Research Article

Topology Identification of Coupling Map Lattice under Sparsity Condition

Jiangni Yu, 1,2 Lixiang Li, 1,2 and Yixian Yang 1,2

1 Information Security Center, State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China
2 National Engineering Laboratory for Disaster Backup and Recovery, Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Lixiang Li; li_lixiang2006@163.com

Received 30 July 2014; Accepted 20 September 2014

Academic Editor: Florin Pop

Copyright © 2015 Jiangni Yu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Coupling map lattice is an efficient mathematical model for studying complex systems. This paper studies the topology identification of coupled map lattice (CML) under the sparsity condition. We convert the identification problem into the problem of solving the underdetermined linear equations. The \( \ell_1 \) norm method is used to solve the underdetermined equations. The requirement of data characters and sampling times are discussed in detail. We find that the high entropy and small coupling coefficient data are suitable for the identification. When the measurement time is more than 2.86 times sparsity, the accuracy of identification can reach an acceptable level. And when the measurement time reaches 4 times sparsity, we can receive a fairly good accuracy.

1. Introduction

The coupled map lattice with nonlocally coupling chaotic characteristic is widely observed and highly involved in many fields, which ranges from complex network [1–3] to neural network, from biological system to ecological system, and from physics to computer science [4–6]. Driven by some practical applications which benefit from the better controlling of CML, a great deal of current research of CML has focused on dynamical analysis, control, and modeling. However, the behavior of CML is largely influenced by the topology of the network which generally is invisible to us. Thus, its identification usually becomes the promise of application.

Up to now, various researches focus on identifying the patterns of coupled map lattice models [7, 8], such as the study of the formation and evolution of spatiotemporal patterns based on a reference model [7], and the identification of CML based on the wavelet [8]. However, few researches have discussed the topology identification of CML. There are two problems of topology identification: how we should get the result with less measurement time (M) and what precondition the measured data need.

In this paper, as CML is discrete, for the facilitation of identification, we transform the CML equation so that the identification problem can be converted into a problem of solving the linear equation

\[ y = \Phi x. \]

The \( \ell_1 \) method is introduced to solve the topology identification of CML. In the process of identification, the relationship between the measurement time (M) and the sparsity (S) of x is taken into consideration. In the experiment, through the analysis of entropy, we discuss the coupling coefficient and the solvability of the equation \( y = \Phi x \). We find that smaller coupling coefficient benefits the identification. Through further research, we study the influence of the measurement time on the identification precision; that is, when the measurement time \( M \geq 2.86S \), we can achieve a decent identification result, and when \( M \geq 4S \), the identification result is very good.

2. Analysis of Identification Process

Considering a CML with \( N \) elements is as follows:

\[ z_{t+1}(i) = (1 - \varepsilon) f(z_t(i)) + \frac{\varepsilon}{2W} \sum_{j=1}^{W} c_j F_i(t, j), \]  
(1)
where $z_i(i)$ is the state of each node, $i = 1, 2, \ldots, N$, and $t$ is the discrete time step. $\epsilon \in (0, 1)$ is the strength of the coupling. $f(z) = 4z(1-z)$ is the standard form of logistic nonlinear function. Also $F_i(t, j) = f(z_i(i-j)) + f(z_i(i+j))$, where $j = 1, 2, \ldots, W$. $c_{ij} \in \mathbb{R}^{N \times (2M+1)}$ is the weighted topology connection from element $i$ to its neighbor elements $i - j$ and $i + j$, and the weighted topology connection obeys the periodic boundary condition. The schematic diagram is shown as in Figure 1.

