Preprints, 10th IFAC International Symposium on
Advanced Control of Chemical Processes
Shenyang, Liaoning, China, July 25-27, 2018

Research and application of causal network modeling based on process knowledge and modified transfer entropy

ZHU Qunxiong, Ya Sitai, GENG Zhiqiang, XU Yuan, HAN Yongming, HE Yanlin*

***Abstract*—**Causal network modeling is an important part of alarm root cause analysis in industrial process. The transfer entropy is an effective method to model the causal network. However, there are some problems in determining the prediction horizon of transfer entropy. To solve the problems, a modified transfer entropy, which consider about the prediction horizon from one variable to another and to itself simultaneously, is proposed to improve the capacity of causality detection. Moreover, based on the data-driven and process knowledge modeling methods, an approach combining the modified transfer entropy with superficial process knowledge is designed to correct false calculations and optimize causal network models. Two case studies including a stochastic process and Tennessee Eastman process are carried out to illustrate the feasibility and effectiveness of the proposed approach.

I. Introduction

Alarm systems play a fundamental role in the modern industrial system. Nowadays, alarm flood is one of the most common problems of the industrial alarm systems due to the large-scale and complicated processes. Nuisance alarms and causal alarms are the main sources of alarm floods [1]. Causal network modeling is necessary to overcome the causal alarms. With a causal model, the causality can be visualized by directed graph, and the alarm root causes can be diagnosed easily and effectively. As a result, alarm floods can be alleviated at source. Therefore, capturing the causality among process variables and establishing the causal model are important and necessary to handle the problem of causal alarms.

Causal model is a model that describes causality among process variables. The methods for modeling the causality can be divided into two kinds: one is knowledge based modeling and the other one is data-driven modeling. Knowledge based modeling is qualitative, which lacks quantitative information to determine the strength of causality. Besides, it's hard to realize without the expertise. With the rapid development of computer and measuring techniques, data-driven modeling methods, such as time-delayed correlation analysis[2], granger causality[3-5], Bayesian network[6-8], interpretive structural model[9,10] and transfer entropy, have been explored and developed broadly. Transfer entropy was proposed by Thomas Schreiber to quantify the information exchange [11]. It essentially describes the causality caused by information flow, which has been applied broadly in many fields like neurology [12] and economics [13]. Margret Bauer used transfer entropy to study the disturbance propagation path in chemical plants [14]. Afterwards, different kinds of modified transfer entropy were put forward. Staniek proposed symbolic transfer entropy to reduce the influence of noise through replacing the original time series by symbolic time series [15]. Duan put forward a kind of direct transfer entropy to determine the direct causality between variables [16]. To avoid the kernel density estimation, Yu calculated the transfer entropy by use of binary alarm data [17]. Although these methods have been proven efficient, they have to meet the presumption that the process is static Markov process. That is to say, the process dynamics should keep unchanged. However, not all the real processes can be approximated as Markov processes. As a result, the transfer entropy analysis would be wrong [18].

Motivated by the above considerations, a new kind of modified transfer entropy is proposed to overcome the shortcomings of parameters optimization in this paper. Besides, the superficial process knowledge is used to adjust the results of our proposed algorithm. Causal network is then established. The feasibility and effectiveness of this approach are proven by a stochastic process and Tennessee Eastman process.

II. Principle of TRANSFER ENTROPY

Transfer entropy originates from information entropy proposed by Shannon in 1948 [19]. Information entropy is aimed at quantifying the process uncertainty. The higher the information entropy is, the greater the uncertainty is, and the more information the process has. On the basis of information entropy, in 2002 Schreiber proposed the transfer entropy theory to measure the information exchange [20]. The formula of transfer entropy is as follows:

$$T\left(X\,|Y\right)=\sum_{x_{i+1},x_i^{(k)},y_i^{(l)}} p\left(x_{i+1},x_i^{(k)},y_i^{(l)}\right)\log\frac{p\left(x_{i+1}\left|x_i^{(k)},y_i^{(l)}\right.\right)}{p\left(x_{i+1}\left|x_i^{(k)}\right.\right)} \quad (1)$$

