A Novel Approach to Alarm Causality Analysis Using Active Dynamic Transfer Entropy

Yi Luo, Bhushan Gopaluni, Yuan Xu, Liang Cao, and Qun-Xiong Zhu

Ind. Eng. Chem. Res., Just Accepted Manuscript • DOI: 10.1021/acs.iecr.9b06262 • Publication Date (Web): 14 Apr 2020

Downloaded from pubs.acs.org on April 21, 2020

Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.
A Novel Approach to Alarm Causality Analysis Using Active Dynamic Transfer Entropy

Yi Luoabc, Bhushan Gopalunic, Yuan Xab, Liang Caoc, Qun-Xiong Zhub*

a.College of Information Science & Technology, Beijing University of Chemical Technology, Beijing, 100029, China;
b. Engineering Research Center of Intelligent PSE, Ministry of Education of China, Beijing 100029, China;
c. Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, BC, Canada;

* Corresponding author: Tel.: +86-10-64413467. Email addresses: buctielab@hotmail.com

ABSTRACT

Alarm flooding is a serious safety problem in the chemical process industries. Bayesian Networks are a set of powerful tools that can be used to trace the root-cause of alarms. For highly integrated complex chemical processes, we propose a Bayesian Network based on Active Dynamic Transfer Entropy (ADTE) to establish an accurate alarm propagation network during an alarm flood. The proposed method has two primary advantages: (1) it circumvents the false causality problem caused by strong correlations and therefore can be used to mine deeper alarm propagation paths like feedback loops. (2) It provides the time of origin of an alarm as it propagates through the process network, allowing operators to respond appropriately. The proposed method involves the following elements: modular segmentation, extraction of common cause variables, calculation of alarm propagation time between variables, calculation of ADTE, identification of an underlying alarm network and tuning of relevant parameters. The Tennessee Eastman Process (TEP) is used to demonstrate the validity and superiority of the proposed ADTE-based alarm causality method.
**Nomenclature**

\[ x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4} \]  
Continuous random variables

\[ x_{i_p} \]  
Parent variables of \( x_i \)

\[ x_c \]  
Common cause variables of \( x_i \)

\[ T_{x_p \rightarrow x_i} \]  
Transfer entropy from \( x_{i_p} \) to \( x_i \)

\[ T_{(x_p \rightarrow x_i)^{nor}} \]  
Normalized transfer entropy from \( x_{i_p} \) to \( x_i \)

\[ D_{x_p \rightarrow x_i} \]  
Direct transfer entropy from \( x_{i_p} \) to \( x_i \)

\[ ATE_{x_p \rightarrow x_i} \]  
Active transfer entropy from \( x_{i_p} \) to \( x_i \)

\[ ADTE_{x_p \rightarrow x_i} \]  
Active dynamic transfer entropy from \( x_{i_p} \) to \( x_i \) with propagation time \( h \)

1. INTRODUCTION

Modern industrial plants are highly integrated and extremely complex\(^1\). As such, it is of utmost importance to operate them safely and efficiently while avoiding catastrophic events that can lead to casualties, significant economic losses and environmental pollution. In the past two decades, the process modeling, process monitoring and control systems, such as Distributed Control System (DCS) and Supervisory Control and Data Acquisition (SCADA), have evolved to incorporate various software tools for this purpose\(^2,3\). As a result, the performance of industrial alarm systems has improved significantly. However, in practice, alarm flooding problem occurs due to an excessive number of alarms\(^4,7\). A typical alarm flooding problem includes invalid and repeated alarms, making it difficult for operators to prioritize and separate critical alarms. Alarm flooding problem can be resolved by finding the
root-cause of alarms from many correlated alarms using causality analysis and fault
diagnosis methods\textsuperscript{8,9}.

Causality analysis is a powerful approach often used to trace the root-cause of alarms and
establish the causal network of alarms in complex chemical processes. When an alarm flood
occurs, causality analysis can be used to provide operators with approximate information
about the root-cause and therefore ensure the safety of underlying processes\textsuperscript{8,9}.

In the research of causal analysis of industrial processes, a causal network of alarms can be
built by symbol directed graph (SDG) using a large amount of expert knowledge. These
methods have been applied to industrial problems with some success\textsuperscript{10,11}. However, in many
practical problems, the expert knowledge is limited and therefore it is difficult to establish a
causal network that covers all existing relationships between process variables. With the
development of the SCADA and the DCS systems, large amounts of data are measured and
stored during process operations. These incredible volumes of data facilitate the studies of
data-driven methods. There are several effective methods for finding relationships between
process variables including interpretable structural models\textsuperscript{12,13}, Bayesian networks\textsuperscript{14-17},
Granger causality analysis\textsuperscript{18-20} and transfer entropy\textsuperscript{20-22}. For Granger causality analysis and
transfer entropy, the latter approach is known to provide information about the causal
networks especially when the data are limited\textsuperscript{23}. However, both of these methods generally
exhibit similar performance when there are enough data\textsuperscript{20,23}. In a focused study to solve the
alarm flood problem, Hu has made noteworthy contributions about the cause-effect
between variables by transfer entropy\textsuperscript{22,24}. Similar to SDG, the Bayesian Network (BN) based
approaches can infer the causality among variables using prior expert knowledge. Moreover,
the Bayesian Network approach is more suitable than SDG for extracting the causal relationships as they provide additional information in the form of probabilistic expressions. Granger causality analysis and transfer entropy are two powerful tools for establishing pairwise causality of variables and therefore have been used in conjunction with Bayesian Networks.

