

SIGNED DIRECTED GRAPH BASED MODELING AND ITS VALIDATION FROM PROCESS KNOWLEDGE AND PROCESS DATA

FAN YANG ^{*,**}, SIRISH L. SHAH ^{*}, DEYUN XIAO ^{**}

^{*} Department of Chemical and Materials Engineering
University of Alberta, 7th Floor, ECERF, 9107 116 Street, Edmonton, AB T6G 2G6, Canada
e-mail: {fyang.cme, Sirish.Shah}@ualberta.ca

^{**} Tsinghua National Laboratory for Information Science and Technology (TNList), Department of Automation
Tsinghua University, 1 Tsinghuayuan, Haidian District, Beijing 100084, China
e-mail: {yangfan, xiaody}@tsinghua.edu.cn

This paper is concerned with the fusion of information from process data and process connectivity and its subsequent use in fault diagnosis and process hazard assessment. The Signed Directed Graph (SDG), as a graphical model for capturing process topology and connectivity to show the causal relationships between process variables by material and information paths, has been widely used in root cause and hazard propagation analysis. An SDG is usually built based on process knowledge as described by piping and instrumentation diagrams. This is a complex and experience-dependent task, and therefore the resulting SDG should be validated by process data before being used for analysis. This paper introduces two validation methods. One is based on cross-correlation analysis of process data with assumed time delays, while the other is based on transfer entropy, where the correlation coefficient between two variables or the information transfer from one variable to another can be computed to validate the corresponding paths in SDGs. In addition to this, the relationship captured by data-based methods should also be validated by process knowledge to confirm its causality. This knowledge can be realized by checking the reachability or the influence of one variable on another based on the corresponding SDG which is the basis of causality. A case study of an industrial process is presented to illustrate the application of the proposed methods.

Keywords: signed directed graph, transfer entropy, process topology, fault diagnosis, process hazard assessment.

1. Introduction

The scale and complexity of process operations have been increasing to meet the business and regulatory compliance demands of industry. These industrial processes, however, are often subject to low productivity due to system faults and/or even hazardous operating conditions because of abnormal conditions such as faulty operations, equipment degradation, external disturbances, and control system failures. Compared with the traditional fault detection in local systems, fault detection in large-scale complex systems is particularly difficult because of the high degree of interconnections in such systems. A simple fault in such systems may easily propagate throughout the plant. Such a type of analysis is similar to hazard operability analysis, which is initially carried out in a qualitative way followed by quantitative analysis. It is important to remember that simple local faults can propagate and spread to a wider do-

main because of the connectivity between units and the topology of the system. Therefore, in root cause analysis of process faults, it is important to capture connectivity and topology to use this information to relate to the dynamics of the systems as captured in process data. This fusion of information can be further utilized to find the probable root causes and resulting consequences according to the current symptom(s). Therefore, the problems of fault detection and propagation require analysis on two fronts: (i) capturing process connectivity and causality, and (ii) validation of such models using both process knowledge and process data.

There are several models to describe process topology and cause/effect relations. Causal graphs describe the causal relations between variables, and they can be extended to AND/OR causal graphs to describe logical causal relations (Ligeza and Kościelny, 2008). A potential conflict structure, as a subgraph of a causal graph,

describes specifically the computational causal relations used for consistency-based diagnosis (Ligeza, 1996). There are also more complex models that describe essentially causal relations. Bond graphs and their extension, temporal causal graphs (Mosterman and Biswas, 1999), use different symbols to further describe dynamic characteristics. More precisely, qualitative transfer functions (Leyval *et al.*, 1994) and differential equations (Montmain and Gentil, 2000) have been integrated into causal graphs, and complex algorithms are introduced to improve their correctness (Cheng *et al.*, 2011). Similar or improved approaches were investigated by many researchers, such as Pastor *et al.* (2000), Alonso *et al.* (2003), Fagarasan *et al.* (2004), and Jan *et al.* (2007). All of these improved methods depend on mathematical models. If Matlab/Simulink models of dynamic systems are available, then the diagnostic models can be generated directly (Górny, 2001). The dependency on models, however, limits the utility in large-scale complex systems.

As a qualitative graphical model, the Signed Directed Graph (SDG) technique has been used to describe the topology and causality of industrial processes, including both material flow and information flow paths (Iri *et al.*, 1979). Compared to classic causal graphs, SDGs add signs to nodes and arcs to include more information and yet remain simple and qualitative. SDGs can help us understand the internal relationships between variables and thus be implemented in hazard assessment (to find the possible consequence of a fault) and also in root cause diagnosis (Yang *et al.*, 2010a).

Although the modeling of SDGs can be based on mathematical equations (Maurya *et al.*, 2003a; 2003b; 2004), it can also be undertaken practically from process knowledge that is readily available from Piping and Instrumentation Diagrams (P&IDs). Although simple, SDGs are able to capture fairly precise process connectivity information from the direct relationship between variables, as based on causality. Some techniques to capture connectivity from P&IDs have been developed by Fedai and Drath (2005), Yim *et al.* (2006), and Thambirajah *et al.* (2009) based on the standard data format of Computer Aided Engineering Exchange (CAEX), though there may also be a lot of irrelevant information in these files. The SDGs constructed in this way should be validated by the process data to confirm their correctness.

When dealing with process data, various pairwise causality capture methods have been proposed (Bauer *et al.*, 2007; Bauer and Thornhill, 2008), whereas constructing a causal network or determining the topology purely from data is difficult, if not impossible, unless one also resorts to process knowledge. In multivariate cases, knowledge discovery in databases has been studied and implemented in the DiaSter system (Korbicz and Kościelny, 2010), including the discovery of qualitative and quantitative dependencies. However, when the scale of a pro-

cess is large, most of the data-based methods are essentially pattern recognition, which is difficult to be combined with knowledge-based or model-based methods. Moreover, by these data-based methods, one is likely to run into a lot of redundancy and errors.

In general, the combination of knowledge and data reinforces the trust in SDG models and leads to their use with a high degree of confidence. In this paper, SDG models are used because they can be built by knowledge as well as data and therefore provide a common description that facilitates their validation.

It should be noted that verification and validation are two important steps to be followed before the model can be used. Palmer and Chung (1999) discussed verification to ensure that the internal structure of the model is complete, correct, and consistent; this is based on the model itself. However, the model should also be tested in the environment in which it is going to be used; in practice, historical data are employed for this test, which is the purpose of this paper.

This paper is organized as follows. Section 2 provides an overview of the SDG modeling and inference method based on process knowledge. Section 3 introduces two data-based methods, time delayed cross-correlation and transfer entropy, to capture causality between two variables. Section 4 describes mutual validation of knowledge description and a data-based causal network. In Section 5 the above methods are evaluated by application to a real industrial process where we build an SDG and validate it by process data, followed by concluding remarks in Section 6.