Normally, the weighted topology connections are composed by the known part $K$ and the unknown part $X$. Our purpose is to identify this unknown part. In order to identify these unknown connections, the dynamical equation (1) can be transformed as the following equation:

$$\frac{2W}{\epsilon} (z_{t+1}(i) - (1-\epsilon)f(z_t(i))) = \sum_{j=1}^{W} c_{ij} F_i(t, j). \quad (2)$$

Now, we expand the matrix $c \in \mathbb{R}^{N \times (2M+1)}$ to $C \in \mathbb{R}^{N \times N}$. In the expanded matrix $C$, if $j > W$, $C_{ij} = 0$; if $j \leq W$, $C_{ij} = c_{ij}$. Under these circumstances, (2) can be expressed as

$$\frac{2W}{\epsilon} (z_{t+1}(i) - (1-\epsilon)f(z_t(i))) = \sum_{j=1}^{N} C_{ij} \cdot f(z_t(j)). \quad (3)$$

We set $R(t, i) = (N/\epsilon)(x_{t+1}(i) - (1-\epsilon)f(x_t(i)))$. If we put $t = 1, 2, \ldots, M$ iterations together, then we get $M$ measurements. Its formula can be expressed as

$$R(1,1) \ldots R(1,N) 
\vdots
R(M,1) \ldots R(M,N) = \begin{pmatrix} f_1(1) & \cdots & f_1(N) \\
\vdots & \ddots & \vdots \\
f_M(1) & \cdots & f_M(N) \end{pmatrix} * C = \begin{pmatrix} C_{11} & \cdots & C_{1N} \\
\vdots & \ddots & \vdots \\
C_{1N} & \cdots & C_{NN} \end{pmatrix}. \quad (4)$$

Here, $C$ can be divided into two parts, that is, the known part $K$ and unknown part $X$, which is $C = K + X$. Thus, to get the solution $X$, we have

$$R(1,1) \ldots R(1,N) - \begin{pmatrix} f_1(1) & \cdots & f_1(N) \\
\vdots & \ddots & \vdots \\
f_M(1) & \cdots & f_M(N) \end{pmatrix} * K = \begin{pmatrix} f_1(1) & \cdots & f_1(N) \\
\vdots & \ddots & \vdots \\
f_M(1) & \cdots & f_M(N) \end{pmatrix} * X. \quad (5)$$

To simplify, the above equation set can be represented as

$$Y = \Phi \ast X, \quad (6)$$

where $Y$ implies the left side of (5) and $\Phi$ implies the left matrix of the right side.

Obviously, the elements in $Y$ and $\Phi$ are decided by $\epsilon$, $W$, and the sampled value $z$. The size of $Y$ and $\Phi$ are decided by $M$ and $N$. Thus, (6) can be expressed as

$$(y_1, \ldots, y_N) = \Phi \ast (x_1, \ldots, x_N), \quad (7)$$

where $x_1 \in \mathbb{R}^{N \times 1}$ and $y_1 \in \mathbb{R}^{M \times 1}$ are the columns of $X$ and $Y$, respectively. Without loss of generality, we use $x \in \mathbb{R}^{N \times 1}$ instead of $x_i$ and use $y \in \mathbb{R}^{M \times 1}$ instead of $y_i$. The equation set finally can be described as

$$y = \Phi \ast x. \quad (8)$$

Eventually, we transform the topology identification problem into a problem of solving the equation set. The vectors $y$ and $\Phi$ can be computed by the sampled value $z_i(i)$ (which is the state value of each element in the $t$th iteration). The character of solutions $x$ depends on the matrix $\Phi \in \mathbb{R}^{M \times N}$. When there are at most $S$ unknown connections of each element, which means the nonzero value of $x$ will be less than $S$, in this case we defined $x$ as $S$-sparse. In the following section, the influences of $\Phi$ and $S$ will be taken into consideration in the identification problem.

### 3. The Analysis of $\Phi$ and $U$

Normally, when $\Phi$ is full rank, which means $\Phi$ is reversible, (5) has a unique solution, which implies that the measurement size $M$ should be larger than the element $N$. However, when $M$ is smaller than $N$, the identification of the unknown connections is an interesting problem. On account of the partial unknown connection, namely, $x$ is $S$-sparse, the solution can be worked out by an underdetermined equation $y = \Phi x$ when $M < N$. The solving of $y = \Phi x$ is related to two aspects: the construction of $\Phi$ and the method to work out the solution $x$.