where X and Y denote two different process variables in Markov process, $x_i^{(k)}=\left[x_i,x_{i-1},\cdots,x_{i-k+1}\right]$ and $y_i^{(l)}=\left[y_i,y_{i\text{-}1},\cdots,y_{i\text{-}l+1}\right]$ represent the measured values at time i, k and l denote the order of X and Y, respectively, log means the logarithm with base 2, $p\left(x_{i+1},x_i^{(k)},y_i^{(l)}\right)$ is the joint probability and $p\left(x_{i+1}\left|x_i^{(k)},y_i^{(l)}\right.\right)$ is the conditional probability. Considering the time delay of information exchange between two variables, Bauer changed the formula as:

$$T\left(X\,|Y\right)=\sum_{x_{i+h},x_i^{(k)},y_i^{(l)}} p\left(x_{i+h},x_i^{(k)},y_i^{(l)}\right)\log\frac{p\left(x_{i+h}\left|x_i^{(k)},y_i^{(l)}\right.\right)}{p\left(x_{i+h}\left|x_i^{(k)}\right.\right)} \quad (2)$$

*Research supported by National Natural Science Foundation of China under Grant No. 61573051, 61472021 and the Fundamental Research Funds for the Central Universities under Grant No. ZY1704, JD1708. A extension version of this conference paper has been published in Industrial & Engineering Chemistry Research, 2017, 56(48), 14282-14289.

ZHU Qunxiong, Ya Sitai, GENG Zhiqiang, XU Yuan, HAN Yongming, and HE Yanlin are with College of Information Science & Technology, Beijing University of Chemical Technology, Beijing 100029, China and the Engineering Research Center of Intelligent PSE, Ministry of Education of China, Beijing 100029, China (corresponding author to provide phone: +86-10-64426960; fax: +86-10-64437805; e-mail: heyl@mail.buct.edu.cn).

where h is the prediction horizon. Being reference, the interval between the current time and the future time should be constant. But in formula (2), the interval will change with h. This is unreasonable, and in some cases the prediction horizon couldn't be decided. So Shu changed the formula as:

$$T\left(X\left|Y\right.\right)=\sum_{x_{i+h},x_{i+h-1}^{(k)},y_i^{(l)}} p\left(x_{i+h},x_{i+h-1}^{(k)},y_i^{(l)}\right)\log\frac{p\left(x_{i+h}\left|x_{i+h-1}^{(k)},y_i^{(l)}\right.\right)}{p\left(x_{i+h}\left|x_{i+h-1}^{(k)}\right.\right)} \quad (3)$$

where $x_i^{(k)}$ in (2) is changed to $x_{i+h-1}^{(k)}$ in (3). In this way, the interval between $x_{i+h-1}^{(k)}$ and $x_i^{(k)}$ keeps constant. Since the effects of two variables make on each other are different, the transfer entropy is asymmetric. The causality between two variables can be measured as:

$$T_{y\to x}=T\left(X\left|Y\right.\right)-T\left(Y\left|X\right.\right) \quad (4)$$

If $T_{y\to x}$ is positive, Y is the cause of X; whereas X is the cause of Y. If $T_{y\to x}$ is close to zero, there is no causality between them.

III. Modified transfer entropy and causal network modeling

A. Modified transfer entropy algorithm

Transfer entropy can be expressed as the different between information entropy. The formula (3) can be rewritten as:

$$\begin{aligned}T\left(X\left|Y\right.\right)=&\sum_{x_{i+h},x_{i+h-1}^{(k)},y_i^{(l)}} p\left(x_{i+h},x_{i+h-1}^{(k)},y_i^{(l)}\right)\log p\left(x_{i+h}\left|x_{i+h-1}^{(k)},y_i^{(l)}\right.\right)\\&-\sum_{x_{i+h},x_{i+h-1}^{(k)}} p\left(x_{i+h},x_{i+h-1}^{(k)}\right)\log p\left(x_{i+h}\left|x_{i+h-1}^{(k)}\right.\right)\\&=H\left(x_{i+h}\left|x\right.\right)-H\left(x_{i+h}\left|x,y\right.\right)\end{aligned} \quad (5)$$