2. SDG modeling by process knowledge

A process variable (or another continuous variable such as a manipulated one) is represented by a node in a SDG, and the cause/effect relationship between two variables is represented by a directed arc between the corresponding nodes (Iri *et al.*, 1979). The source node means the cause or parent, and the destination node means the effect or child. For example, an arc from node *A* to node *B* implies that deviation or perturbation in *A* may cause deviation in *B*. A sign '+', '-', or '0' is assigned to a node in comparison with thresholds to denote higher than, lower than, or within the normal operating region, respectively. Positive or negative influence between nodes is distinguished by the sign '+' (promotion) or '-' (suppression) of the arc, shown as a solid or dotted arc respectively.

The construction procedure of SDGs can be accomplished based on mathematical models, i.e., Differential and Algebraic Equations (DAEs). For example, the following Differential Equation (DE):

$$\frac{dx_i}{dt} = f_i(x_1, \dots, x_n) \quad (1)$$

can be transformed into a set of arcs with the sign

$$\text{sgn}(x_j \rightarrow x_i) = \text{sgn} \left. \frac{\partial f_i}{\partial x_j} \right|_{x_1^0, \dots, x_n^0}, \quad (2)$$

where (x_1^0, \dots, x_n^0) denotes the steady state. Algebraic Equations (AEs) can be transformed into arcs in much the same way.

In most cases, the SDG is established by qualitative process knowledge and experience because precise equations are usually unavailable and unnecessary. Nevertheless, process knowledge is available in P&IDs, and it is suggested that one should capture process topology from this information, where upstream variables influence downstream variables, external environmental variables influence internal state variables, and manipulated variables influence controlled variables in a control system.

2.1. SDG modeling within a process unit. For a single unit, several physical quantities reflect the characteristics of the process. The following three types of relationships between these variables can be described as DAEs: DEs, which reflect dynamic causal relationships (Type I), AEs with causal meaning, which include driving force equations, functional relationships, and other algebraic equalities (Type II), and AEs with no causal relationships (Type III) (Yang *et al.*, 2010b). For the first two types, Maurya *et al.* (2003a) summarized the modeling methods. The third type can be used as redundant constraints to remove irrelevant solutions (Oyeleye and Kramer, 1988).

Based on these methods, an SDG of a unit can be obtained from qualitative knowledge instead of DAEs. For example, consider a tank process as shown in Fig. 1, where L is the level in the tank, K is the valve position in the outlet pipe, F_1 and F_2 are inlet and outlet flowrates, respectively. The DAEs of this process are given below:

$$A \frac{dL}{dt} = F_1 - F_2, \quad (3)$$

$$F_2 = K\sqrt{L}, \quad (4)$$

$$K = \alpha L, \quad (5)$$

where A is the cross sectional area of the tank and α is the proportional coefficient of the control law. By these DAEs, an SDG can be built as shown in Fig. 2, where 2(a)

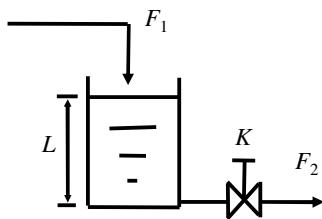


Fig. 1. Scheme of a tank process.

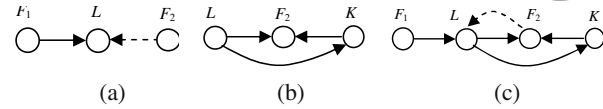


Fig. 2. SDGs of the tank process: SDG obtained from Eqn. (3) (a), SDG obtained from Eqns. (4) and (5) (b), combination of (a) and (b) (c).

is obtained from Eqn. (3) (Type I), 2(b) from Eqns. (4) and (5) (Type II), and 2(c) is a combination of them. This SDG can also be obtained from process knowledge that the level is determined by the inlet and outlet and the outlet flow rate is affected by the level and the valve position. Irrespective of the type of control (P, PI, or PID), it is shown as an arc from L to K , and it is unnecessary to write the corresponding DAEs for this case. Accurate parameters and functions are not important for this usage.

In the process industry, standard units, such as tanks and exchangers, can be modeled and saved as templates or modules for reuse. One can build SDGs for special units by using the qualitative information obtained from process knowledge as well as first-principles or mathematical models. When all the units are modeled, they should be connected according to the linkage information.

2.2. SDG modeling between process units. Units are connected because of material and signal flow paths. The directions of arcs are consistent with the transportation paths. The arcs should be connected to corresponding variables in different units.

An efficient way of showing these transportations is via P&IDs in which both material flow and information flow (control signal) are shown. SDGs can be built by unfolding the units in P&IDs as unit SDGs and connecting them according to flow directions. Thus SDGs bring out not only the connectivity and topology between the units but also the causality between variables.

This plant topology information can be expressed by a text file that can be automatically generated by CAD tools (such as the DXF format developed by AutoCAD for enabling data interoperability). The program should extract related information from it to identify process units and flowsheets (Palmer and Chung, 2000). For a standard and uniform extraction of connectivity information, XML (eXtensible Markup Language) is one possible source of information. Some CAD tools can export XML texts, and some techniques of computer science can be employed to capture the connectivity information from them (Fedai and Drath, 2005; Yim *et al.*, 2006; Thambirajah *et al.*, 2009). This format of connectivity information forms the basis for causality and can be converted into matrices or SDGs essentially.

2.3. SDG-based inference and model improvement.

In the safety area, fault diagnosis, especially root cause diagnosis, and process hazard assessment, especially HAZard and OPerability analysis (HAZOP), are two different tasks. The former is to correctly identify the root cause of a fault. It is based on measurements and used in real-time. The latter is used off-line to find the possible consequences due to every cause. By assuming various faults, one can analyze possible consequences that are likely to be triggered. Both tasks need a model and mechanism to describe how a fault propagates, and thus the SDG model can be employed.

Based on an SDG, fault propagation can be explained by searching for consistent paths. An arc is called consistent if the product of the sign of its source, the sign of its destination, and the sign of the arc is equal to '+'. Take arc $A \rightarrow B$ with sign '-' (suppression), for example. If A is '+', and B is '-', then this arc is consistent, which means the fault on A can be propagated to B along this arc because the signs show the suppression effect. A fault cannot be propagated along an inconsistent arc even though there may be a connection. Consistent arcs are connected successively to compose a consistent path, i.e., a reasonable fault propagation path. In Fig. 3, a consistent path is highlighted, while there is another node with a sign '-' enclosed in a circle that does not belong to any consistent path, because the arc adjacent to it is inconsistent in spite of the existing connection.

According to this rule of consistency, one can search for root causes and consequences by backtracking and forward tracking, respectively. A common search algorithm is the depth-first or width-first traversal on the graph (Iri et al., 1979), by which the root causes or affected variables as well as the propagation paths can be obtained.

One disadvantage of the aforementioned unit-based method is incorrect or ambiguous inference due to the existence of divider/header combinations or recycle loops, because there may exist multiple paths with the same origin and terminal. This is unavoidable because of the incomplete information of qualitative SDGs. This disadvantage, however, can be partially overcome through model improvement. Maurya et al. (2003b) analyzed model description in detail by distinguishing an initial response and a steady-state response and modified the SDG by adding redundant equations. Palmer and Chung (2000) developed

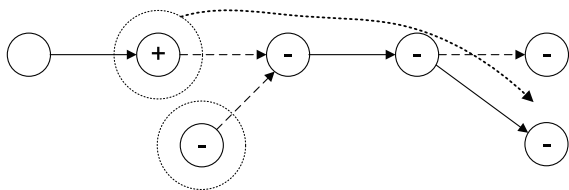


Fig. 3. Example of a consistent path.

the modular approach to remove redundant paths. Neither of these efforts can solve the problem perfectly or efficiently without investigating real data, especially for large processes. This issue will be discussed in Section 4.1.