The research on using Bayesian Networks for alarm causality analysis includes two subareas: parameter training and structure learning\textsuperscript{24,25}. In parameter training, the underlying structure of the Bayesian network is fixed and the relevant parameters are estimated using data. It is possible to use either raw data or extracted features using Principal Component Analysis (PCA) to train the Bayesian Network parameters\textsuperscript{26}. For chemical processes with nonlinear data, a kernel PCA approach\textsuperscript{27} is generally used on the raw data before training the parameters of the Bayesian Network. In structure learning, the objective is to construct the structure of Bayesian Network as is done by “scoring and structure” search methods\textsuperscript{14}. The basic premise of these scoring algorithms is to quantify the accuracy of a given Bayesian Network using a score. The more accurate the scoring calculation and the more efficient the structure search algorithms are, the better the reconstruction of Bayesian Networks. Two widely used scoring metrics are Bayesian information criterion (BIC)\textsuperscript{28} and the Bayesian Dirichlet Equivalence (BDE)\textsuperscript{29}. Antoniak\textsuperscript{30} proposed a Bayesian Network construction model combining a hybrid search method and the BDE score criterion. By adding a penalty to the complexity of the model using the BIC score, a more simplified model is obtained. However, BIC and BDE criteria are often not good enough to provide acceptable scoring results in complex chemical processes. Hence, the entropy value is used as an alternative scoring
criterion to estimate the Bayesian Network structure in these processes. Meng established a criterion called “family transfer entropy probability” and verified the effectiveness of using transfer entropy based on the score and structure search method.

Methods that rely on causality analysis such as Granger causality, transfer entropy and cross-correlation functions simply provide correlation information and are therefore not reliable in constructing the underlying process structures. The alarm networks are often constructed by these methods focusing on the similarity of the alarm and get false causal results. To this end, Duan proposed the Direct Transfer Entropy (DTE) method to analyze whether the relationship between process variables is direct or indirect. Duan also proposed the Transfer zero-entropy method to analyze the dataset which does not follow a well-defined distribution. In order to further solve the causal propagation of alarms, not just the similarity of alarms. In this paper, we propose a multi-blocks Bayesian Network model based on Active Dynamic Transfer Entropy (ADTE). The industrial process can be divided into several blocks which include several variables in each block. The purpose of multi-blocks segmentation is to adapt to the complexity of chemical data and to ensure accuracy, unlike the traditional Bayesian Networks based on transfer entropy. The ADTE-based multi-blocks Bayesian network model focuses on finding the causality relationship rather than only the alarm similarity relationship, which has important implications in the modeling of alarm systems.

The main contribution of this paper is an algorithm to identify the propagation of alarms through a complex process network using the concepts of Active Dynamic Transfer Entropy and Bayesian Networks. The operators can use the path of alarm propagation to mitigate
alarm flooding. In addition, we propose the concepts of alarm propagation time and alarm
propagation structures. With the proposed approach, operators can not only identify the
root cause for alarm flooding but also predict the variables to which alarms will spread in
future.

This article is organized as follows: the traditional Bayesian Networks, DTE and block
segmentation methods for chemical processes will be introduced in Section 2; In Section 3,
an approximation of ADTE will be introduced. In addition, the calculation of alarm
propagation time and the construction of alarm propagation structure with time information
will be introduced in detail. In Section 4, the TEP benchmark is used to illustrate the
advantages of the proposed method. Conclusions are presented in Section 5.

2. PRELIMINARIES

In this section, the traditional Bayesian Network will be introduced and the BDE, BIC and TE
criteria for scoring a structure will be briefly compared. In addition, this section will also
introduce the basic DTE methods and the block segmentation method for chemical
processes. These methods are then used in developing the ADTE-based multi-blocks
Bayesian Network.

2.1 Bayesian Network

A Bayesian Network shows the relationship between different process variables and the
strength of those relationships by using graph models and probability. Bayesian Networks
are widely used in chemical processes. Several data-driven methods can be used to
construct a Bayesian Network among which the score and structure search approach is
better than other methods.\(^{34}\)
Let us consider a process structure with variables $x_1$ to $x_n$ connected as shown in Figure 1. The measurement of a variable $x_i$ at time $t = l$ is denoted by $x_i^l$. Given measurements of the $n$ variables from time $t = 1$ to $t = m$, we define the data matrix $X \in R^{m \times n}$ as follows:

$$X = \begin{bmatrix}
  x_1^1 & x_2^1 & \cdots & x_n^1 \\
  x_1^2 & x_2^2 & \cdots & x_n^2 \\
  \vdots & \vdots & \ddots & \vdots \\
  x_1^m & x_2^m & \cdots & x_n^m
\end{bmatrix}$$

If a connection exists from $x_1$ to $x_2$, then $x_1$ is referred as parent variable, and $x_2$ is referred as child variable, and the transfer correlation is from $x_1$ to $x_2$. All the parent variables of $x_i$ are represented by $x_{ip}$.

In addition, we also define an adjacent matrix $A$ with elements $a_{ij}$,

$$A = \begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix} \in R^{n \times n}$$

where $i$ refers to the $i^{th}$ variable $x_i$ and $j$ refers to the $j^{th}$ variable $x_j$. If a direct connection exists between $x_i$ and $x_j$ then $a_{ij} = 1$ otherwise $a_{ij} = 0$. The matrix $A$ simply provides information about the existence of connections between different variables.

Using a new matrix $\Theta$ we capture the strength of these connections, which can be
measured using a variety of metrics such as BIC, BDE, and TE, etc.

\[ \Theta = \begin{bmatrix} \theta_{11} & \theta_{12} & \ldots & \theta_{1n} \\ \theta_{21} & \theta_{22} & \ldots & \theta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{n1} & \theta_{n2} & \ldots & \theta_{nn} \end{bmatrix} \in \mathbb{R}^{n \times n} \]

With the above definitions, any given process structure can be characterized by the tuple 
\[ G = \{X, A, \Theta\} \] where \( G \) represents the corresponding structure and the relevant information. Given a set of structures \( G^{all} = \{G_1, G_2, \ldots, G_r\} \), for every structure \( G_s \), we define a score which is calculated by one of the scoring functions in Equations. 2, 3 or 5.