where $H\left(x_{i+h}\left|x\right.\right)$ is the conditional information entropy of x to x_{i+h}, and $H\left(x_{i+h}\left|x,y\right.\right)$ is the conditional information entropy of x and y to x_{i+h}. Form formula (5) we can see that the information propagates not only between two different variables but also between the same variable from current time to the future time. In formula (5), the interval between the current time and the future time is one sample interval. Different process may have different sampling time. In this way, the interval can't be adjusted automatically. Hence, we propose a new kind of modified transfer entropy by introducing h_x to fully consider the prediction horizon itself. The formula is as follows:

$$\begin{aligned}T\left(X\left|Y\right.\right)=&\sum_{x_{i+h},x_{i+h-h_x}^{(k)},y_i^{(l)}} p\left(x_{i+h},x_{i+h-h_x}^{(k)},y_i^{(l)}\right)\log\frac{p\left(x_{i+h}\left|x_{i+h-h_x}^{(k)},y_i^{(l)}\right.\right)}{p\left(x_{i+h}\left|x_{i+h-h_x}^{(k)}\right.\right)}\\=&\sum_{x_{i+h},x_{i+h-h_x}^{(k)},y_i^{(l)}} p\left(x_{i+h},x_{i+h-h_x}^{(k)},y_i^{(l)}\right)\log p\left(x_{i+h}\left|x_{i+h-h_x}^{(k)},y_i^{(l)}\right.\right)\\&-\sum_{x_{i+h},x_{i+h-h_x}^{(k)}} p\left(x_{i+h},x_{i+h-h_x}^{(k)}\right)\log p\left(x_{i+h}\left|x_{i+h-h_x}^{(k)}\right.\right)\\&=H\left(x_{i+h}\left|x\right.\right)-H\left(x_{i+h}\left|x,y\right.\right)\end{aligned} \quad (6)$$

$$\begin{aligned}H\left(x_{i+h}\left|x\right.\right)&=-\sum_{x_{i+h},x_{i+h-h_x}^{(k)}} p\left(x_{i+h},x_{i+h-h_x}^{(k)}\right)\log p\left(x_{i+h}\left|x_{i+h-h_x}^{(k)}\right.\right)\\&=-\sum_{x_{i+h_x},x_i^{(k)}} p\left(x_{i+h_x},x_i^{(k)}\right)\log p\left(x_{i+h_x}\left|x_i^{(k)}\right.\right)\end{aligned} \quad (7)$$

where h_x denotes the prediction horizon between x and x_{i+h}. Using formula (6), h can be fixed and the interval between x and x_{i+h} can be optimized by changing h_x.The parameters k,l,h and h_x in formula (6) should be determined. According to the study of Overbey [21, 22]and Nichols[23], for simplicity, the process can be regarded as first order Markov process, so $k=l=1$. h_x and h can be decided by maximizing the information entropy $H\left(x_{i+h}\left|x\right.\right)$ and $H\left(x_{i+h}\left|x,y\right.\right)$,respectively.

After the calculation, significance level test should be conducted to validate the causality. By creating the new time series $N_s=\left(X_{new},Y_{new}\right)$ [24] and calculating the transfer entropy of the new time series $\lambda_i=T^i_{y_{new}\to x_{new}}\left(i=1,\cdots N_s\right)$, the significance level is formulated based on the 3σ rule:

$$s_\lambda=\mu_\lambda+3\sigma_\lambda \quad (8)$$

where μ_λ and σ_λ are the means and variance of λ_i , respectively. If $T_{y\to x}$ is larger than s_λ , then y is the cause of x.

B. Causal network modeling based on process knowledge and the modified transfer entropy

The method of transfer entropy must meet the assumption that the process is static Markov process, that is to say, the current state at time t is decided by the states at time t-1 or t-k with certain k. The process dynamics should keep steady in an enough long time. However, this assumption is hard to meet. Therefore, only based on the data-driven transfer entropy, the results are not accurate. By combining superficial process knowledge with our proposed modified transfer entropy, the results would be credible.