3. Causality capture based on process data

There are many methods to capture causality information from process data. Two of them are given below. One is the simplest and yet a very practical one, the other is a general but computationally complex one.

3.1. Causality capture based on cross-correlation.

In a process, variables are interacting. By analyzing the correlation between them, we can capture the causality. Although correlation does not imply causality, it can be used as a validation. Moreover, when computing Pearson's correlation, causes and effects can be recognized by introducing lags in a time series to find the maximum correlation.

Assume that x and y are time series of n observations with means μ_x and μ_y and standard deviations σ_x and σ_y , respectively. Then the Cross-Correlation Function (CCF) with an assumed lag k is

$$\phi_{xy} = \frac{E[(x_i - \mu_x)(y_{i+k} - \mu_y)]}{\sigma_x \sigma_y} \quad (6)$$

The expectation can be estimated by the sample CCF as

$$\hat{\phi}_{xy}(k) = \begin{cases} \frac{\sum_{i=1}^{n-k} (x_i - \mu_x)(y_{i+k} - \mu_y)}{(n-k)\sigma_x \sigma_y} & \text{if } k \geq 0, \\ \frac{\sum_{i=1-k}^n (x_i - \mu_x)(y_{i+k} - \mu_y)}{(n+k)\sigma_x \sigma_y} & \text{if } k < 0. \end{cases} \quad (7)$$

A value of the CCF is obtained by assuming a certain time delay for one of the time series. Thus the absolute maximum value can be regarded as the real cross-correlation and the corresponding lag as the estimated time delay between these two variables. For mathematical description, one can compute the maximum and minimum values $\phi^{\max} = \max_k \{\phi_{xy}(k), 0\} \geq 0$ and $\phi^{\min} = \min_k \{\phi_{xy}(k), 0\} \leq 0$, and the corresponding arguments k^{\max} and k^{\min} . Then the time delay from x to y is

$$\lambda = \begin{cases} k^{\max} & \text{if } \phi^{\max} \geq -\phi^{\min}, \\ k^{\min} & \text{if } \phi^{\max} < -\phi^{\min} \end{cases} \quad (8)$$

(corresponding to the maximum absolute value) and the actual time delayed cross-correlation is $\rho = \phi_{xy}(\lambda)$ (between -1 and 1). If λ is less than zero, then it means that the actual delay is from y to x . Thus the sign of λ provides the directionality information between x and y . The

sign of ρ corresponds to the sign of the arc in the SDG indicating that the correlation is positive or negative.

By this definition, ρ is a statistical estimate inevitably prone to some uncertainty due to disturbances, noise and the size of data windows. Therefore its value should be judged with care. Even if the two time series are uncorrelated random noises, ρ may still likely be different from zero. Therefore the value of the CCF between two variables should be checked against a threshold. If the correlation between the two series is very weak, then the effect of the noise will dominate the results. Therefore, in correlation analysis, only those values which are significantly larger than a user-defined threshold (e.g., ± 0.2) are considered to be evidence of correlations.

To sum up, based on estimation, the maximum CCF is defined as the time delayed correlation coefficient (or correlation in brief), and the corresponding argument k is defined as an estimate of the time delay from x to y , from which we can find the cause and the effect (Bauer and Thornhill, 2008). Of course, correlations are based on statistics under the assumption of linearity, and thus they need hypothesis tests to obtain the level of significance. Although the estimates of correlation coefficients are not accurate, the directions are believable for most cases. Although this method is practical and easy for computation, it has many shortcomings, some of which are explained below.

- The nonlinear causal relationship does not necessarily show up in correlation analysis. For example, if y equals the square of x with the time delay of one sampling time, then, based on the time-delayed cross-correlation, this obvious causality cannot be found because all the values are small relative to a threshold. This can be explained because the true correlation should be zero.
- Correlation simply gives us an estimate of the time delay. The sign of the delay is an estimate of the directionality of the signal flow path. The time delay obtained, however, is only an estimate. In addition, the trend in a time series is ignored, and values at different time instances are regarded as samples of the same random event. Thus the causality obtained by this measure is purely the time delay based on the estimate of the covariance.

Given all the correlations between any two variables, a correlation color map can be constructed (Tangirala *et al.*, 2005), whose horizontal and vertical coordinates are both variables in the same order and the color of each pixel shows the correlation between the two corresponding variables according to a scaled color bar. This provides an intuitive way for human beings to observe correlation, especially for identification of similar groups of variables.

Based on these correlations with directions, one can construct a causal network. Because correlation has the

property of a transitive relation, some correlations can be explained by sequential direct causal relations. For example, the causality from A to C can be a combined result of causal relations from A to B and from B to C . Pairwise data analysis cannot recognize this difference without process knowledge. In addition, other related variables can affect the correlation between two variables to lessen the intensity, which may interfere the analysis results. In any case, a correlation high enough, whether direct or indirect, should be explained by an arc or a path in the corresponding SDG. This is a means of validating SDGs.

3.2. Causality capture based on transfer entropy.

From another point of view, causal relation means information transfer. Thus the measure of transfer entropy (Schreiber, 2000) in information theory can also be employed. The transfer entropy from y to x is defined as

$$t(x|y) = \sum p(x_{i+h}, \mathbf{x}_i, \mathbf{y}_i) \cdot \log \frac{p(x_{i+h} | \mathbf{x}_i, \mathbf{y}_i)}{p(x_{i+h} | \mathbf{x}_i)}, \quad (9)$$

where p means the complete or conditional Probability Density Function (PDF), $\mathbf{x}_i = [x_i, x_{i-\tau}, \dots, x_{i-(k-1)\tau}]$, $\mathbf{y}_i = [y_i, y_{i-\tau}, \dots, y_{i-(l-1)\tau}]$, τ is the sampling period, and h is the prediction horizon. The transfer entropy is a measure of information transfer from y to x done by measuring the reduction of uncertainty while assuming predictability. It is defined as the difference between the information about a future observation of x obtained from the simultaneous observation of past values of both x and y , and the information about the future of x obtained from the past values of x alone. It gives us a good sense of the causality information without having to require the delay information. Experience suggests that one takes $\tau = h \leq 4$, $k = 0$, and $l = 1$ for the initial trial, while the usual way is to test and compare various results from different values of τ and h . If the transfer entropies in two directions are considered, then $t(x \rightarrow y) = t(y|x) - t(x|y)$ is used as a causality measure to decide the quantity and direction of information transfer (Bauer *et al.*, 2007).