The class of Score and Search (SS) methods are based on the data matrix \( X \) and score for each structure \( G_s \in G^{all} \). These methods choose the structure \( G_{max} \) with the highest score as the optimal structure. The evaluation criterion is shown as follows,

\[ G_{max} = \arg \max_{G_s \in G^{all}} \text{score}(G_s) \]  \hspace{1cm} (1)

where \( \text{score}(G_s) \) represents the score for structure \( G_s \). The score is determined using criteria such BIC, BDE and TE. BIC and TE focus more on the information flow of Bayesian structure. While the optimal structure obtained by BIC is often relatively simple, TE on the other hand often captures more accurate nonlinear relationships. The score using BIC for structure is defined as follows,

\[ BIC(G_s) = \sum_{i=1}^{n} \sum_{k=1}^{q_i} \left( p_{ijk} \log \left( \frac{p_{ijk}}{p_{ij}} \right) \right) - \frac{1}{2} \log(m) \sum_{i=1}^{n} ((r_i - 1)q_i) \] \hspace{1cm} (2)

where \( q_i \in R^+ \), and \( r_i \in R^+ \) represent the number of parent variables and the number of directed edges of each child variable \( x_i \), respectively. \( n \) and \( p_{ij} = \sum_{k=1}^{r_i} p_{ijk} \) represent the number of variables and the probability of different states \( k \) of the edge \( A_{ij} \), respectively. \( p_{ijk} \) means the probability of the \( i^{th} \) variable being in the \( k \)th state while the
parent variable is in the $j^{th}$ state.

The TE\textsuperscript{25} criterion for evaluation is as follows,

\[
TE(G) = \sum_{i=1}^{n} \left( T_{x_{ij}\rightarrow x_i} - \lambda \log (q_i (r_i - 1)) \right)
\]

where $q_i \in R^+$ and $r_i \in R^+$ represent the number of parent variables and the number of directed edges of each child variable $x_i$, respectively. $\lambda$ is a scalar parameter used to penalize the model complexity. And $T_{x_{ij}\rightarrow x_i}$ represents the transfer entropy value of the directions from $x_{ij}$ to $x_i$. The TE score is calculated as follows,

\[
T_{x_{ij}\rightarrow x_i} = \sum_{x_{ij}, x_i, x_j} p(x_{ij}^{t+1}, x_i^t, x_j^t) \log_2 \frac{p(x_{ij}^{t+1} | x_i^t, x_j^t)}{p(x_{ij}^{t+1} | x_j^t)}
\]

Since there always exists transfer entropy value between any two variables, a significance test is needed to determine whether a transfer relationship does exist between variables.

Therefore, the 95% significance test is used for further screening the effective relationships in this paper, which are described in Yu\textsuperscript{21} and Hu's\textsuperscript{22} papers. Unlike BIC and TE algorithms based on information flow, BDE\textsuperscript{29} calculates the prior probability between variables based on the sampled data, and establishes a model that maximizes the posterior probability of the data, as shown below:

\[
BDE(G) = \sum_{i=1}^{n} \sum_{j=1}^{n} \left[ \log \frac{\Gamma (a_{ij})}{\Gamma (a_{ij} + m_{ij})} + \sum_{k=1}^{r_i} \log \frac{\Gamma (a_{ij} + m_{jk})}{\Gamma (a_{ij})} \right] + \log (d_{G_i})
\]

where $a_{ij} = \sum_{k=1}^{r_i} a_{jk}$ represents the prior probability between variables, and $d_{G_i}$ represents the number of directed edges in the structure $G_i$. $\Gamma$ represents the Gamma function. Compared with the Bayesian network construction based on expert knowledge, the above three structure construction methods have better performance in practical
applications.

2.2 Direct Transfer Entropy

DTE is an extended version of TE that is used to determine whether the information transfer between variables is direct or indirect. The calculation of TE is shown in Equation.4. The TE is a measure of the prediction accuracy of $x_{t+1}$ using $x'_t$ without the information of $x'_t$. A large TE implies that $x_p$ can predict $x_i$. However, the TE method cannot determine whether there are other variables exists in the propagation path of $x_p$ to $x_i$.

The DTE is calculated as follows:

$$D_{x_p \rightarrow x_i} = \int \left[ p(x_{t+1} | x_t, x_j, x_p) \right] \log \frac{p(x_{t+1} | x'_t, x'_j, x'_p)}{p(x_{t+1} | x'_t, x'_j)} dv$$

where $v$ denotes the random vector $[x_{t+1}, x'_t, x'_j, x'_p]$. In the DTE method, we analyze the information of $x_{t+1}$ that is obtained from $x'_t$, but without the information from $x'_t$ and $x'_j$, where $x_j \in X$ is the middle variable. Similarly, DTE also needs to be tested for significance. If the value of DTE is significant, it indicates that $x_p$ can directly predict $x_i$; if the value of DTE is not significant, but the value of TE is significant, it indicates that the information transferring from $x_p$ to $x_i$ occurs through the variable $x_j$. This observation helps in further analysis of the relationship between variables in the structure.

2.3 System Blocks Segmentation

Considering that typical chemical plants have many units, we can reduce the workload of building Bayesian structure and improve the accuracy of the structure through block segmentation. In the actual chemical process, we can utilize priori process knowledge to divide the chemical process into multiple blocks according to following four steps:
Step 1: The overall chemical process is roughly divided into existing critical process units, such as reactors, condensers etc.

Step 2: According to the expert knowledge of variable correlation and process information such as energy flow or control loop, variables which are not well classified are added to each sub-block;

Step 3: Define the variables connecting adjacent sub-blocks as the associated variables of the two sub-blocks, and these associated variables are used in both adjacent sub-blocks for Bayesian network structure learning;

Step 4: Connect multiple sub-blocks through associated variables between two sub-blocks to restore the structure of the complete chemical process.

