Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Causal network construction based on convergent cross mapping (CCM) for alarm system root cause tracing of nonlinear industrial process

Chang Tian[1], Chunhui Zhao[1,*], Haidong Fan[2], Zhenwei Zhang[2]
*1. State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering, Zhejiang University, Hangzhou, CO 310027, China (*corresponding author: e-mail: chhzhao@zju.edu.cn).*
2. Zhejiang Energy Group Research Institute, Hangzhou, 310000, China

Abstract: In terms of an alarm system, the propagation of a fault is identified as the main reason for low efficiency and the leading cause of dramatic industrial accidents. Thus, tracing the root causes of faulty conditions that lead to alarm floods is necessary. For root cause tracing, a widely accepted method is to characterize the process by causality at first and then trace the root causes. This work focuses on the former part. The conventional techniques to deal with causal analysis of industrial processes have difficulty in handling the nonlinearity of variables, obtaining accurate probability density and time lag, etc. In this work, a novel causal network construction method based on convergent cross mapping (CCM) that accurately describe process causality was proposed to deal with the above problems. First, the original monitoring variables were determined by a maximum Lyapunov method to determine whether they were chaotic time series, which aims to judge whether the application conditions of CCM can be satisfied. Then, some characteristic variables are selected from original variables through data preprocessing and descending dimension methods, which are defined as nodes that constitute the causal network. Second, the CCM-based methods are used to identify the causal direction and indirect causal relationship between variables, so as to construct the structure of the causal network. Since the CCM based on deterministic systems theory, it can handle nonlinearity and does not rely on the sample distribution. Finally, the weight of the edges in the graph is calculated to obtain the causal network which describes the process causality and serves as the basis for subsequent root causes tracing of alarms. The effectiveness of the proposed method is illustrated via a real industrial case study.

Keywords: convergent cross mapping, causal network construction, process causal analysis, alarm root causes tracing

1. INTRODUCTION

Alarm system is an indispensable task in complex industrial facilities, which are usually deployed with a large number of sensors, actuators, communication devices, and control modules (Guo et al., 2017). It is an important part to ensure the safe operation of modern industrial plants. (Yu & Zhao, 2020). Modern industrial plants consist of many interconnected units, and thus a single fault may can propagate throughout plant devices to cause alarm floods in alarm system (Hu, W et al., 2015). Consequently, the identification of the root cause of alarm floods in alarm systems is necessary to support decision and help operators to take immediate and valid measures to handle abnormal conditions and prevent any major accident due to fault propagation. Alarm root tracing is a method based on the causal-effect relationship between variables to find the abnormal propagation path and locate the root cause of the faulty condition. For root cause tracing, a widely used way is to characterize the process by causality at first and then trace the root causes. Thus, it is desirable to accurately describe the cause-effect relationship between variables in the process to provide the basis for root traceability.

In recent years, several methods for identifying the causality between variables from historical data have been proposed. The cross correlation function determines the causal relationship between variables through the similarity and time characteristics of two time series (Bauer et al., 2008). Another commonly used method is Granger causality (Granger, 1969). Specifically, one time series can be called Granger casual to the other if we can better predict the second time series by incorporating knowledge of the first one (Maciej Kamiński et al., 2001). Several methods based on frequency domain information also have been developed, such as the directed transfer function (DTF) (Kaminski & Blinowska, 1991) and the partial directed coherence (PDC) (Luiz & Sameshima, 2001). However, those methods are only applicable to linear systems and cannot handle the complex nonlinearity of real industrial processes. To deal with the nonlinear problem, Schreiber et al. (2000) proposed transfer entropy based on the information entropy theory. Other scholars have done further research on this basis (Zhao & Feng, 2017). However, since the calculation of information entropy depends on the theory of probability calculation, it is hard to ensure the accuracy of the results. In addition, the transfer entropy cannot calculate the time-delay relationship between variables, which is very important in causality mining.

