tim t_d one can predict the value in relation with error, the distribution of errors for each time t_d is plotted in Fig. 11. As the time leap t_d increases, the distribution of errors starts to flatten out, because the error increases. This is shown in Fig. 11 (g), where the curve is the mean squared error of the prediction values versus time leap t_d.

5 Conclusions

We proposed a novel information dynamic spectrum framework for automated detection of critical transitions and identification of directional influences. We have shown that the framework is able to: (1) provide an effective measure fro quantifying associative, asymmetric directional influence rather than symmetric influence, (2) provide an effective formulation that capture system-wise directional influence, rather than pair-wise influence and (3) provide an effective measure for detecting instability in systems with directional influence dynamics. Within this framework, we further proposed a method that analyzes the time series of complex systems to predict instability trajectory toward critical transitions. This enables the advancement of predicting dynamics of complex systems. The advantages include: (1) provide system trajectory prediction using instability signals (TE/ATE±) which capture information that cannot be discerned by looking at (or predicting) individual signal alone and (2) provide model-based prediction with $x\%$ (e.g. 95%) confidence interval prediction of system trajectories beyond conventional deterministic predictions or complicated probability state space predictions.

References

1. Bandt, C., Pompe, B.: Permutation Entropy: A Natural Complexity Measure for Time Series. Physical Review Letters 88(17) (2002)
2. Clauset, A.: http://www.cs.unm.edu/ aaron/blog/archives/2012/03/milestone.htm
3. Jurgens, D., Lu, T.-C.: Temporal Motifs Reveal the Dynamics of Editor Interactions in Wikipedia. In: ICWSM (2012)
4. Kwon, O., Yang, J.-S.: Information Flow between Stock Indices, 2008 EPL 82 68003 (2008)
5. Lizier, J.T., Prokopenko, M., Zoomaya, A.Y.: Coherent Information Structure in Complex Computation. Theory in Biosciences Special Issue on Guided Self-organization (2012)
6. Moon, H., Lu, T.-C.: Early Warning Signal of Complex Systems: Network Spectrum and Critical Transitions. In: WIN (2010)
7. Moon, H., Lu, T.-C.: Network Catastrophe: Self-Organized Patterns Reveal both the Instability and the Structure of Complex Networks (2012) (preprint)
8. Scheffer, M., Bascompte, J., Brock, W.A., Brovkin, V., Carpenter, S.R., Dakos, V., Held, H., van Nes, E.H., Rietkerk, M., Sugihara, G.: Early-warning signals for critical transitions. Nature 461 (2009)
9. Scheffer, M., Carpenter, S.R., Lenton, T.M., Bascompte, J., Brock, W., Dakos, V., van de Koppel, J., van de Leemput, I.A., Levin, S.A., van Nes, E.H., Pascual, M., Vandermeer, J.: Anticipating Critical Transitions. Science 338, 344 (2012)
10. Schreiber, T.: Measuring Information Transfer. Phys. Rev. Lett. 85, 461 (2000)
11. Staniek, M., Lehnertz, K.: Symbolic Transfer Entropy. Physical Review Letters 100, 15801 (2008)
12. White, C.R., Colburn, J.S., Nagele, R.G.: A Non-Foster VHF Monopole Antenna. IEEE Antennas and Wireless Propagation Letters 11 (2012)