In a large-scale process, the whole system is composed of multiple sub-systems. Each sub-system represents a specific unit, in which there are more than one process variables. Through process division based on superficial process knowledge, we can firstly model each of the sub-system using modified transfer entropy, and then the whole causal model can be established by combining these sub causal models. In division, all the process variables will be allocated into different sub-systems. The variables in each sub-system consist in the variables belonging to this sub-system, the inflow variables, the outflow variables and the variables belonging to its adjacent sub-system which have influence on it. After division, the variables should be processed according to the superficial process knowledge. The rules are listed as follows:

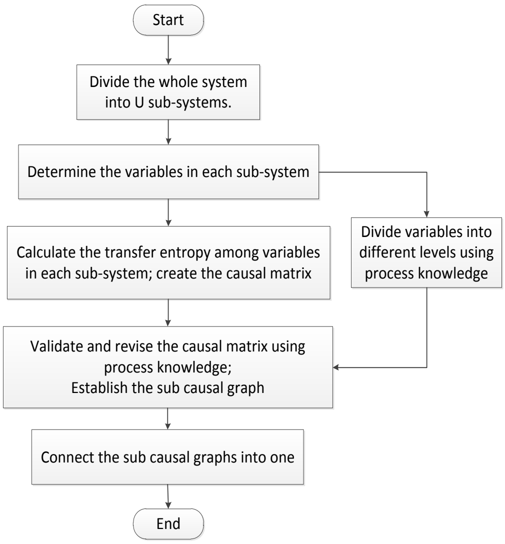

Figure 1 The flow chart of causal network modeling based on process knowledge and modified transfer entropy

(1) According to the inflow and outflow relationship, the inflow variables are defined as the root variables, and the outflow variables are defined as the leaf variables.

(2) Determine the related variables between adjacent sub-systems, such as the variables with same attribute or variables with obvious causality.

(3) Analyze the time order of each sub-system according to the process flow chart.

(4) Based on the time order, divided the related variable into high level and low level.

Using these rules, the variables are now endowed with the characteristics of input, output and time order. From the perspective of information flow, the input variables are the causes and the output variables are the effects; from the perspective of time order, the cause precedes effect. These findings from process knowledge can validate and revise the results of modified transfer entropy. Besides, system division largely decreases the computation complexity. The flow chart for causal network modeling based on process knowledge and modified transfer entropy is shown in Figure 1.

IV. Case study

This section provides the test of our proposed causal network modeling in two case studies, stochastic process and TE process, to verify its superiority and efficiency.

A. *Stochastic process*

The stochastic process is denoted by the formula (9) including both linear relationship and nonlinear relationship.

$$\begin{cases} Y_{k+1} = 3X_k^2 + 9X_k + 1.6\sqrt{\left|Z_k + 2.5\right|} + v_{1k} \\ Z_{k+1} = 0.3Z_k + 0.86\left(X_k - 2\right) + v_{2k} \end{cases} \tag{9}$$

where X_k, Y_k and Z_k are three random variables. $X_k \sim \left(0,1^2\right)$. Y_k is nonlinearly related with X_k and Z_k. Z_k is linearly related with X_k. v_{1k} and v_{2k} are white noise. Calculate the values of transfer entropy among these three variables according to formula (6) and the significance level according to formula (8). The results are shown in Table1.

Table 1 The values of transfer entropy and the correponding significan levels

	X_k	Y_k	Z_k
X_k	N/A	0.4184(0.4293)	0.2037(0.2905)
Y_k	1.2558(0.4922)	N/A	0.3225(0.2799)
Z_k	0.7986(0.3975)	0.3059(0.4330)	N/A

The value outside the bracket means the transfer entropy from the variable in the column to that in the row, and the value inside the bracket means the corresponding significance level. According to Table 1, the causality among these three variables is shown in Figure 2. It can be seen that X_k influences Z_k, and X_k and Z_k influence Y_k. The causality shown in Figure 2 is consistent with that in formula (8). It proves that our proposed causal network modeling method is accurate and efficient.