Therefore, it is desirable to deal with causality between variables in a nonlinear system. The above-mentioned methods for describing the causality of industrial plants are from a pure statistical point of view, which all face high requirements on sample quality. If the observed sample does not represent the overall characteristics of the process, or its statistical properties change over time, it is difficult to obtain an accurate and reliable result. However, the actual industrial process is running according to certain rules, we can use the model to describe an industrial process, thus an industrial process is more appropriate than a completely random process with a deterministic system description. Sugihara et al. (2012) proposed a Convergence Cross-Map (CCM) method from the perspective of deterministic systems, which was originally applied in the field of ecology. This method exploits the characteristics of differential homeomorphism mapping between variables in a nonlinear coupled system to transform the original causal relationship recognition into a comparison of the mutual prediction effects of embedded manifolds. The characteristics of differential homeomorphism are also common in the reconstructed manifold of industrial process data (Cheng & Zhao, 2016). So the idea of CCM is also applicable to industrial processes but it has not been applied to establishment of causal networks for real industry.

In the present work, in order to provide a basis for root cause tracing of the alarm system, a CCM-based causal network construction method is proposed to accurately describe the causalities of nonlinear industrial process. As basis, a novel CCM-based indirect causation identification method is proposed to identify indirect causal relationships in transitive causal chains. Then, a CCM-based causal network construction method is introduced. Here, the causal network is a weighted directed acyclic graph that describes causalities of process.

The main contributions of this paper are summarized as below:
1) The causality identification problem in nonlinear industrial process is solved from the perspective of deterministic system, which does not involve the problem of sample distribution, thus reducing the requirement for sample quality;

2) A convergent cross mapping (CCM) based causal network construction method is proposed for the first time in terms of nonlinear industrial processes. The effect of time lag in causal transmission is considered in the identification of causal directionality, and the indirect causation in the transitive causal chain can be identified.

This paper is organized as follows. Section 2 describes the details of proposed CCM-based causal network construction method. Section 3 shows an industrial case to validate the efficacy of the method. Finally, the conclusions are in Section 4.

2. METHODOLOGY

2.1 Framework of CCM-based causal network construction strategy

In this subsection, the framework of proposed CCM-based causal network construction strategy is introduced. The construction of CCM-based causal network contains two major issues: (i) how to identify the structure of the causal network. (ii) how to identify the causal network parameters which can also be referred as the weights of the edge.

The CCM-based causal network of a nonlinear industrial system is a directed-weighted graph, which is a set of triples of nodes and adjacency relationships. The causal network can reflect causal relationship of the system without any prior knowledge about the mechanism of the system, and can be represented in mathematic form as:

$$CN = (V, A_d, A_w) \quad (1)$$

where V is the set of nodes in causal network, A_d is a square matrix describing the adjacency between nodes, and A_w is the parameters of causal network in which the elements represent the weights of the edges in the network. Notably, A_w is a square matrix for a specific condition, and its parameters can be interval numbers. A flow chart of the proposed method is shown in Fig. 1.

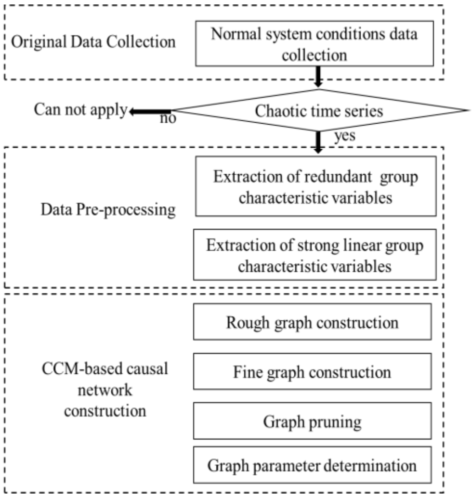

Fig. 1. The framework of the CCM-based causal network construction strategy

2.2 CCM-based causal network construction

The data used to construct causal network are monitoring data collected from industrial systems. Mathematically, the collected monitoring dataset contains N attributes which represent N sensors used in an industrial system constitute an N-dimensional matrix can be presented as follow:

$$D = [X_1, X_2, \dots, X_N]^T \quad (2)$$

Each column in the dataset is a time series of a specific variable. It should be ensured that the collected data are obtained from normal operating sensors under normal system conditions.