In Eqn. (9), the PDF can be estimated from a histogram or kernel methods (Silverman, 1986), which are nonparametric methods, to fit any shape of the distributions. Here the Gaussian kernel method is used because it is more robust than the naive histogram-based method. The Gaussian kernel function is defined as

$$K(v) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}v^2}. \quad (10)$$

Thus a univariate PDF can be estimated by

$$\hat{p}(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - x_i}{h}\right), \quad (11)$$

where N is the number of samples, and h is the bandwidth chosen to minimize the mean square error of the

PDF estimation calculated by $h = c \cdot \sigma \cdot N^{-1/5}$, where $c = (4/3)^{1/5} \approx 1.06$ according to a ‘normal reference rule-of-thumb’ (Li and Racine, 2007).

For the multivariate case (q dimensional), the estimation of PDF is

$$\hat{p}(x_1, x_2, \dots, x_q) = \frac{1}{Nh_1 \cdots h_q} \cdot \sum_{i=1}^N K\left(\frac{x_1 - x_{i1}}{h_1}\right) \cdots K\left(\frac{x_q - x_{iq}}{h_q}\right), \quad (12)$$

where

$$h_s = c\hat{\sigma}(x_{is})_{i=1}^N \cdot N^{-1/(4+q)}$$

for $s = 1, \dots, q$, and the notation is the same as the univariate case.

Compared to the approach based on cross-correlation, this one can be applied to more general conditions such as nonlinear relations. In the above (nonlinear) example, the causality cannot be validated based on the cross-correlation. Given a lag of 1, however, the transfer entropies from x to y and vice versa are 0.27 and 0.01, respectively, and thus the causality from x to y can be confirmed, which is consistent with the actual setting.

Transfer entropy shows the information transfer in each direction. Thus this method provides more insight into causality for complex systems especially for the case with recycles. In Fig. 4(a), x and y are connected directly via a forward path and a recycle path. In the forward channel from x to y , the relation is via an AR model, i.e., $y(i) + 0.5y(i-1) = x(i-5)$, whereas there also exists a feedback channel from y to x , i.e., $x(i) = 1 - 0.5y(i-1)$. Thus the information transfer lies in the two channels. If one is only concerned with the causality measure, then $t(x \rightarrow y)$ (in Fig. 4(b)) indicates that the causality is from x to y . However, if transfer entropies $t(y|x)$ (in Fig. 4(c)) and $t(x|y)$ (in Fig. 4(d)) are studied, then both arcs can be validated.

To sum up, this method is more general than the previous one and is able to reveal the essential relationship even when the lag is inexplicit. There are several parameters to be set by users, which provide some freedom. It is, however, highly dependent on the estimation of PDFs (although it can have any non-Gaussian forms). Thus the computational burden is higher. However, when using this method, the time delay cannot be estimated, and the arc signs in SDGs cannot be obtained. In real applications, one may mainly choose correlation analysis for validation but sometimes the transfer entropy can bring in additional insights.

There are other alternative methods to capture the causality between time series such as Granger causality and predictability improvement (Lungarella et al., 2007).

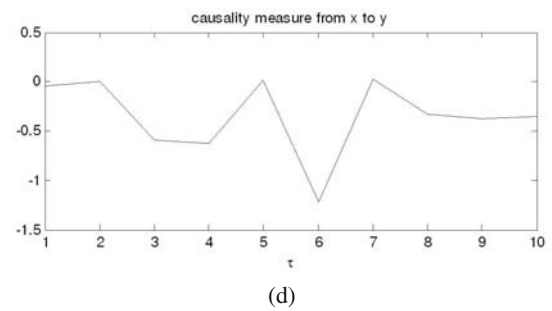
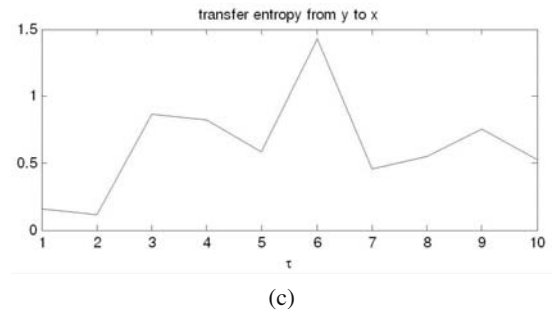
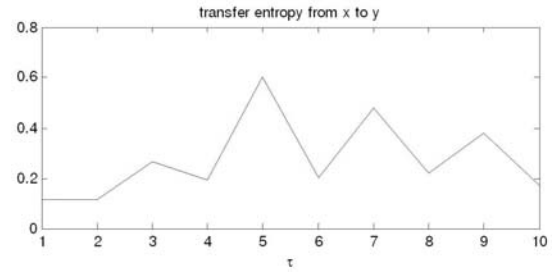
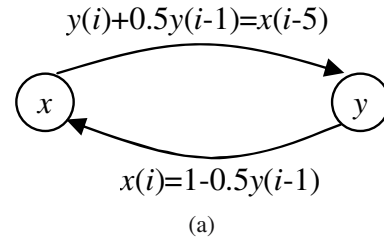


Fig. 4. Transfer entropy measures: two variables and their relations (a), transfer entropy from x to y (b), transfer entropy from y to x (c), causality measure which is the difference (d).

Each has its own advantages and limitations. They complement each other and no one method is powerful enough to replace the others. Hence data analysis is a combination of various methods.

4. Mutual validation of process description

Process knowledge and process data are two means to capture causality information in a process. Neither, however, is sufficient for practical use because of redundancy and errors in the resulting models. We should combine them by mutual validation.

4.1. Using process data to validate knowledge description. A causal network constructed based on process knowledge is essentially closer to the real causality. Ideally, it covers all the possible paths. It also has the ability to avoid indirect relations by avoiding parallel paths. For example, if A can reach B and B can reach C , then A can reach C . However, if the reachable path from A to C is just realized by the serial connection of the two paths from A to B and from B to C , then the direct arc from A to C can be deleted, otherwise there should exist sound reasons to include different parallel paths to achieve the reachability from A to C , resulting in a necessary arc. The major problem of this causal network constructed by process knowledge is the lack of quantitative information to confirm its reliability. In fact, there are usually redundant and irrelevant arcs that should be deleted.

The reasons for this redundancy include many cases whereas the essential reason is that connectivity is merely a necessary but insufficient condition for causality. While detailed information can also be obtained to exclude the inexistent arcs or choose the dominant path, the effort would be multiplied because quantitative process knowledge is needed in addition to qualitative information. In the modular approach (Palmer and Chung, 2000), this task is left to the user, although semi-automated; it still highly relies on experience. Thus we limit the process knowledge of concern to P&IDs, that is to say, qualitative connectivity information. Under this circumstance, complex structures, such as dividers, headers, or recycle loops, often make the qualitative algebra provide ambiguous results that are difficult to improve according to qualitative information because the intensity of each arc is unknown. This is an intrinsic problem of the SDG model, although a few researchers have made some improvements (Palmer and Chung, 2000; Maurya *et al.*, 2003b).

In general, it is difficult to exclude physically broken (e.g., valve block), behaviourally nondescript loop (e.g., control loop), or extremely weak (due to attenuation of signals) paths. To validate the causal network, one should resort to process data for quantitative evidence. If there is no data support for reachability, then the causality should be excluded. For the tank system shown in Fig. 1, if the data of F_1 , F_2 , L , and K (take controller output as an alternative) are available and sufficiently excited, then the arcs in Fig. 2(c) can be validated except $F_2 \rightarrow L$ because the control determines the value of L .