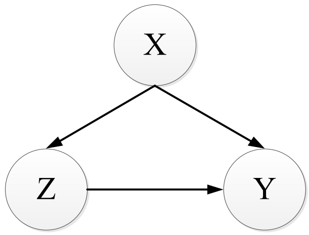

Figure 2 The causal network of stochastic process

B. *TE process*

Proposed by Downs and Vogel in 1993, Tennessee Eastman process is a simulation model of a real chemical process [25]. This model consists of five main units: Reactor, Condenser, Separator, Stripper and Compressor. The flow chart of TE process is shown in Figure 3. TE process contains 12 manipulated variables, 22 continuous process measurements, and 19 composition measurements. In our work, 22 continuous process measurements are chosen to model (see Table 2). The number of samples is 200 and the sampling time is 1.8s.

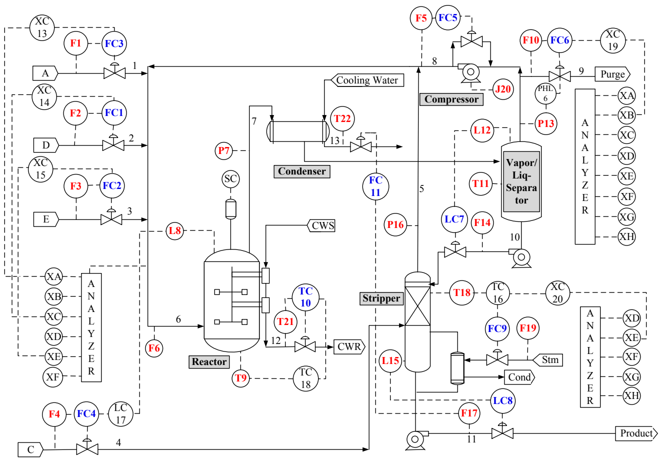

Figure 3 Flow chart of TE process

According to the flow chart, TE process can be divided into five sub units. Let $U = \{u_1, u_2, u_3, u_4, u_5\}$ represent the whole TE process. The variables in each sub unit are listed as follows:

Reactor $u_1=\{F_1, F_2, F_3, F_5, F_6, P_7, L_8, T_9, T_{21}\}$

Condenser $u_2=\{T_9, T_{22}\}$

Separator $u_3=\{P_7, L_8, T_9, T_{11}, L_{12}, P_{13}, F_{14}, T_{22}\}$

Stripper $u_4=\{F_4, T_{11}, L_{12}, P_{13}, F_{14}, L_{15}, P_{16}, F_{17}, T_{18}, F_{19}\}$

Compressor $u_5=\{F_5, F_{10}, T_{11}, L_{12}, P_{13}, J_{20}\}$.

Table 2 The 22 continuous process measurements

Variable Number	Variable Description	Variable Number	Variable Description
F_1	A feed	L_{12}	Separator level
F_2	D feed	P_{13}	Separator pressure
F_3	E feed	F_{14}	Separator underflow
F_4	A and C feed	L_{15}	Stripper level
F_5	Recycle flow	P_{16}	Stripper pressure
F_6	Reactor feed rate	F_{17}	Stripper underflow
P_7	Reactor pressure	T_{18}	Stripper temperature
L_8	Reactor level	F_{19}	Stripper steam flow
T_9	Reactor temperature	J_{20}	Compressor work
F_{10}	Purge rate	T_{21}	Reactor ooling water outlet T
T_{11}	Separator temperature	T_{22}	Condenser cooling water outlet T

After the division, the transfer entropy between any two variables in each sub unit should be calculated. The parameters h_x and h should be determined firstly. Taking variables F_5 and F_6 in u_1 for an example, based on formula (7), the trend of the transfer entropy from F_5 to itself with different interval is shown in Figure 4.