CCM is a method of analysing the causal relationship between time series on the assumption of deterministic systems, references (Sugihara et al., 2012) for specific methods. Although it is demonstrated that most industrial processes is a nonlinear deterministic system(Li , Zhao , & Gao, 2018), it still requires to judge whether the data obtained can be analysed using a CCM-based approach to causality analysis. In order to judge whether the method is applicable, it must be determined whether the time series is a chaotic time series generated by a nonlinear deterministic system. In this paper, whether the maximum Lyapunov exponent of time series is greater than zero is used as a criterion for judging whether a time series is a chaotic time series (Lu, 2002). If the above conditions are not met, this method is not suitable.

Furthermore, in modern industrial processes, in order to ensure the reliability of the process monitoring results, redundant sensors are deployed for the same attribute with the same physical meaning and coupling relationships. Thus, simple variable preprocessing is needed. First, the average value of redundant sensors is used as a characteristic variable to represent the monitoring results of an attribute. Based upon the above, feature extraction is required for variable groups in process monitoring variables that have a strong linear relationship that CCM can hardly identify their causalities correctly. Because the variables in these groups are mostly from a certain component in process and the correlation coefficient of these variables is close to 1 thus they can be represented by a characteristic variable. In this paper, the principal component analysis (PCA) method (Wold, 1987) is used to extract the characteristic variables of variable groups.

Through above operations, all basic elements for causal modelling are obtained and they provide a suitable foundation for the following analysis and modelling. The variables obtained are used as nodes to build the network. Causal network construction integrates the decentralized nodes to an organic whole which involves two major issues: the determination of network structure, and parameters. In the following, the above two problems will be solved.

Given a specific industrial system with N variables which contains the original monitoring data from sensors, the characteristic variables vr representing redundant sensors, and characteristic variables vl of subgroups of strong linear relationship, the CCM-based causal network can be constructed according to the following procedures:

(1) Rough graph construction

The data set D after preprocessed has N variables, each X_i is a time series:

$$D = [X_1, X_2, \dots, X_N] \tag{3}$$

The starting of the algorithm is a fully connected undirected graph:

$$P = \{< x_i, x_j >\}, where\ i, j = 1,2, \dots, N, i \neq j \tag{4}$$

Initialize the adjacency matrix A_d and the weight matrix A_w to zero matrix. Let $V = \{x_1, x_2, \dots, x_N\}$ be the complete set of nodes in causal network model, perform CCM in both directions for each pair of nodes: $x_i, x_j \in V$, where $i \neq j$. Specifically, choose a threshold ε:

(i) If cross-map estimate of x_i well-behaved, but the cross-map model for x_j does not converge to a number greater than ε, which means x_i drives x_j unilaterally, then connect the nodes x_i, x_j,and the direction is $x_i \rightarrow x_j$, let the element $ad_{i,j}$ of row i and column j of adjacency matrix A_d be 1, and $ad_{j,i}$ be 0, vice versa.

(ii) If both cross-map estimates of x_i and x_j do not converge to a number greater than ε, then there is no causal relationship between the nodes x_i, x_j, and the connection between nodes is deleted, let the element $ad_{i,j}, ad_{j,i}$ of adjacency matrix A_d be 0.

(iii) If the CCM estimates in both directions converge to a number greater than the threshold, the causal relationship between the variables requires further analysis. Put nodes x_i and x_j as pending nodes into the pending set T:

$$T = \{x_i, x_j\}, i, j \in N \tag{5}$$

(2) Fine graph construction

The primary CCM method cannot recognize the causal direction of the variables in the set T. Cause with moderately strong forcing from X, even though Y exerts no effect, there may still be partial cross mapping of Y arising from the contemporaneous dependence of Y on X. With extremely strong forcing, the intrinsic dynamics of the forced variable become subordinate to the forcing variable, leading to the well-studied phenomenon of “synchrony” (Josic, 2000).

Select a reasonable maximum delay range, and perform extended-CCM (Ye et al., 2015) in both directions for each pair of nodes: $x_i, x_j \in T$. Then determine the direction of causality of the two variables by the time lag where the optimal cross-map estimate appears.