3. THE PROPOSED METHODS

Using the definitions from the previous section, we propose an ADTE-based multi-blocks Bayesian network and use this method to construct a Bayesian network structure, which describes the causal relationships between process variables in a real plant. The available causal analysis methods are prone to several challenges such as two variables having high similarity but no real causal relationship. Traditional methods such as Granger causal analysis, Transfer entropy, and Bayesian networks are not effective at separating variables with high similarity from those with causal relationships. To solve this problem, we introduce the ADTE-based multi-blocks Bayesian network in detail below. This method is less prone to false causality relationships but identifies the causal networks with better accuracy. In addition, the obtained alarm propagation paths can be used to identify the source of alarm flooding.
3.1 Identify Common Cause Variables

In the process of constructing a Bayesian network, it is important to identify common cause variables to minimize the risk of falsely assigning causal relationships between variables. As shown in Figure 2, assuming there are four variables \( \{x_i, x_j, x_y, x_o\} \) in chemical process, and two transfer relations \( x_i \rightarrow x_o \) and \( x_j \rightarrow x_y \), then there will be higher probability that similar data trends exists between \( x_j, x_y \) and \( x_i, x_o \). As a result, the correlation between \( x_j \) and \( x_o \) is strong. We define \( x_i \) as a common cause variable of \( x_j, x_o \). Therefore, we can determine whether there is a common cause variable on the propagation path according to the transfer entropy value. Since the value of the transfer entropy will not be 0 in the actual process, it is necessary to determine whether there is a common cause variable through a significance test by the following criterion,

\[
T_{(x_j \rightarrow x_o)^{nor}} = T_{x_j \rightarrow x_o} - T_{(x_j \rightarrow x_o)^{nor}} \tag{7}
\]

where \( T_{(x_j \rightarrow x_o)^{nor}} \) represents the threshold of significant test \(^{31, 32}\), \( T_{(x_j \rightarrow x_o)^{nor}} \) represents the normalized transfer entropy value. All variables \( x_{j_p} \) that satisfy Equation 8 are considered as common cause variables for the propagation paths from \( x_j \) to other variables.

\[
T_{(x_{j_p} \rightarrow x_j)^{nor}} > 0 \tag{8}
\]

where \( T_{(x_{j_p} \rightarrow x_j)^{nor}} \) represents the transfer entropy value from \( x_{j_p} \) to \( x_j \) after normalized by Equation 7, if \( T_{(x_{j_p} \rightarrow x_j)^{nor}} > 0 \), it means the transfer relationship from \( x_{j_p} \) to \( x_j \) is significant, \( x_j \) has common cause variables and we should remove them when analyzing the causality from \( x_j \) to \( x_o \). The reason for this judgment is that when calculating the active transfer entropy of \( x_j \) to \( x_o \), it is necessary to ensure that \( x_j \) does not receive information from other variables, and the information received by \( x_y \) has no effect on the
active transfer entropy calculation of $x_j$ to $x_o$.  

![Diagram](image)

Fig.2 Common cause variable interpretation

where $x_i$ is the parent variable of $x_j$ and $x_y$ is the intermediate variable of $x_j$ and $x_o$.  

When calculating the active transfer entropy from $x_j$ to $x_o$, $x_i$ is regarded as a common cause variable, but there is no need to regard $x_y$ as a common cause variable. If the transfer path is from $x_j$ to $x_o$, instead of from $x_j$ to $x_y$ and then to $x_o$. When calculating the path $x_j$ to $x_o$, $x_j$ is regarded as the parent variable of $x_y$ and also the common cause variable of $x_y$ and $x_o$, and the path from $x_y$ to $x_o$ will generate a smaller transfer entropy value compared with the path $x_j$ to $x_y$ and then to $x_o$.

### 3.2 Active Transfer Entropy

In the alarm system, the data is divided into alarm data and normal data, wherein the alarm data is divided into high alarm data and low alarm data. In this paper, the alarm data of $X_i$ at the sample time $t$ is defined as $A'_i \in R$, which is generated by the alarm system is discretized as follows:

$$
a'_i = \begin{cases} 
1, & x'_i < A'_i \\
2, & A'_i < x'_i < A'_i \\
3, & x'_i > A'_i 
\end{cases} \tag{9}
$$

Where $A'_i \in R$ and $A'_i \in R$ represent the low threshold and high threshold of $x_i$ in the process. For discrete data, the discretized ATE can be defined as follows:
\begin{equation}
ATE_{x_p \rightarrow x_i} = \sum_{x_i^{t+1}, x_i', x_i^{t-k}, x_i^t} p(x_i^{t+1}, x_i', x_i^{t-k}, x_i^t) \log_2 \frac{p(x_i^{t+1} \mid x_i', x_i^{t-k}, x_i^t)}{p(x_i^{t+1} \mid x_i', x_i^t)}
\end{equation}

where \( x_i \) is the common cause of \( x_i \) obtained by Equation.7 and Equation.8.

Equation.10 not only can accurately calculate the transfer entropy of \( x_p \) to \( x_i \) without the influence of common cause \( x_i \), but also indicate the probability of causing \( x_i \) to generate an alarm if an alarm occurs in \( x_p \). Like DTE and TE, ATE also needs to be tested for significance.

### 3.3 Active Dynamic Transfer Entropy

In the actual chemical process, the transfer of information between variables takes a certain amount of time. The proposed approach is also capable of extracting information related to propagation time in the structure. The active propagation time between variables by the following equations,

\begin{align}
ATE_{x_p \rightarrow x_i \rightarrow x_i' \rightarrow x_p} &= \sum_{x_i^{t+1}, x_i', x_i^{t-k}, x_i^t} p(x_i^{t+1}, x_i', x_i^{t-k}, x_i^t) \log_2 \frac{p(x_i^{t+1} \mid x_i', x_i^{t-k}, x_i^t)}{p(x_i^{t+1} \mid x_i', x_i^t)} \\
ATE_{x_p \rightarrow x_i' \rightarrow x_i \rightarrow x_p} &= \sum_{x_i^{t+1}, x_i', x_i^{t-k}, x_i^t} p(x_i^{t+1}, x_i', x_i^{t-k}, x_i^t) \log_2 \frac{p(x_i^{t+1} \mid x_i', x_i^{t-k}, x_i^t)}{p(x_i^{t+1} \mid x_i', x_i^t)} \\
ATE_{x_p \rightarrow x_i \rightarrow x_i' \rightarrow x_p} &= \sum_{x_i^{t+1}, x_i', x_i^{t-k}, x_i^t} p(x_i^{t+1}, x_i', x_i^{t-k}, x_i^t) \log_2 \frac{p(x_i^{t+1} \mid x_i', x_i^{t-k}, x_i^t)}{p(x_i^{t+1} \mid x_i', x_i^t)}
\end{align}

Equation.11 represents the amount of information that \( x_i \) flows from \( x_p \) is more than that \( x_i \) is transmitted from \( x_p \), and Equation.12 represents the amount of information that \( x_i \) flows from \( x_p \) is more than \( x_i \) flows from \( x_p \). Due to the detection asymmetry of the information entropy, we need to confirm that the prediction performance of \( x_p \) to \( x_i \) is improved compared with the prediction performance of \( x_p \) to \( x_i \).
subtract Equation.12 from Equation.11 to obtain the DTE value in Equation.13. If the DTE value is significant at the time $h$, it proves that $x_{i_p}^{t-h}$ is more suitable for predicting $x_i$.