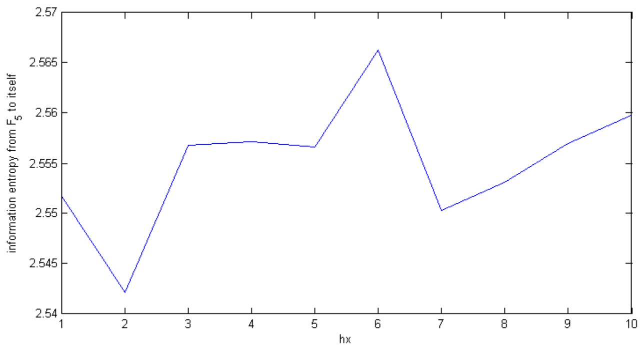

Figure 4 The transfer entropy of variable F_5 to itself

When $h_x=6$, $H(F_{5(i+h_x)}|F_{5(i)})$ reaches the maximum value. So h_x is set as 6. Then calculate the transfer entropy $T(F_6|F_5)$ based on the formula (6). The trend with different h is shown in Figure 5.

When $h=9$, $T(F_6|F_5)$ reaches the maximum. So h is set as 9. In the same way, the transfer entropy between any two variables in u_1 can be calculated. The causality between them can be measured using formula (4) and determined by comparing with its corresponding significance level. If $T_{y\to x}$ is larger than s_λ , y is the cause of x; otherwise, there is no causality form y to x. The causal matrix about u_1 is shown in Table 3. '1' denotes that the variable at the column is the cause of the variable at the row. '0' denotes that the variable at the column is not the cause of the variable at the row.

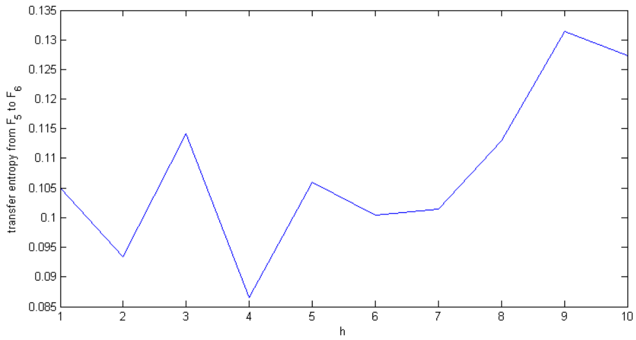

Figure 5 The transfer entropy from F_5 to F_6

Afterwards, the process knowledge is used to validate and revise the causal matrix. For sub unit u_1, the inflow variables are $\{F_1, F_2, F_3, F_5, F_6\}$. They are regarded as the root variables. The direction of influence should be from $\{F_1, F_2, F_3, F_5, F_6\}$ to $\{P_7, L_8, T_9, T_{21}\}$. Besides, the flow rate of feed A, D and E are independent. Based on these process knowledge, the optimized causal matrix of u_1 is shown in Table 4. The casual network of u_1 is shown in Figure 6.

Table 3 The causal matrix of u_1

Variable number	F_1	F_2	F_3	F_5	F_6	P_7	L_8	T_9	T_{21}
F_1	0	1	1	1	0	1	1	1	1
F_2	0	0	1	1	0	0	0	1	0
F_3	0	0	0	1	0	0	0	1	1
F_5	0	0	0	0	0	0	0	0	0
F_6	1	1	1	1	0	1	1	1	1
P_7	0	1	1	1	0	0	1	1	0
L_8	0	1	1	1	0	0	0	1	0
T_9	0	0	0	1	0	0	0	0	0
T_{21}	0	1	0	1	0	1	1	1	0

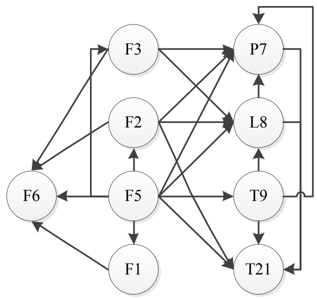

Figure 6 The causal network of sub unit u_1

Table 4 The optimized causal matrix of u_1

Variable number	F_1	F_2	F_3	F_5	F_6	P_7	L_8	T_9	T_{21}
F_1	0	0	0	1	0	0	0	0	0
F_2	0	0	0	1	0	0	0	0	0
F_3	0	0	0	1	0	0	0	0	0
F_5	0	0	0	0	0	0	0	0	0
F_6	1	1	1	1	0	0	0	0	0
P_7	0	1	1	1	0	0	1	1	0
L_8	0	1	1	1	0	0	0	1	0
T_9	0	0	0	1	0	0	0	0	0
T_{21}	0	1	0	1	0	1	1	1	0