If x_i is best able to cross map x_j forward in time, whereas cross mapping in the true direction shows optimal prediction for negative time lags. Connect the nodes x_i, x_j, and the direction is $x_i \rightarrow x_j$, let the element $ad_{i,j}$ of row i and column j of adjacency matrix A_d be 1, and $ad_{j,i}$ be 0, vice versa. Because the amount of computation of the extended CCM method is much larger than the primary CCM method, this coarse-to-fine network structure construction method can solve the problem with less computation. Through the above steps, a directed graph $G = (V, A_d)$ is obtained.

(3) Graph pruning

There may be the following causal relationship as shown in Fig 2: y_1 causes y_2 causesy_3, such that indirect causation from y_1 to y_3 occurs. However, only the direct links are required in the causal network model, which not only simplifies the network structure, but also facilitates the subsequent root causes tracing. Therefore, such transitive causal chains should be identified and the indirect causal connection should be deleted in the network.

Think of each potential transitive causal chain as a subgraph, in this step, firstly find the pending subgraph set $GT = \{g_1, g_2, \dots g_m\}$ in the graph G through the adjacency matrix, and each element g_m in set GT is a subgraph $g_m =< v_m, ad_m >$, where v_i represents the set of nodes in subgraph g_m and ad_m represent the adjacency list, and the pseudocode is presented in Algorithm 1. The Algorithm 1 is to find the transitive causal chains as shown in Fig. 4 through the adjacency matrix $A_d = \{a_{i,j}, i, j = 1,2, \dots N\}$. Directed graph

$G = (V, A_d)$ has N nodes in total, then let $\mathbb{N} = \{1,2,\dots,\mathrm{N}\}$. The length of the transitive causal chains L is represented by the number of nodes in the transitive causal chains. The algorithm uses a traversal search idea, starting with a node and searching for any possible transitive causal chains from the node that are less than N in length. Think of each transitive causal chain as a subgraph g, recording the set of subgraph nodes v and the adjacency lists ad. Then iterate through all N nodes to find all the transitive causal chains, their collection is set GT.

Algorithm 1 Transitive Causal Chains Recognition

```
Input Argument: G =< V, A_d >
Output Argument: GT
for i, j in ℕ, i ≠ j do
  if a_ij == 1 then
    Q = {j}, P = {p∈ℕ, p ≠ i, j}, L = 2, Q_it = ∅
    REPEAT
      L = L + 1, Q_it = ∅
      for q in Q do
        for p in P do
          if a_q,p == 1 then
            add p in Q_it
            if a_i,p == 1 then
              build node set v, build adjacency list ad
              add g =< v, ad > in set GT
            end if
          end if
        end for
      end for
      Q = Q_it
    UNTIL P = ∅ or L ≥ N
  end if
end for
```

Secondly, perform CCM-based indirect causation identification method for $\{g_i \in GT, i = 1,2,\dots,|GT|\}$ to determine whether the causal relationships in the transitive causal chains are direct or indirect causation, and delete the indirect causation, thereby simplifying structure of directed graph G. At this point, the structural learning of the causal network graph is completed, and finally a simplified directed graph $G = (V, A_d)$ is obtained.

(4) Graph parameter determination

The final step in the construction of causal networks is the learning of network parameters, that is, the learning of the weight of the edges. It is believed that under normal operating conditions, the CCM estimate value between variables should fluctuate within a certain range, with that range as the weight of the edges. As mentioned at the beginning of this section, the weight matrix, A_w is a square matrix for a specific condition, and its parameters are interval numbers.

For a pair of connected vertices in the network G, choose m randomly segments of length $\mathcal{T}$ selected for the CCM calculation from entire time series, and perform m times CCM estimation. The same segment was estimated using leave-one out cross-validation for all cross-mapping calculations. So far, m CCM estimate results are obtained, and the distribution is estimated by kernel density estimation, take their confidence interval with a given confidence level as the network parameter and assign value to weight matrix A_w. After updating the whole weight matrix, the final causal network graph model is obtained:

$$CN = (V, A_d, A_w) \tag{6}$$

2.3 CCM-based indirect causation identification

In this subsection, a CCM-based indirect causation identification method is proposed to identify indirect causal relationships in transitive causal chains. Consider a minimal transitive causal chain unit as shown in Fig. 2, y_1 causes y_2 causes y_3 such that indirect causation from y_1 to y_3 occurs. Ye et al.(2015) shows that the direct links (y_1 to y_2 and y_2 to y_3) are strongest with the highest cross map skill and the most immediate effects, the indirect link separated by one node (y_1 to y_3) has weaker cross map skill and longer delayed effects. When constructing causal network models, only direct causal connections are desired, which facilitates the subsequent root causes tracing. Based on the above knowledge, this paper propose an algorithm for distinguishing direct and indirect causality in the transitive causal chain like in Fig.2.

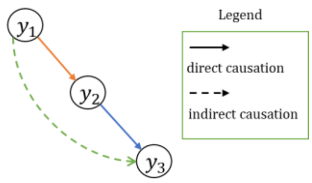

Fig. 2. Minimal Transitive Causal Chain

Consider transitive causal chain as a subgraph, because the transitive causal chain is a sparse graph, this method use the set of nodes and adjacency lists to represent it.

$$CC = (U, E_d) \tag{7}$$

where U is set of nodes in the transitive casual chain and E_d is an adjacency list which represents an array consisting of $|V|$ link lists. For $\exists u \subset U$, $E_d[u]$ is a list contains all nodes that connect to node u. The length of the transitive causal chains is represented by the number of nodes in them.

The extended-CCM method (Ye et al., 2015) is used to identify the time-delay relationships and CCM is used to calculate the CCM predictability between nodes in the transitive casual chains. The collection of time-delay between directly connected nodes in the transitive casual chain is denoted by $\mathcal{L} = \{l_i, i = 1,2,\dots,|L-1|\}$. A set of CCM predictability between directly connected nodes in the transitive casual chain is denoted by $\mathcal{W} = \{w_i, i = 1,2,\dots|L-1|\}$. Let ℓ denotes the time delay relationship between the head and tail nodes in the transitive casual chain, and $\mathscr{w}$ represents the CCM predictability, if (8) is satisfied, the causal relationship between the head and tail nodes is considered to be due to causality and is an indirect causal relationship, the pseudocode is presented in Algorithm 2.

$$\begin{cases} \ell \geq \sum l_i, i = 1,2,\dots,|L-1| \\ \mathscr{w} > \forall w_i \subset \mathcal{W} \end{cases} \tag{8}$$

Algorithm 2 Identification Of Indirect Causation In Transitive Causal Chain

for u **in** U **do**
  find head node of the transitive causal chain$|E_d[u]| \geq 2$
end for
for node **in** $\mathrm{E}_d[headnode]$ **do**
  find end-node of the transitive causal chain use $\mathrm{E}_d[\text{endnode}] = \emptyset$
  delete end-node in $\mathrm{E}_d[headnode]$
end for
calculate the time-delay relationship ℓ between head-node and end-node
calculate the CCM predictability $\mathcal{w}$ between head-node and end-node
for u **in** U **do**
  for element **in** $E_d[u]$ **do**
    calculate the time-delay relationship l_i between nodes
    calculate the CCM predictability w_i between nodes
  end for
end for
if $\ell \geq \sum l_i, i = 1,2,\dots,|L-1|$ & $\mathcal{w} > \forall w_i \subset W$, **then**
  the indirect causation between the head and tail nodes is confirmed
end if

3. CASE STUDY

This section presents a case study to demonstrate the effectiveness of the proposed methods using real industrial process data. The monitoring data in the example were taken from DCS of coal mill in a thermal power plant. The maximum Lyapunov exponent method is used to judge whether the time series of all variables are chaotic time series, the variables of the non-chaotic time series are deleted. Then, the variables are preprocessed to obtain the applicable variables as nodes that constitute the causal network. Ultimately, 6 variables for causal network construction are obtained. The descriptions of variables are listed in Table1.

Table 1. Variables for causal network construction

No.	Description
1	rotating classifier bearing temperature
2	export gas-solid mixture temperature
3	grinding bowl differential pressure
4	inlet primary air volume
5	characteristic variable of motor
6	characteristic variable of lubricating oil

For causal network construction, first, a graph of mixed directed and undirected edges is obtained using primary CCM, and a set of nodes $T = \{(1,4),(3,2)\}$ that cannot determine the causal direction using primary CCM.