#### 3.1. The Solving Condition for $\Phi$

According to [9], the necessary and sufficient condition for the solution of the underdetermined problem is that, for any vector $x$ sharing the same sparsity $s$, there exists a $\delta_s \in (0, 1)$ such that

$$(1 - \delta_s) \|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta_s) \|x\|_2^2, \quad (9)$$

where $\delta_s \in (0, 1)$ implies that the columns of $\Phi$ should be linearly independent. The property is generally satisfying the law of random matrix such as i.i.d Gauss and Bernoulli. Recently, owning to the pseudorandom property of chaotic system, the chaos sequence is used for the construction of $\Phi$. It was proved in [10] that when the sampling distance guarantees a certain value, the logistic system satisfies the RIP condition. In [11], a series of chaos systems, such as Lorenz system and Chua’s circuit, are used to produce $\Phi$. Thus, in the model of this paper, it was promising to identify the unknown connections.
3.2. The Solving Process of $x$. As illustrated above, the unknown vector is $S$-sparse; then the resolution problem can be transformed into an optimization problem which is to find the most sparse solution of (8). Thus the $\ell_0$ norm optimization is taken into consideration. However, to find a $x = \text{argmin} \|x\|_0$ to satisfy $y = \Phi x$ is a NP hard problem that requires an exhaustive enumeration of all $C(S,N)$ possible combinations for the locations of the nonzero entries in $x$. Then the application of $\ell_2$ norm optimization is tried, which is to find a $x = \text{argmin} \|x\|_2$ to satisfy $y = \Phi x$. It is also denied for the lack of precision. After that, $\ell_1$ norm is considered; it is more efficient than $\ell_0$ and more precise than $\ell_2$. The optimization method is expressed as follows:

$$x = \text{argmin} \|x\|_1 \quad \text{s.t.} \quad y = \Phi x. \quad (10)$$

The $\ell_1$ optimization problem can be transformed to a linear programming known as basis pursuit (BP). As [12] illustrated, the matrix $\Phi$ can be shown to have the RIP, when the measurement time satisfies the condition as follows:

$$M > cS \times \log\left(\frac{N}{S}\right), \quad (11)$$

where $c$ is a constant; we can exactly reconstruct $S$-sparse vector $x$ via the BP method. Plenty of experimental results show that most of $S$-sparse signals can be worked out when $M \geq 4S$ [13]. In this paper, we will discuss the relation between $M$ and $S$ in this identification circumstance by experimental analyses.

4. Experiments and Discussions

In the simulation, we set $N = 100$, $W = 40$, $z_0(i) \in [0, 1]$, $i = 1, 2, \ldots, N$, $\varepsilon = 0.01$. The chaotic state trajectory of (1) by 100 iterations is shown in Figure 2.

The character of random likeness and boundedness of chaotic system can be observed from Figure 2, and for a better perceptual intuition the state of the 50th element is shown in Figure 3.

4.1. The Influence of $\varepsilon$ on the Identification. The chaotic behavior is obvious when $\varepsilon$ is low, while the system may exist in many states, such as chaos, periodicity, and synchronization. To observe the changing of the network state, we use the information entropy to describe the state as illustrated in [14], by which higher value of entropy implies better chaos character of each node and vice versa. The state of CML is influenced by the parameters $\varepsilon$ and $W$. With the changing of these factors, the information entropy is shown as Figure 4.

Figure 4 shows the entropy with different $\varepsilon$ and $W$; it is obvious that the value of $\varepsilon$ is the main influence factor to the entropy.

(i) When $\varepsilon < 0.2$, the degree of coupling is low. It shows a high value of entropy, which means the high randomness of the network. CML system is under a nonsynchronous state, and the sampled sequence benefits the identification.

(ii) When $0.2 < \varepsilon < 0.4$, it shows a low value of entropy, and the CML system is under a complex state between the nonsynchronous state and the synchronous one.
Some behavior such as periodicity, which indicates the randomness is weaker, may lead to a low entropy of the sequence. Thus, it is not the favorable condition for identification.

(iii) When $0.4 < \varepsilon$, it arrives at a high value of entropy as chaos because of the synchronization of CML. Considering the synchronization of elements, the correlation between the columns of $\Phi$ is high and does not conform to the RIP condition. Therefore, despite the high entropy, it is not suitable for identification.