4.2. Using process knowledge to validate data-based relations. Although there are several data-based methods to capture causality, they are developed to find the causality between two variables. Real systems, however, are multivariate, the causality within which is shown as a network with weights based on the causality measures between every two variables. Thus a causality matrix is obtained to reflect the magnitude of the causality of each pair of variables, and the direction is determined by the time delay in correlation analysis or the sign of measure in information transfer computation. The topology of the causal network, however, also relies on the propagation relations by screening the indirect relations. According to Bauer and Thornhill (2008), one of the two typical topologies (extreme cases) is generated according to the number of nonzero entries in the first row and above the main diagonal of the causality matrix, while the real topology is the combination of these two topology forms. With the aid of a correlation test and a directionality test, one can select the evident relations first and construct the network (Yang *et al.*, 2010b). A method based on correlation is proposed, which introduces the resulting quantitative information obtained by CCF computations.

- *Step 1:* In matrix \mathbf{P} , select the maximum value in the elements that has not been used and tested.
- *Step 2:* Check the results of the correlation test and directionality test. If the correlation value fails to pass the tests, then stop.
- *Step 3:* Check the result of the consistency test for all the variables in the existing arcs. If it fails, then go to Step 5.
- *Step 4:* Add an arc corresponding to this element with an estimated time delay. The sign of the arc is determined by the sign of the element.
- *Step 5:* Go to Step 1.

Nevertheless, the resulting network can be incorrect and may be inconcise without validation by process knowledge.

One possible way to validate an acceptable causal network is to check the reachability between the two variables for evident causality; this can be realized by searching consistent paths. This treatment, however, cannot exclude indirect relations; the arcs are selected according to their magnitudes of causality. For example, if causal relations exist from A to B and from B to C , there may be strong causality between A and C that can be confirmed by reachability, yet it is actually redundant. The study to identify whether the path is direct or indirect is underway in the temporal as well as the spectral domain (Gigi and Tangirala, 2010). Although it is possible, it is not efficient without looking at process knowledge. Thus, in most cases,

one should make good use of process knowledge in the modeling procedure.

For the tank system as shown in Fig. 1, based on the data of F_1 , F_2 , L , and K , one can easily obtain the causal relations from F_1 to F_2 , from K to F_2 , and from F_1 to K when the level control is in effect. These arcs can be validated by reachability check. If the data set also includes a transient response, then other causal relations, such as from F_1 to L , can be detected. The SDG can be constructed accordingly and validated. The structure related to L is slightly different from Fig. 2(c) because of the level control.

5. Industrial case study

Consider a real industrial system, the Final Tailings Pump House (FTPH) process at Suncor Energy Inc. in Fort McMurray, Alberta, Canada.

5.1. Process description. The flow sheet for this process is shown in Fig. 5, where some texts have been made illegible and the control strategy is omitted for confidentiality reasons. The tailings from upstream are pumped into a distributor and then processed in parallel cyclo-packs and pump boxes, and finally discharged into the ponds. There are five parallel lines from the cyclo-pack downwards, where Lines A/B/C/E are structurally identical while Line D is distinct. Based on the pressure of the distributor, a prioritization program is implemented on the parallel lines and Line A is therefore the most important.

The single-loop, cascade, and selective control strategy are applied, including distributor pressure control, cyclo-pack pressure and underflow control by adjusting the number of cyclones opened, gypsum addition flow rate control, pump box level control and discharge density control by adjusting Cold Process Water (CPW), pump box level control by adjusting pump speed, pump box discharge flow rate control by adjusting pump speed, and Mature Final Tailing (MFT) addition flow rate control. For this process an SDG constructed from process knowledge is shown in Fig. 6.

For simplicity, six key variables $x_i (i = 1, \dots, 6)$ in Line A are chosen, which are $y10$, $y16$, $y18$, $y21$, $y28$, and $y30$, respectively, as indicated in Table 1. The corresponding subsection of the SDG (Fig. 7(a)) is extracted from the SDG in Fig. 6. This can be easily done by checking reachability in the SDG model.

5.2. Using process data to validate knowledge description. Based on the process data of the variables in Line A over a one-week long period (with one minute sampling frequency), the correlation color map obtained is displayed in Fig. 7(c). We use these data and corresponding time delays for SDG validation. For example, the bottom left corner is a cluster of correlated variables. By checking

the P&ID, they are found to be associated with the level and density controls in the same pump box.

One can focus on the cross-correlations and time delays between the six variables for further analysis. They are shown by the following correlation matrix \mathbf{P} (comprising all the correlation coefficients between two variables) and the causality matrix $\mathbf{\Lambda}$ (comprising all the estimated time delay λ s from one variable to another).

Table 1. Key process variables in the FTPH process: Line A for causality analysis.

Notation	Tag name	Description
x_1	y10	distributor pressure
x_2	y16	gypsum addition flow rate
x_3	y18	gypsum density
x_4	y21	sludge header pressure
x_5	y28	sump density
x_6	y30	sump level

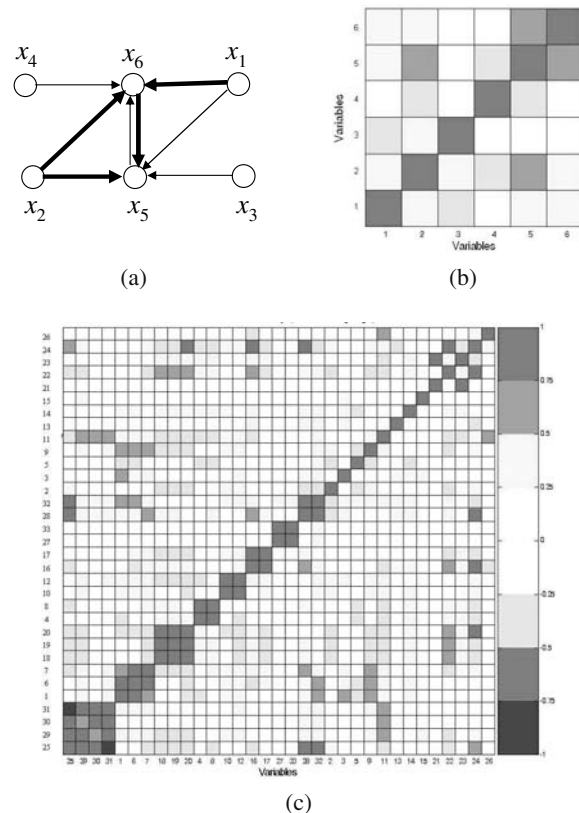


Fig. 7. SDG and its validation by cross-correlation: subsection of the SDG concerning six important variables in which the thick lines are validated (a), correlation color map of the six important variables (b), correlation color map of all the variables in Line A (c). Readers can contact the authors (yangfan@tsinghua.edu.cn) for a colored version of this figure for better visualization.