Since the chemical process is unstable and periodic, there may be multiple $h$ corresponding DTE values are significant on the path $x_{i_p}$ to $x_i$. Combine the results in Equation.13 with Equation.10, the ADTE for different propagation times is calculated as follow.

$$ADTE_{x_{i_p}^{t-h} \rightarrow x_i} = \sum_{x_{i_p}^{t-h}, x_i'} p(x_i^{t+1}, x_i', x_{i_p}^{t-h}) \log \frac{p(x_i^{t+1} | x_i', x_{i_p}^{t-h})}{p(x_i^{t+1} | x_i', x_{i_p}^{t-h})}$$  \hspace{1cm} (14)$$

This paper retains all ADTE values corresponding to DTE values that are significant, which is useful in the later structural search process.

3.4 Structure Searching

After obtaining the transfer entropy values for each propagation time, we propose a structure search method using the greedy search algorithm, and the score of each structure is given as follow:

$$Score_{ADTE}(G_{greedy}) = \sum_{ed} (ADTE_{ed} - \lambda_1 \log(h_{ed})) - \lambda_2 \log(n)$$  \hspace{1cm} (15)$$

where $Score_{ADTE}(G_{greedy})$ represents the ADTE-based score of each structure searched by greedy algorithm. $n$ represents the number of directed edges in the structure obtained by greedy search, and $\lambda_1$ and $\lambda_2$ are two penalty scale parameters, respectively. The penalty coefficients are used to limit the complexity of the structure and reduce the effect of the long propagation time caused by periodic data. $ADTE_{ed}$ and $h_{ed}$ represent the ADTE value and the propagation time of the $ed^{th}$ direction edge in each structure $G_{greedy}$. After scoring each structure, according to Equation.16, the highest-ranking structure is obtained as the final Bayesian network structure.
\[ G_{\text{max}} = \arg \max_{G \in G^m} \text{score}_{\text{ADTE}}(G_{\text{proof}}) \]  \hspace{1cm} (16)

### 3.5 Structure Integration

After obtaining the Bayesian network structure of each sub-block, we add virtual variables to integrate different sub-blocks.

In Figure 3, it takes one sampling time for the alarm of \( x_i \) to effect \( x_j \), and two sampling times for the alarm in \( x_j \) to effect \( x_o \). In the parameter training process, the data \( x_i', x_j', x_o' \), and \( x_i'^{+1}, x_j'^{+1}, x_o'^{+3} \) should be included. However, it requires two sampling times from \( x_j \) to \( x_o \), which is same as the time it takes for \( x_j \) to effect \( x_o \). In this case, we use the concept of virtual variable to achieve structural integration, as shown in Figure 4.

Adding the virtual variable significantly improves the accuracy of the alarm propagation structure. However, there is an additional computational cost incurred due to the increased number of variables.

### 3.6 Alarm Source Tracing Strategy

After establishing the structure and training the variable parameters, a simple alarm source
tracing strategy can be developed as shown in Equations.17 and Equations.18:

\[ P_{\text{single},i} = \sum_{j=1}^{k} P_j / k \]  

\[ P_{\text{multi}} = \sum_{i=1}^{n} P_{\text{single},i} \]  

Equation.17 represents a single alarm source tracing strategy, where \( k \) represents the number of all variables which the alarm path has passed through, and \( P_j \) represents the alarm probability of variable \( x_j \) in structure \( G_{\text{max}} \). \( P_{\text{single},i} \) represents the score of alarm propagation path of alarm \( x_i \). When multiple alarms occur simultaneously, according to Equation.18, the scores corresponding to each path of the alarm root variable are calculated. The alarm root variable is determined by the largest \( P_{\text{multi}} \) of the multi-alarm propagation path and \( n_a \) represents the number of alarm variables. The overall alarm root analysis step is shown in Figure.5.

**Fig.5 Alarm source tracing step of Bayesian network based on ADTE**

**4. SIMULATION**

This section illustrates the importance of eliminating false causality during the construction of alarm structures based on Tennessee Eastman Problem (TEP)\(^{35-38}\). We use BIC, BDE, and
MBTE methods to compare with and further demonstrate the superiority of the proposed method.

The TEP process consists of 41 measured variables XMEAS (1-41) and 12 control variables XMV (1-12) and the variable distribution and variable information are shown in Figure.6 and Table.1, respectively. In the experimental data set of TEP, there are 21 types of IDV (1-21) faults and each fault contains 960 samples. In this paper, the data set of IDV1 is used to verify the effectiveness of proposed method. IDV1 is a step fault and the root cause of this fault is FF4, which represents the change of the A/C feed ratio.

![Fig.6 The process of Tennessee Eastman](image)

<table>
<thead>
<tr>
<th>Variable number</th>
<th>Variable symbol</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF1</td>
<td>XMEAS(1)</td>
<td>A feed (stream 1)</td>
</tr>
<tr>
<td>FF2</td>
<td>XMEAS(2)</td>
<td>D feed (stream 2)</td>
</tr>
<tr>
<td>FF3</td>
<td>XMEAS(3)</td>
<td>E feed (stream 3)</td>
</tr>
<tr>
<td>FF4</td>
<td>XMEAS(4)</td>
<td>Total feed (stream 4)</td>
</tr>
</tbody>
</table>
4.1 Causality Structure learning

In this section, causality structure learning will be divided into four parts: sub-blocks segmentation and variable selection, ADTE value calculation, optimal structure extraction and virtual variable expansion.