Secondly, the extended-CCM method is used determine the causal direction of the nodes in the set T, thereby complete the construction of the directed graph G, resulting in a directed graph as shown in Fig.3.

Third, use algorithm proposed in section 2.3 to determine whether there is an indirect causation in each transitive causal chain. Find the transitive causal chain in the graph, in this case, $GT = < g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8 >$, where the nodes and their connections of each chain are listed in Table 2. It turns out that there are indirect causations in g_3, g_4, g_7, g_8 which are {2->5} and {4->5}. Indirect causations are shown in Fig.3 by dashed lines. Therefore, delete the edges between node 2,5 and node 4,5 in G. The resulting causal network structure model is shown in the Fig.3 with solid line. The causations between variables are shown in the Table 3, the number 1 in table indicates that the variable represented by the number of row is the cause of the variable represented by the number of column, the number 0 indicates that the variable represented by the number of row is not the cause of the variable represented by the number of column.

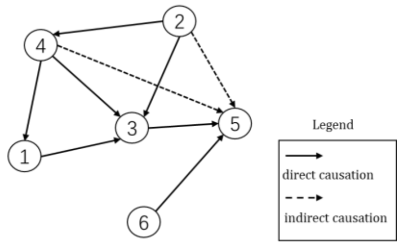

Fig. 3. Directed graph with direct and indirect causation

Table 2. Transitive causal chains and their descriptions

Chain	Nodes and connections
g_1	{2->4->3,2->3}
g_2	{2->4->5,2->5}
g_3	{2->3->5,2->5}
g_4	{4->3->5,4->5}
g_5	{4->1->3,4->3}
g_6	{2->4->1->3,2->3}
g_7	{4->1->3->5,4->5}
g_8	{2->4->1->3->5,2->5}

According to the mechanism knowledge of the coal mill, the primary air (i.e.④) is adjusted to the change of outlet gas-solid mixture temperature (i.e.②). The drying process of the coal powder is bottom-up, the change of the primary air (i.e.④) directly affects the grinding bowl differential pressure (i.e.③). But there is no direct evidence that the primary air (i.e.④) affects the rotating classifier (i.e.①), so the causality here may be an inappropriate result. The outlet gas-solid mixture temperature (i.e.②) represents the quality of the pulverized coal, the working state of the rotating classifier (i.e.①) affects the quality of the pulverized coal, resulting in the change of the pressure difference of the grinding bowl (i.e.③). The differential pressure change of grinding bowl (i.e. ③) represents a certain degree of load change and output change, which causes the working state of the motor (i.e.⑤) to change. Meanwhile, the lubricating oil (i.e.⑥) also affects the working state of the motor (i.e.⑤).

Finally, learn the weight of the network and get the weight matrix A_w. Then the final causal network $CN = (V, A_d, A_w)$ of this cause is obtained.

From another point of view, according to the actual abnormal case of this power plant, the mill bowl differential pressure alarms caused by the abnormality of the primary air temperature are selected as an example to illustrate the feasibility of the method. Although the anomaly is caused by a primary wind anomaly according to the fault record of the power plant, the fault characterizes only the mill bowl differential pressure alarm. It can be learned from the constructed causal network model that the anomaly may be caused by an abnormal primary air inlet (i.e.④), which is consistent with the actual fault cause, or it may be caused by the abnormal outlet gas-solid mixture temperature (i.e.②). The network has the ability to provide a foundation for root cause tracing of alarm system. However, the abnormality is caused by which root cause and which propagation route is need further root cause tracing, but requires further efforts which is beyond the scope of this paper.

Table 3. Result of CCM-based causal network construction

No.	1	2	3	4	5	6
1	--	0	1	0	0	0
2	0	--	1	1	0	0
3	0	0	--	0	1	0
4	1	0	1	--	0	0
5	0	0	0	0	--	0
6	0	0	0	0	1	--

In conclusion, the causal network construction method proposed in this paper can appropriately describe the causal relationship in nonlinear industrial processes.