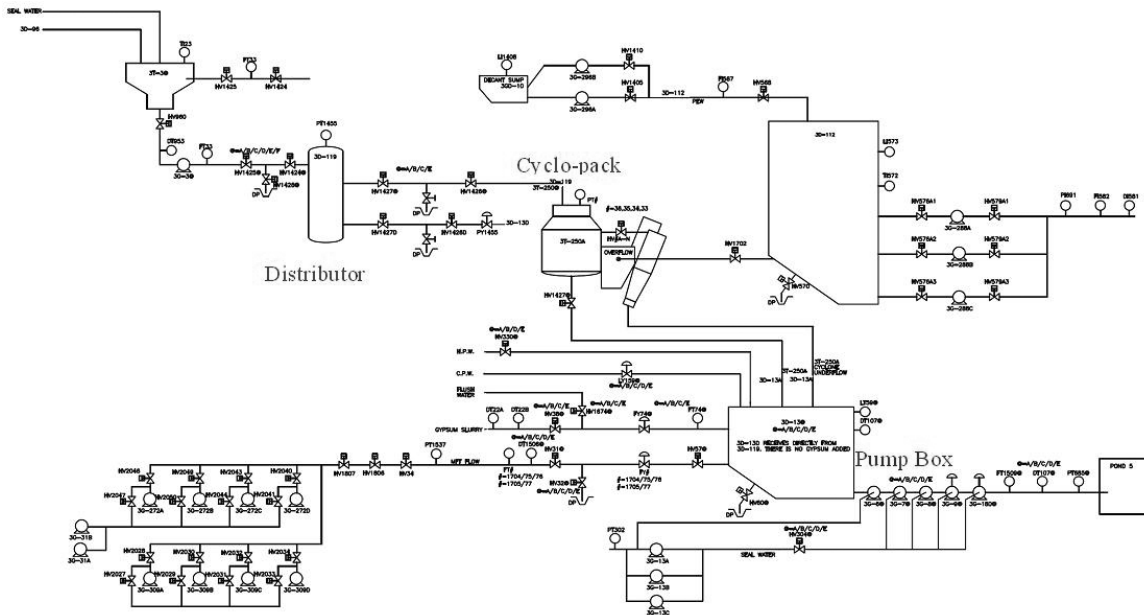


Fig. 5. Partial flowsheet of the FTPH process (texts made illegible for confidentiality reasons).

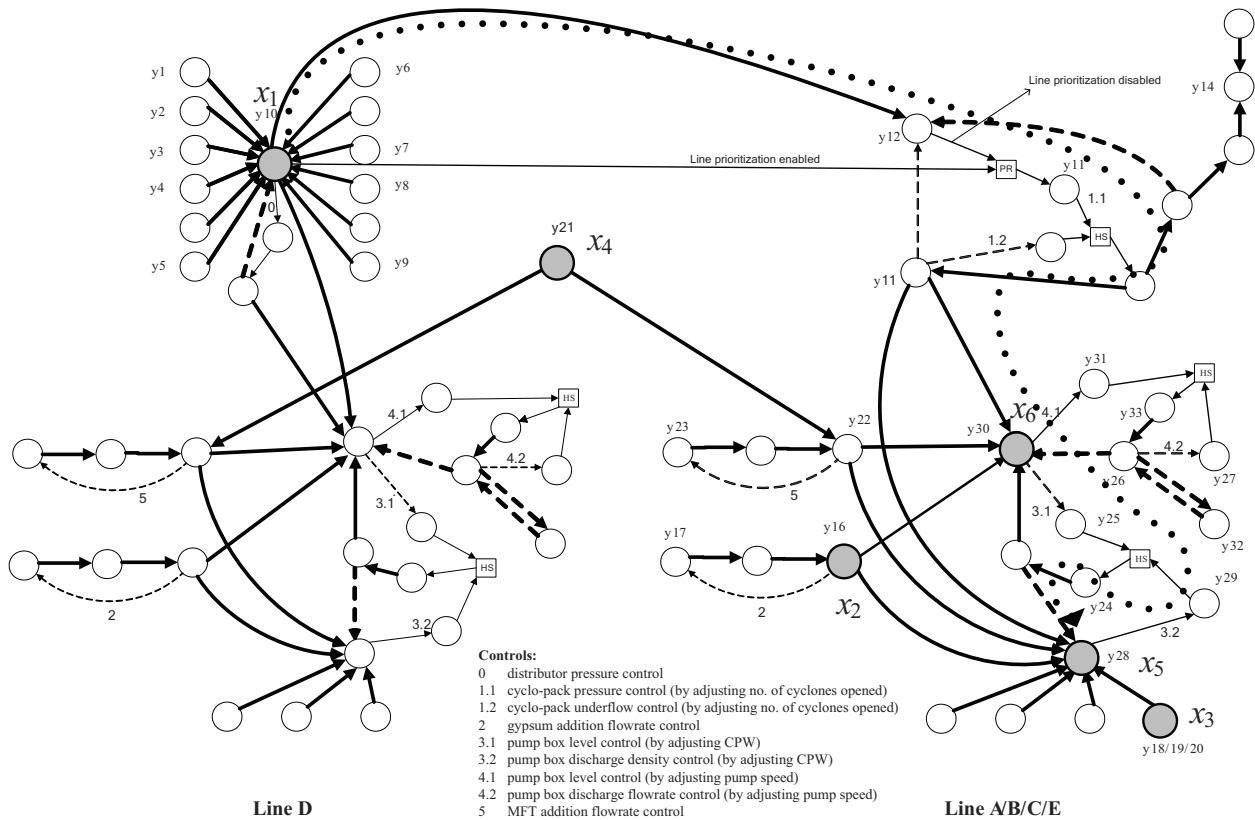


Fig. 6. SDG of the FTPH process. Relationships between variables are shown as thick lines while control signal paths are shown as thin lines. Solid and dashed lines show positive (reinforcement) and negative (reduction) causal relations, respectively. A dotted line is a fault propagation path.

$$\mathbf{P} = \begin{bmatrix} 1 & .28 & -.28 & -.18 & .39 & .41 \\ & 1 & .36 & -.31 & .74 & .50 \\ & & 1 & -.14 & .10 & -.11 \\ & & & 1 & -.25 & -.24 \\ & & & & 1 & .75 \\ & & & & & 1 \end{bmatrix}, \quad (13)$$

$$\mathbf{\Lambda} = \begin{bmatrix} 0 & -271 & 220 & 32 & 49 & 5 \\ & 0 & -360 & 2 & 1 & 1 \\ & & 0 & -357 & 359 & -360 \\ & & & 0 & 20 & 60 \\ & & & & 0 & -1 \\ & & & & & 0 \end{bmatrix}. \quad (14)$$

For example, the (i, j) -th elements of \mathbf{P} and $\mathbf{\Lambda}$ are the correlation and the estimated time delay between variables i and j , respectively. Due to the symmetric and anti-symmetric properties of these two matrices, only the elements above the diagonal are computed and shown here. Rather than looking at the correlation matrix in the numerical form, it is better to look at the color coded correlation matrix as shown in Fig. 7(b) which is a portion of Fig. 7(c).