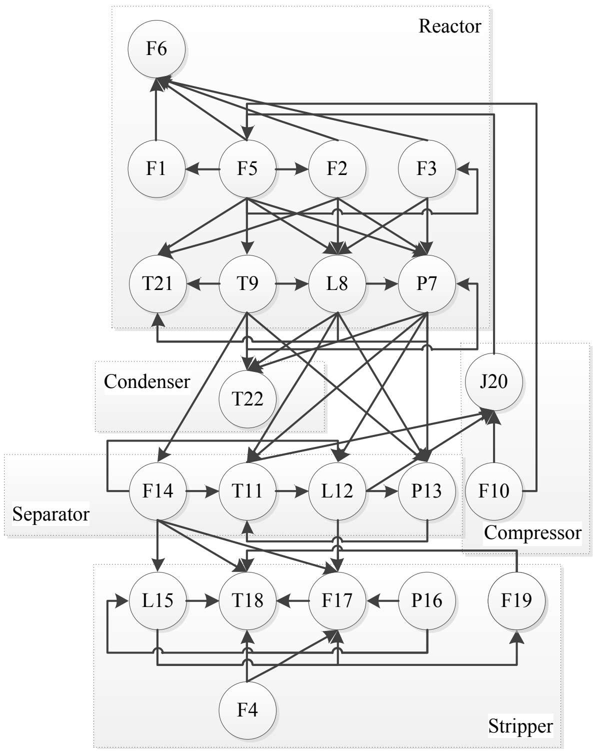

Figure 7 The causal network of TE process

Accordingly, the causal networks of the other four sub units are created. Combining these sub causal networks, the whole causal model of TE process can be established, see as Figure 7.

Compared with the traditional transfer entropy method, our proposed modified algorithm could find some causality which are hard to be captured by the traditional algorithm. That's because in our proposed algorithm, the prediction horizon h_x is introduced to fully consider the information transfer between variable itself. The comparison is shown in Table 5. Besides, the sub-system division based on process knowledge greatly decreases the computation complexity.

Table 5 The comparison between proposed algorithm and traditional transfer entropy

Causality	Traditional	Proposed
$F_1 \to F_6$	No	Yes
$F_2 \to F_6$	No	Yes
$F_3 \to F_6$	No	Yes
$F_5 \to F_6$	No	Yes
$F_6 \to L_8$	No	No
$T_9 \to T_{21}$	No	Yes
$F_{14} \to L_{15}$	No	Yes
$F_{17} \to L_{15}$	No	No
$F_{19} \to T_{18}$	No	Yes

V. Conclusions

A new kind of modified transfer entropy algorithm is proposed in this paper. The information transfer between variables themselves and different variables is fully taken into consideration by optimizing the prediction horizon. The discovering ability of causality is improved a lot. What's more, the superficial process knowledge is utilized to divide the whole process into multiple sub systems, which facilitates the calculation of transfer entropy among variables and validates the results of data-driven methods. The case studies of a stochastic process and TE process confirm the feasibility and effectiveness of our proposed scheme for causal network modeling.