4.1.1 Sub-blocks Segmentation and Variable Selection

23 common variables in Table 1 are divided into three sub-blocks by the segmentation principle in section 2.3. The sub-blocks and the associated variables between adjacent sub-blocks are given in Table 2:

<table>
<thead>
<tr>
<th>Sub-block</th>
<th>Units</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub1</td>
<td>feed, reactor, condenser,</td>
<td>FF1, FF2, FF3, FF4, RF6, RF7, RL8, RT9, RT21,</td>
</tr>
<tr>
<td></td>
<td>compressor, venting</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CF23, CF5, CW20</td>
</tr>
<tr>
<td>Sub2</td>
<td>Condenser, compressor,</td>
<td>CF23, CF5, CW20</td>
</tr>
<tr>
<td></td>
<td>Venting, separator</td>
<td>PR10, ST11, SL12, SP13, SF14,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ST22</td>
</tr>
</tbody>
</table>
After dividing the TEP into three sub-blocks, we select alarm variables to establish the causality structure. Table 3 shows the alarm probability of each variable in normal case and IDV1 abnormal case.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Alarm probability under normal condition</th>
<th>Alarm probability under IDV (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF1</td>
<td>1.63</td>
<td>92.09</td>
</tr>
<tr>
<td>FF2</td>
<td>0.93</td>
<td>2.56</td>
</tr>
<tr>
<td>FF3</td>
<td>4.53</td>
<td>28.95</td>
</tr>
<tr>
<td>FF4</td>
<td>0.35</td>
<td>89.07</td>
</tr>
<tr>
<td>CF5</td>
<td>0.47</td>
<td>0</td>
</tr>
<tr>
<td>RF6</td>
<td>0.23</td>
<td>4.19</td>
</tr>
<tr>
<td>RP7</td>
<td>1.05</td>
<td>31.51</td>
</tr>
<tr>
<td>RL8</td>
<td>0.7</td>
<td>23.26</td>
</tr>
<tr>
<td>RT9</td>
<td>0.7</td>
<td>1.28</td>
</tr>
<tr>
<td>PR10</td>
<td>0.47</td>
<td>30.07</td>
</tr>
<tr>
<td>ST11</td>
<td>0.93</td>
<td>27.56</td>
</tr>
<tr>
<td>SL12</td>
<td>0.7</td>
<td>0.35</td>
</tr>
</tbody>
</table>

In the data pre-processing step, variables which alarm probability do not change substantially are unrelated to IDV1 fault and will not be considered in this simulation.

Through the information in Table 3, the sub-blocks of Table 2 can be simplified, as shown in Table 4. The later simulation in this paper is based on the sub-blocks division results in Table 4.

### Tab4. The optimal sub-blocks of TE process

<table>
<thead>
<tr>
<th>Sub-block</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub1</td>
<td>FF1, FF2, FF3, FF4, RF6, RP7, RL8, RT21, CW20</td>
</tr>
<tr>
<td>Sub2</td>
<td>CW20, PR10, ST11, SP13, ST22</td>
</tr>
<tr>
<td>Sub3</td>
<td>ST11, SP13, ST22, SP16, ST18, SF19</td>
</tr>
</tbody>
</table>

### 4.1.2 ADTE

After obtaining the variables in each sub-block, we calculate the potential common cause
variables. To this end, the simulation results of Sub1 will be described in detail, and similar results were obtained using Sub2 and Sub3. Calculated by the method proposed in section 3.1, the results of the nine variables in Sub1 are shown in Figure 7.

![Fig.7 Potential common cause variables analysis](image)

The transfer relationship between the variables can be seen from Figure 7. The white area represents that the value of the transfer entropy between the variables has not passed the significance test. The red area represents that there is a transfer relationship between the variables, and the deeper the color, the stronger the transfer relationship. Further, the potential common cause variables of each variable in sub1 are listed in Table 5.

<table>
<thead>
<tr>
<th>Variable</th>
<th>The potential common cause variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF1</td>
<td>RP7,RT21</td>
</tr>
<tr>
<td>FF2</td>
<td>RP7,CW20</td>
</tr>
<tr>
<td>FF3</td>
<td>FF4</td>
</tr>
<tr>
<td>FF4</td>
<td>RP7,RL8</td>
</tr>
<tr>
<td>RF6</td>
<td>FF2,RL8,CW20</td>
</tr>
<tr>
<td>RP7</td>
<td>RL8,RT21</td>
</tr>
<tr>
<td>RL8</td>
<td>None</td>
</tr>
<tr>
<td>RT21</td>
<td>FF2,CW20</td>
</tr>
<tr>
<td>CW20</td>
<td>FF2,RF6,RL8,RT21</td>
</tr>
</tbody>
</table>

The information related to potential common cause variables is brought into ADTE for
calculation. Considering the computational complexity of TEP, the maximum alarm propagation time between variables is defined as ten samples. Figure 8 shows the relationship strength between variables according to the ADTE values at different propagation time, the maximum time lag is settled as 8 for calculation. In Figure 7, the white area represents that the value of the transfer entropy between the variables has not passed the significance test. The deeper color in Figure 8, the stronger causality.

![Fig.8 The ADTE values of Sub1 at different propagation time](image)

It can be found that the same path between two variables may correspond to different propagation times, where RL8 to FF4 have five choices for propagation time, i.e. 2, 3, 5, 6 and 8. The reason for this phenomenon is the periodic characteristics of some variables or control loops in the chemical process. All this information should be preserved. Another advantage is that the ADTE can find relationships that were missed by previous studies. For example, the path from FF3 to CW20 does not exist in the initial data set, but there is a causal relationship when the propagation time is three, four, five, and six sample times. In the structure search step, the proposed method will automatically select the optimal
propagation time to build the Bayesian network structure based on Equation 16 and Equation 17. After obtaining the ADTE value between the variables, it will search the optimal structure through the method in section 3.4, in which the time penalty term $\lambda_1 = 0.17$ and the structural complexity penalty term $\lambda_2 = 0.01$. The result of traditional BIC, BDE, MBTE and ADTE are compared with Sub1 data, as shown in Figure 9-12.
The black line indicates that the relationship identified by data is consistent with the actual process, the dotted line indicates that the relationship identified from data is inconsistent with the actual process, and the red line indicates an important variable relationship that is not denied from the data. Compared to the structures obtained by BIC and BDE, MBTE and ADTE have significant advantages. However, the structure of MBTE is relatively simple. It may miss several relationships worthy of analysis, and it is difficult to distinguish between similarity and causality. For example, RF6, the feed rate for FF1, FF2, and FF3, varies synchronously with the reactor pressure RP7, and there only exists similarity between RF6 and RP7 rather than causal relationship. The causality relationship between RF6 and RP7 based on MBTE method is very strong, while ADTE can correctly find there is no causality relationship between them. Table 6 is a summary of the four methods to find the relationship between variables, where Y represents the correct result, O represents the opposite result and N represents lost important relationship result.