In Eqn. (13), those values exceeding a pre-set threshold (such as 0.4) can be used to validate some of the arcs in the subsection of the SDG shown in Fig. 7(a) as thick lines. For example, correlation from x_5 to x_6 is 0.75 and the time delay is -1 , so the arc from x_6 to x_5 is validated. Similarly, the arcs from x_2 to x_5 , x_2 to x_6 , x_1 to x_6 , and x_1 to x_5 are also validated. Note that very large time delays do not make sense because they show the ineffectiveness of this measure and the computed correlation is considered invalid.

There are still some arcs that have not been validated. Thus we resort to a more general but more complex measure—transfer entropy. To obtain a rough insight, if we look at the time trends, then we observe that the trends of x_2 and x_4 do not look like the other four variables. This shows that, during this period, the relations associated with x_2 and x_4 are not strong enough to be validated. We have to examine more data with sufficient excitement to check if there is a discernible relation between these variables. However, the trends of the other four variables have apparent similarities that should be definitely due to causality. In order to reduce the computational load, we extract from the previous data set only 200 minute records of continuous data shown in Fig. 8(a).

The transfer entropy measure is used to compute the information transfer between these four variables where τ is assumed to lie between 1 and 10. When τ is 9, the transfer entropies from x_1 to x_6 , from x_3 to x_5 , and from x_5 to x_6 all reach their individual maximum values: 2.01, 1.48, and 1.41, respectively. Thus the bidirectional relationship

between x_5 and x_6 is validated, and the reachability from x_3 to x_5 is also validated. They are marked in Fig. 8(b) as thick lines.

By combing the above two methods, we found that most of the arcs have been validated except from x_4 to x_6 and from x_1 to x_5 . The former can be explained by process knowledge because x_6 is the sump level, affected by quite a few variables due to various feeds. To validate this, we have to extract the influence of each variable by signal conditioning, which is an ongoing work in the framework of multivariate systems. The latter is an indirect relation, i.e., x_1 and x_5 are related with the relay of x_6 . Since quantitative information is missing in the SDG model, the transitive property may be weakened due to the attenuation during the propagation. The time-delayed cross-correlations from x_1 to x_6 and from x_6 to x_5 are 0.41 and 0.75, respectively, but from x_1 to x_5 it is 0.39, less than the above two. Moreover, the time delays of the above two are 5 and 1, respectively, while the indirect one is 49, making the validation result of this arc unacceptable. Therefore, data analysis is usually used to only validate direct arcs.

5.3. Using process knowledge to validate data-based relations. Starting from data-based methods, for exam-

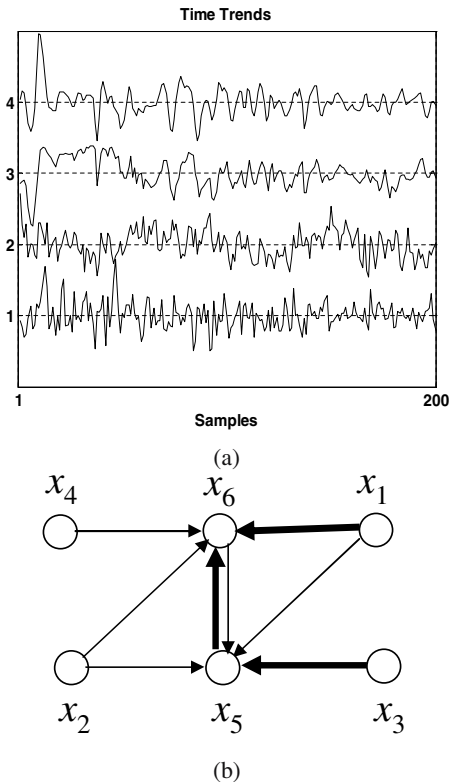


Fig. 8. Time trends and validation results by transfer entropy: time trends of four important variables (1: x_1 ; 2: x_3 ; 3: x_5 ; and 4: x_6) in the process (a), SDG and validated arcs (thick lines) (b).

ple, correlation analysis, one can obtain the causal network. According to the procedure in Section 4.2 and Eqn. (13), 0.74 is selected, so there is an arc between x_5 and x_6 , and the sign is '+'. The corresponding element in Λ is -1 as shown in Eqn. (14). Thus the direction is from x_6 to x_5 . Similarly, the second arc is $x_2 \rightarrow x_5$, and the third one is $x_2 \rightarrow x_6$. For the latter, because the two associated nodes have been used, the consistency test should be undertaken. The time delays associated with the three arcs are 1s that are reliable. Other arcs are built one by one. Note that arc $x_1 \rightarrow x_5$ is ignored because the time delay is 49, which fails to pass the consistency test. The obtained SDG is shown in Fig. 9.

However, from reachability analysis based on the SDG, arc $x_2 \rightarrow x_4$ does not make sense and should be deleted because there is no path connecting them. From process knowledge, they are parameters on different pipes, and in the P&ID of Fig. 5, they are independent. One explanation for this link is that there is another cause or upstream unit resulting in the changes in both of them. This incorrect arc may bring out wrong results if the root cause is x_2 because x_4 does not depend on it.

5.4. Application of SDGs in fault propagation analysis. Given a validated SDG, fault propagation can be analyzed qualitatively based on consistent paths. This is important in any conclusive significant HAZOP analysis and also in fault detection and isolation, especially root cause analysis, where SDGs can help.

In this case, a fault propagation path is shown by the dotted lines in Fig. 6 meaning that the pressure change in the distributor (x_1) can affect the parameters in cyclo-packs and pump boxes (x_5 and x_6) in turn. During the HAZOP study, all the consistent paths should be considered and the corresponding consequences should be evaluated. If the domain of influence of one variable is large, the intensity of influence is strong, or the consequence is severe, then some appropriate measures should be taken. In this case study, x_1 is important because it has wide influence on almost all downstream variables. Thus the controller on it is well tuned and the line prioritization is implemented

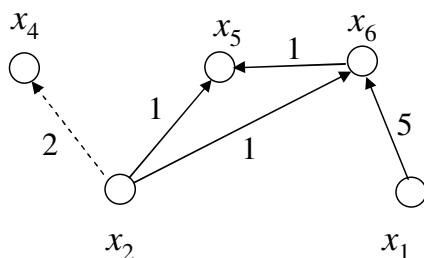


Fig. 9. Causal network obtained by correlation analysis. Solid and dashed lines are positive and negative correlations, respectively. Numbers are the estimated time delays.

to lessen the risk. On the other hand, root cause analysis is undertaken online. When the variables on this path are showing disturbances, then one can trace immediately the starting point and that may be identified as the root cause of fault propagation. The automation of this procedure will help operators quickly identify the symptom of the abnormal situation.

6. Concluding remarks

In this paper, modeling and fault propagation analysis based on SDGs are proposed for application to large-scale industrial systems. The SDG has its advantage of simplicity and good matching with process data. It is feasible even when no model is available. Thus it has great potential in many process analysis applications. Although both process knowledge and process data can be used to capture the causality between variables, neither of them can be solely used to find the true causality without validation.