References

[1] Qunxiong, Zhu, Gao Huihui, and Liu Feifei. "Research progress of alarm systems in the process industries," *Computers and Applied Chemistry*, vol. 31, Iss.2, pp.129-134, 2014
[2] Bauer, Margret, and Nina F. Thornhill. "A practical method for identifying the propagation path of plant-wide disturbances," *Journal of Process Control*, vol.18, Iss.7, pp. 707-719, 2008.
[3] Bressler S L, Seth A K. "Wiener-Granger causality: A well-established methodology," *NeuroImage*, vol 58, Iss. 2, pp. 323-329, 2011
[4] Yuan T, Qin S J. "Root cause diagnosis of plant-wide oscillations using Granger causality," *Journal of Process Control*, vol.24, Iss.2, pp. 450-459, 2014
[5] Friston K, Moran R, Seth A K. "Analysing connectivity with Granger causality and dynamic causal modeling,". *Current opinion in neurobiology*, vol 23, Iss.2, pp. 172-178, 2013.
[6] Dey S, Stori J A. "A Bayesian network approach to root cause diagnosis of process variations,". *International Journal of Machine Tools and Manufacture*, vol.45, Iss.1, pp.75-91, 2005.
[7] Alaeddini A, Dogan I. "Using Bayesian networks for root cause analysis in statistical process control," *Expert Systems with Applications*, vol.38, Iss. 9, pp. 11230-11243, 2011.
[8] Wee Y Y, Cheah W P, Tan S C, Wee k k. "A method for root cause analysis with a Bayesian belief network and fuzzy cognitive map," *Expert Systems with Applications*, vol.42, Iss.1, pp. 468-87, 2015.
[9] Gao, Huihui, Yuan Xu, and Qunxiong Zhu. "Spatial Interpretive Structural Model Identification and AHP-Based Multimodule Fusion for Alarm Root-Cause Diagnosis in Chemical Processes," *Industrial & Engineering Chemistry Research*, vol.55, Iss.12, pp. 3641-3658, 2016.
[10] Gao H, Xu Y, Gu X, et al. "Systematic rationalization approach for multivariate correlated alarms based on interpretive structural modeling and Likert scale," *Chinese Journal of Chemical Engineering*, vol. 23, Iss.12, pp. 1987-1996, 2015.
[11] Schreiber T. "Measuring Information Transfer," *Physical Review Letters*, vol. 85, Iss.2 , pp.461-464 ,2000.
[12] Choi H. "Localization and regularization of normalized transfer entropy," *Neurocomputing*, vol. 139, pp.408-414, 2014.

[13] Sensoy A, Sobaci C, Sensoy S, Alali F. "Effective transfer entropy approach to information flow between exchange rates and stock markets," *Chaos, Solitons & Fractals*, vol. 68, pp. 180-185, 2014.
[14] Bauer M., Cox J W, Caveness M H, Downs J J, Thornhill N F. "Finding the direction of disturbance propagation in a chemical process using transfer entropy," *IEEE Transactions on control systems technology*, vol.15 , Iss. 1, pp. 12-21,2007.
[15] Staniek M, Lehnertz K. "Symbolic transfer entropy," *Physical review letters*, vol. 100, Iss. 15, pp. 158101, 2008.
[16] Duan P, Yang F, Chen T, Shah S L. "Direct Causality Detection via the Transfer Entropy Approach," *IEEE Transactions on control systems technology*, vol. 21, Iss. 6, pp. 2052-2066, 2013.
[17] Yu W, Yang F. "Detection of Causality between Process Variables Based on Industrial Alarm Data Using Transfer Entropy," *Entropy*, vol. 17, Iss. 8, pp. 5868-5887, 2015.
[18] Shu Y, Zhao J. "Data-driven causal inference based on a modified transfer entropy," *Computers & Chemical Engineering*, vol. 57, pp. 173-180, 2013.
[19] Shannon C E. "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.
[20] Schreiber T, Kaiser A. "Information transfer in continuous processes," *Physica D*, vol. 166, pp. 43-62, 2002.
[21] Overbey L A, Todd M D. "Dynamic system change detection using a modification of the transfer entropy," *Journal of Sound and Vibration*, vol. 322, Iss. 1-2, pp. 438-453, 2009.
[22] Overbey L A, Todd M D. "Effects of noise on transfer entropy estimation for damage detection," *Mechanical Systems and Signal Processing*, vol. 23, Iss. 7, pp. 2178-2191, 2009.
[23] Nichols J M, Seaver M, Trickey S T, Todd M D, Olson C, Overbey L. "Detecting nonlinearity in structural systems using the transfer entropy," *Physical Review E*, vol. 72, Iss. 4, pp. 046217, 2005.
[24] Duan P, Yang F, Shah S L, Chen T. "Transfer Zero-Entropy and Its Application for Capturing Cause and Effect Relationship Between Variables," *IEEE Transactions on Control Systems Technology*, vol. 23, Iss. 3, pp. 855-867, 2015.
[25] Down J J, Vogle E F. "A plant-wide industrial process control problem," *Computers & Chemical Engineering*, vol. 17, Iss. 3, pp. 245-255, 1993.