Thus, for the requirement of identification, smaller $\varepsilon$ is better for the identification.

4.2. Identification Result of Unknown Connections. Here we set $\varepsilon = 0.01$, $N = 100$, $S = 20$, $M = 60$, $W = 50$. Figure 5 shows the connections of an element to the others achieve a successful identification with accuracy $\|x - \tilde{x}\| = 1.8153e - 007$.

The optimization process is shown in Figure 6.

From Figure 6, the estimated value $\tilde{x}$ converges to the stable state, and the result is correct when $\|x - \tilde{x}\|$ is closed to zero. The experiment is implemented in Matlab 2013 on Windows 8 Service Pack 1 64-bit operation system. We use a desktop which has a 4-core Intel(R) Core(TM) i3-2120 processor running at 3.30 GHz and 2 GB of RAM. It can be observed that with 4 iterations the result can be almost accurate, and it only takes 0.073322 seconds in our experimental simulation.

According to the above experiments, the identification process works perfectly when $\varepsilon = 0.01$. But does the algorithm work besides this value? As illustrated in Figure 4, higher coupling coefficient leads to a low entropy or a synchronized state as b and c. The algorithm works well when $\varepsilon < 0.1$; thus the premise for completing the identification progress smoothly depends on the coupling coefficient of the network. The result just corresponds with the conclusion that smaller $\varepsilon$ benefits the identification.

4.3. The Influence of Measurement Time and Sparsity. As we discussed above, to solve the underdetermined equation, the measurement time and the solution sparsity should fulfill a certain condition. For a more detailed understanding of these correlations, the three-dimensional diagram was drawn and is shown in Figure 7.

From Figure 7 we discover that with smaller measurement time and bigger sparsity the solution turns bad and vice versa. To be more detailed, we showed the aerial view in Figure 8.

For Figure 8, we can divide the areas into four parts a, b, c, and d by the distribution of solution effect. The red line $y = 0.35$ means the ratio $M/S = 2.86$ and the green line $y = 0.25$ means the ratio $M/S = 4$.

(i) For area a, when $M < S$, it does not meet the solution conditions.

(ii) For area b, when $S < M < 2.86S$, $x$ almost cannot be calculated by $\Phi$ accurately.

(iii) For area c, when $2.86S < M < 4S$, $x$ can be calculated out with acceptable accuracy.

(iv) For area d, when $M > 4S$, $x$ can be calculated out exactly.

The area d corresponds with the traditional condition $M > 4S$ discussed above. However, from the experiment we learn that the area c is regarded as the acceptable one. Such result may have a guiding significance in the identification process.

5. Conclusion and Future Work

In this paper, we propose a new approach for the identification of CML. We work out the unknown connections vector through the $\ell_1$ method. The requirement of sampling time and data characters are discussed in detail in this paper, and the identification result meets the precision of the experiment. Moreover, also some possibility of further researches exists.

On one side, assuming there is an undirected network, besides all the values of the nodes’ state, we only know the total sparsity of all the connections (some nodes may have sparse connections, while some nodes’ connections may be not sparse); how can we get the correct identification result? This problem may be solved by the following steps. Firstly, we should find the connections between two nodes with sparse unknown connections and estimate them. Secondly, owning to the symmetry characteristic of undirected network, with more and more unknown connections being identified, some of the connections of nodes which were not sparse become sparse; thus we should find these nodes and estimate the connections. Thirdly, continue the second step until we know that remaining nodes’ connections cannot be estimated.

On the other side, comparing with traditional method of topology identification, $\ell_1$ method can finish the process
of identification with lesser sampling time, but it requires knowing the whole states of the network. Its applications may be limited in the network which are able to know all the states of the nodes. Thus, in the future work, the topology identification with the deficiency of a part of nodes’ state value may be discussed.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This paper is supported by the National Natural Science Foundation of China (Grant nos. 61170269, 6112061), the Beijing Higher Education Young Elite Teacher Project (Grant no. YETP0449), the Asia Foresight Program under NSFC Grant (Grant no. 6116140320), and the Beijing Natural Science Foundation (Grant no. 4142016).

References