<table>
<thead>
<tr>
<th>Method</th>
<th>Y</th>
<th>O</th>
<th>N</th>
<th>Correct rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIC</td>
<td>15</td>
<td>1</td>
<td>2</td>
<td>83.3%</td>
</tr>
<tr>
<td>BDE</td>
<td>18</td>
<td>2</td>
<td>2</td>
<td>81.8%</td>
</tr>
<tr>
<td>MBTE</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>90.9%</td>
</tr>
<tr>
<td>ADTE</td>
<td>22</td>
<td>2</td>
<td>0</td>
<td>91.7%</td>
</tr>
</tbody>
</table>

The ADTE-based Bayesian network method has the best performance and can mine deeper
relationships. Table 7 is the adjacency matrix after the integration of three sub-blocks, where the number represents propagation time in optimal structure.

<table>
<thead>
<tr>
<th></th>
<th>FF1</th>
<th>FF2</th>
<th>FF3</th>
<th>FF4</th>
<th>RF6</th>
<th>RP7</th>
<th>RL8</th>
<th>CW20</th>
<th>RT21</th>
<th>PR10</th>
<th>ST11</th>
<th>SP13</th>
<th>ST22</th>
<th>SP16</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FF2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FF3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FF4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RF6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RP7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RL8</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CW20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RT21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PR10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ST11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SP13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ST22</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SP16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The propagation time of path FF3 to FF2 is 7 sample times, which is not consistent with the common sense. The correct case is that the propagation time of path FF2 to FF3 should be zero. The reason is that, FF2 has less impact on IDV1 fault, while FF3 affects RL8 (reactor grade) and control loop causes FF2 to generate alarms. From Tab.7, we can find that propagation time of the path FF3 to RL8 is 1 sample time and path RL8 to FF2 is 6 sample times, which forms a closed loop among FF2, FF3, RL8. This phenomenon also proves that the method proposed in this paper can deeply explore the dynamic propagation relationship between variables. Based on the adjacency matrix, the alarm propagation structure of TEP is shown in Figure.13.
Blue, yellow and green colors represent the variables of the first sub-block, the second sub-block and the third sub-block respectively. Blue-yellow represents the associated variables of the first sub-block and the second sub-block. Yellow-green represents the associated variables of the second sub-block and the third sub-block. In the parameter training, this work will explore the root alarm of IDV1 alarm according to Figure.13.

4.2 Alarm Source Tracing Analysis

As the propagation time between variables in Table.7 is different, after getting the alarm propagation structure in Figure.13, it is necessary to perform virtual variable expansion according to the method proposed in section 3.5, and the virtual variable should be regarded as an independent variable for parameter training. The result of the expansion is shown in Figure.14.
The blue dotted line represents the propagation path between variables, the green dotted line represents the extended propagation path.

4.2.1 Single Alarm Source Tracing Analysis

The 28 variables (including the virtual variables) in Figure 14 are trained through the TEP data set. Assume that one alarm occurs in SP13 and substitute SP13=1 as evidence into the Bayesian network structure, the alarm probability of each variable is shown in Table 8.

<table>
<thead>
<tr>
<th>Variable</th>
<th>FF3</th>
<th>FF4</th>
<th>RF6</th>
<th>RL8</th>
<th>CW20</th>
<th>RF6-1</th>
<th>RP7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.6911</td>
<td>0.0928</td>
<td>0.9569</td>
<td>0.7495</td>
<td>0.6024</td>
<td>0.9647</td>
<td>0.6674</td>
</tr>
<tr>
<td>Alarm</td>
<td>0.3089</td>
<td>0.9072</td>
<td>0.0431</td>
<td>0.2505</td>
<td>0.3976</td>
<td>0.0353</td>
<td>0.3326</td>
</tr>
<tr>
<td>Variable</td>
<td>CW20-1</td>
<td>RT21</td>
<td>FF1</td>
<td>FF4-1</td>
<td>RP7-1</td>
<td>CW20-2</td>
<td>RT21-1</td>
</tr>
<tr>
<td>Normal</td>
<td>0.5921</td>
<td>0.7564</td>
<td>0.7153</td>
<td>0.0913</td>
<td>0.6665</td>
<td>0.5789</td>
<td>0.7558</td>
</tr>
<tr>
<td>Alarm</td>
<td>0.4079</td>
<td>0.2436</td>
<td>0.2847</td>
<td>0.9087</td>
<td>0.3335</td>
<td>0.4211</td>
<td>0.2442</td>
</tr>
<tr>
<td>Variable</td>
<td>RP7-2</td>
<td>FF2</td>
<td>FF4-2</td>
<td>RT21-2</td>
<td>CW20-3</td>
<td>RP7-3</td>
<td>CW20-4</td>
</tr>
<tr>
<td>Normal</td>
<td>0.6662</td>
<td>0.9890</td>
<td>0.0875</td>
<td>0.7610</td>
<td>0.5617</td>
<td>0.6696</td>
<td>0.5421</td>
</tr>
<tr>
<td>Alarm</td>
<td>0.3338</td>
<td>0.0110</td>
<td>0.9125</td>
<td>0.2390</td>
<td>0.4383</td>
<td>0.3304</td>
<td>0.4579</td>
</tr>
<tr>
<td>Variable</td>
<td>ST11</td>
<td>ST22</td>
<td>ST11-1</td>
<td>SP13</td>
<td>SP16</td>
<td>RT21-3</td>
<td>PR10</td>
</tr>
<tr>
<td>Normal</td>
<td>0.6619</td>
<td>0.5870</td>
<td>0.6608</td>
<td>0.0000</td>
<td>0.1919</td>
<td>0.7557</td>
<td>0.6236</td>
</tr>
<tr>
<td>Alarm</td>
<td>0.3381</td>
<td>0.4130</td>
<td>0.3392</td>
<td>1.0000</td>
<td>0.8081</td>
<td>0.2443</td>
<td>0.3764</td>
</tr>
</tbody>
</table>
According to Table 8, when SP13 alarms, the alarm probability of each variable can be obtained. The propagation path and score corresponding to each alarm roots are shown in Appendix A.