4. CONCLUSIONS

This paper mainly proposes a CCM-based causal network construction method which accurately describes the causalities of a process to provide a basis for root cause tracing of an alarm system. By introducing CCM method into establishment of process causal networks, process nonlinearity and sample quality requirements can be solved. A CCM-based indirect causation identification method is proposed to identify indirect causal relationships in transitive causal chains and construct a more accurate and concise causal network. Besides, a practical industrial case is used to illustrate the effectiveness and feasibility of the proposed method. Some problems may need to be solved, such as how to trace the root causes through the constructed causal network, which will be the focus of our future research.

ACKNOWLEDGMENT

This work is supported by the NSFC-Zhejiang Joint Fund for the Integration of Industrialization and Informatization (No. U1709211), and the Zhejiang Key Research and Development Project (2019C01048)

REFERENCES

Bauer, M., & Thornhill, N. F. (2008). A practical method for identifying the propagation path of plant-wide disturbances. Journal of Process Control, 18(7-8), 707-719.

Cheng, F., & Zhao, J. (2016). Convergent cross mapping method in analysis of disturbances in chemical processes. Ciesc Journal.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37(3), 424-438.

Guo, C., Hu, W., Lai, S., Yang, F., & Chen, T. (2017). An accelerated alignment method for analyzing time sequences of industrial alarm floods. Journal of Process Control, 57, 102-115.

Hu, W., Wang, J., & Chen, T. (2015). A new method to detect and quantify correlated alarms with occurrence delays. Computers & Chemical Engineering, 80, 189-198.

Josic, & Kresimir. (2000). Synchronization of chaotic systems and invariant manifolds. Nonlinearity, 13(4), 1321-1336.

Kaminski, M. J. , & Blinowska, K. J. (1991). A new method of the description of the information flow in the brain structures. Biological Cybernetics, 65(3), 203-210.

Luiz A. Baccalá, & Sameshima, K. (2001). Partial directed coherence: a new concept in neural structure determination. Biological Cybernetics, 84(6), 463-474.

Li, W., Zhao, C. , & Gao, F. (2018). Linearity evaluation and variable subset partition based hierarchical process modeling and monitoring. IEEE Transactions on Industrial Electronics, 65(3), 2683-2692.

Maciej Kamiński, Ding, M., Truccolo, W. A., & Bressler, S. L. (2001). Evaluating causal relations in neural systems: granger causality, directed transfer function and statistical assessment of significance. Biological Cybernetics, 85(2), 145-157.

Odgaard, P. F., Lin, B., & Jorgensen, S. B. (2008). Observer and data-driven-model-based fault detection in power plant coal mills. IEEE Transactions on Energy Conversion, 23(2), 659-668.

Schreiber, & Thomas. (2000). Measuring information transfer. Physical Review Letters, 85(2), 461-464.

Sugihara G1, May R, Ye H, Hsieh CH, Deyle E, Fogarty M, &Munch S. (2012). Detecting causality in complex ecosystems. Science, 338(6106), 496-500.

Wold, S. (1987). Principal component analysis. Chemometrics & Intelligent Laboratory Systems, 2(1), 37-52.

Wang, R., Gao, X., Gao, J., Gao, Z., & Kang, J. (2018). An information transfer based novel framework for fault root cause tracing of complex electromechanical systems in the processing industry. Mechanical Systems and Signal Processing, 101, 121-139.

Ye, H., Deyle, E. R., Gilarranz, L. J., & Sugihara, G. (2015). Distinguishing time-delayed causal interactions using convergent cross mapping. Scientific Reports, 5, 14750.

Yu, W., & Zhao, C. (2020). Robust Monitoring and Fault Isolation of Nonlinear Industrial Processes using Denoising Autoencoder and Elastic Net, IEEE Transactions on Control Systems Technology, 28(3), 1083-1091.

Feng, L., & Zhao, C. (2020) Fault Description Based Attribute Transfer for Zero-Sample Industrial Fault Diagnosis, IEEE Transactions on Industrial Informatics, DOI: 10.1109/TII.2020.2988208.