As seen from Appendix A, path 7 has the highest score and should be defined as the alarm propagation path. Through the method proposed in this paper, FF4 is detected as the root alarm variable, which is in line with the actual situation and has practical significance. The red line represents the identified propagation path in Figure 15

![Fig.15 Propagation path for the alarming of SP13 based on ADTE](image)

As can be seen from Figure 15, when SP13 alarms, FF4 is the source of the alarm flood, and the time FF4 generating an abnormality is about 14 sample times before the SP13 alarms. At the same time, there is 80% chance of alarming for SP16 at the next sample time. This information will help the operator to handle the abnormal situation.

4.2.2 Multiple Alarms Source Tracing Analysis

Like the analysis in 4.2.1, in multiple alarm source tracing, we assume that RT21-3, ST22 and SP16 simultaneously alarm. RT21-3=1, ST22=1 and SP16=1 are substituted into the Bayesian
network structure as evidence, and the alarm probability of each variable is shown in Table.9.

<table>
<thead>
<tr>
<th>Variable</th>
<th>FF3</th>
<th>FF4</th>
<th>RF6</th>
<th>RL8</th>
<th>CW20</th>
<th>RF6-1</th>
<th>RP7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.5999</td>
<td>0.0923</td>
<td>0.9551</td>
<td>0.6736</td>
<td>0.3630</td>
<td>0.9596</td>
<td>0.6302</td>
</tr>
<tr>
<td>Alarm</td>
<td>0.4001</td>
<td>0.9077</td>
<td>0.0449</td>
<td>0.3264</td>
<td>0.6370</td>
<td>0.0404</td>
<td>0.3698</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>CW20-1</th>
<th>RT21</th>
<th>FF1</th>
<th>FF4-1</th>
<th>RP7-1</th>
<th>CW20-2</th>
<th>RT21-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.3305</td>
<td>0.6539</td>
<td>0.6885</td>
<td>0.0916</td>
<td>0.6322</td>
<td>0.2894</td>
<td>0.6318</td>
</tr>
<tr>
<td>Alarm</td>
<td>0.6695</td>
<td>0.3461</td>
<td>0.3115</td>
<td>0.9084</td>
<td>0.3678</td>
<td>0.7106</td>
<td>0.3682</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>RP7-2</th>
<th>FF2</th>
<th>FF4-2</th>
<th>RT21-2</th>
<th>CW20-3</th>
<th>RP7-3</th>
<th>CW20-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.6328</td>
<td>0.9833</td>
<td>0.0894</td>
<td>0.5887</td>
<td>0.1850</td>
<td>0.6470</td>
<td>0.0778</td>
</tr>
<tr>
<td>Alarm</td>
<td>0.3672</td>
<td>0.0167</td>
<td>0.9106</td>
<td>0.4113</td>
<td>0.8150</td>
<td>0.3530</td>
<td>0.9222</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>ST11</th>
<th>ST22</th>
<th>ST11-1</th>
<th>SP13</th>
<th>SP16</th>
<th>RT21-3</th>
<th>PR10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.4324</td>
<td>0.0000</td>
<td>0.4352</td>
<td>0.0121</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.4524</td>
</tr>
<tr>
<td>Alarm</td>
<td>0.5676</td>
<td>1.0000</td>
<td>0.5648</td>
<td>0.9879</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.5476</td>
</tr>
</tbody>
</table>

According to Table.9, when RT21-3, ST22 and SP16 simultaneously alarm, the alarm probability of each variable can be obtained. The propagation path and score corresponding to each alarm root are shown in Appendix. B.

As seen from Appendix. B, path 7 is the alarm propagation path because of the highest score.

In the case where multiple alarms occur at the same time, FF4 is still detected as the root alarm variable, which is in line with the fact. The propagation path is indicated by the red dot line in Figure 16.
According to the simulation results of 4.2.1 and 4.2.2, the ADTE-based Bayesian network proposed in this paper can accurately analyze the alarm source under the alarm flood and assist the operator to handle the alarm quickly. In addition, when ST22, SP16 and RT21 alarms, according to Table.10, we can predict that PR10 has 54% chance of raising an alarm at the next sample time. It is easy to find that the ADTE-based Bayesian network can discover deeper transmission relationships, which is vital for solving the alarm flooding problems.

5. CONCLUSIONS

An ADTE-based Bayesian network method is proposed in this paper. The proposed method can accurately trace the alarm source. It is a complete alarm traceability analysis system and can help operators to respond to alarm flooding conditions quickly and accurately. Compared with the traditional alarm source tracing analysis methods, our method can distinguish the correlation and causality, and effectively solve the false causality problem caused by strong correlation between variables. Furthermore, this work establishes an alarm
active propagation structure based on the causal relationship between variables, which has
great significance to the alarm system of the actual chemical process. In order to prove the
effectiveness of the proposed method, some algorithms, such as BIC, BDE and MBTE, are
used as comparative studies in the same benchmark, TEP. The simulation results
demonstrate the superiority of the proposed method.

ACKNOWLEDGEMENT

This research is supported by the National Natural Science Foundation of China under Grant Nos.
61973022, 61573051 and 61473026.

REFERENCE

plant considering process operation and structure. *Industrial & Engineering Chemistry Research.*
*2020,* 59, 1202-1217.

[2] Boyer S A. *SCADA: supervisory control and data acquisition;* International Society of
Automation, *2009.*

integrated with singular value decomposition: Application to complex chemical processes. *Chemometrics and Intelligent Laboratory Systems,* *2020,* 200, 103981.