For the validation of SDGs modeled by process knowledge, especially by P&IDs, two typical temporal-domain data-based methods are described. Cross-correlation is relatively easy to compute and thus very practical for determining signs and estimating time delays. As a more general but more costly computational measure, transfer entropy gives a complementary validation from the information transfer perspective. In both methods, arcs can be validated one by one, as matrices \mathbf{P} and Λ are not necessary, and hence the total computational effort is proportional to the number of arcs. The above two points of view are illustrated in this paper to show the fusion of two data sources in process modeling. Based on the validated model, fault propagation analysis can be performed.

Sometimes, process data are easy to obtain while process knowledge is insufficient. For this case, data-based methods can be applied first to build a causal network, and then reachability check can be applied to validate this network by process knowledge. Moreover, if the causal network is not important whereas only some causality relations are of concern, then this method is more useful.

Acknowledgment

The authors are grateful for the financial aid from NSFC (60736026 and 60904044), NSERC SPG and IRC at the University of Alberta and the Tsinghua National Laboratory for Information Science and Technology (TNList) Cross-discipline Foundation. Data access from Dr. Ramesh Kadali at Suncor Energy Inc. is gratefully acknowledged.

References

- Alonso, C.J., Llamas, C., Maestro, J.A. and Pulido, B. (2003). Diagnosis of dynamic systems: A knowledge model that

- allows tracking the system during the diagnosis process, in P.W.H. Chung, C. Hinde and M. Ali (Eds.), *Developments in Applied Artificial Intelligence*, Lecture Notes in Artificial Intelligence, Vol. 2718, Springer, Berlin, pp. 208–218.
- Bauer, M., Cox, J.W., Caveness, M.H., Downs, J.J. and Thornhill, N.F. (2007). Finding the direction of disturbance propagation in a chemical process using transfer entropy, *IEEE Transactions on Control Systems Technology* **15**(1): 12–21.
- Bauer, M. and 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, H., Tikkala, V.-M., Zakharov, A., Myller, T. and Jamsa-Jounela, S.L. (2011). Application of the enhanced dynamic causal digraph method on a three-layer board machine, *IEEE Transactions on Control Systems Technology* **19**(3): 644–655.
- Fagarasan, I., Ploix, S. and Gentil, S. (2004). Causal fault detection and isolation based on a set-membership approach, *Automatica* **40**(12): 2099–2110.
- Fedai, M. and Drath, R. (2005). CAEX—A neutral data exchange format for engineering data, *Automation Technology in Practice—ATP International* **1**(3): 43–51.
- Gigi, S. and Tangirala, A.K. (2010). Quantitative analysis of directional strengths in jointly stationary linear multivariate processes, *Biological Cybernetics* **103**(2): 119–133.
- Górny, B. (2001). *Consistency-Based Reasoning in Model-Based Diagnosis*, Ph.D. thesis, AGH University of Science and Technology, Cracow.
- Iri, M., Aoki, K., Oshima, E. and Matsuyama, H. (1979). An algorithm for diagnosis of system failures in the chemical process, *Computers and Chemical Engineering* **3**(1–4): 489–493.
- Jan, A., Jonas, B., Erik, F., Krysander, M. and Lars, N. (2007). Safety analysis of autonomous systems by extended fault tree analysis, *International Journal of Adaptive Control and Signal Processing* **21**(2–3): 287–298.
- Ligeza, A. (1996). A note on systematic conflict generation in CA-EN-type causal structures, *LAAS Report No. 96317*, Laboratory for Analysis and Architecture of Systems, Toulouse.
- Ligeza, A. and Kościelny, J.M. (2008). A new approach to multiple fault diagnosis: A combination of diagnostic matrices, graphs, algebraic and rule-based models. The case of two-layer models, *International Journal of Applied Mathematics and Computer Science* **18**(4): 465–476, DOI: 10.2478/v10006-008-0041-8.
- Korbicz, J. and Kościelny, J.M. (Eds.) (2010). *Modeling, Diagnostics and Process Control: Implementation in the Diaster System*, Springer-Verlag, Berlin/Heidelberg.
- Leyval, L., Gentil, S. and Feray-Beaumont, S. (1994). Model-based causal reasoning for process supervision, *Automatica* **30**(8): 1295–1306.
- Li, Q. and Racine, J.S. (2007). *Nonparametric Econometrics: Theory and Practice*, Princeton University Press, Princeton, NJ.
- Lungarella, M., Ishiguro, K., Kuniyoshi, Y. and Otsu, N. (2007). Methods for quantifying the causal structure of bivariate time series, *International Journal of Bifurcation and Chaos* **17**(3): 903–921.
- Maurya, M.R., Rengaswamy, R. and Venkatasubramanian, V. (2003a). A systematic framework for the development and analysis of signed digraphs for chemical processes: 1. Algorithms and analysis, *Industrial and Engineering Chemistry Research* **42**(20): 4789–4810.
- Maurya, M.R., Rengaswamy, R. and Venkatasubramanian, V. (2003b). A systematic framework for the development and analysis of signed digraphs for chemical processes: 2. Control loops and flowsheet analysis, *Industrial and Engineering Chemistry Research* **42**(20): 4811–4827.
- Maurya, M.R., Rengaswamy, R. and Venkatasubramanian, V. (2004). Application of signed digraphs-based analysis for fault diagnosis of chemical process flowsheets, *Engineering Applications of Artificial Intelligence* **17**(5): 501–518.
- Montmain, J. and Gentil, S. (2000). Dynamic causal model diagnostic reasoning for online technical process supervision, *Automatica* **36**(8): 1137–1152.
- Mosterman, P.J. and Biswas, G. (1999). Diagnosis of continuous valued systems in transient operating regions, *IEEE Transactions on Systems, Man, and Cybernetics: Part A* **29**(6): 554–565.
- Oyeleye, O.O. and Kramer, M.A. (1988). Qualitative simulation of chemical process systems: Steady-state analysis, *AIChE Journal* **34**(9): 1441–1454.
- Palmer, C. and Chung, P.W.H. (1999). Verifying signed directed graph models for process plants, *Computers & Chemical Engineering* **23**(Suppl. 1): S394–S391.
- Palmer, C. and Chung, P.W.H. (2000). Creating signed directed graph models for process plants, *Industrial and Engineering Chemistry Research* **39**(20): 2548–2558.
- Pastor, J., Lafon, M., Trave-Massuyes, L., Demonet, J.-F., Doyon, B. and Celsis, P. (2000). Information processing in large-scale cerebral networks: The causal connectivity approach, *Biological Cybernetics* **82**(1): 49–59.
- Schreiber, T. (2000). Measuring information transfer, *Physical Review Letters* **85**(2): 461–464.
- Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London/New York, NY.
- Staroswiecki, M. (2000). Quantitative and qualitative models for fault detection and isolation, *Mechanical Systems and Signal Processing* **14**(3): 301–325.
- Tangirala, A.K., Shah, S.L. and Thornhill, N.F. (2005). PSC-MAP: A new tool for plant-wide oscillation detection, *Journal of Process Control* **15**(8): 931–941.
- Thambirajah, J., Benabbas, L., Bauer, M. and Thornhill, N.F. (2009). Cause-and-effect analysis in chemical processes utilizing XML: Plant connectivity and quantitative process history, *Computers and Chemical Engineering* **33**(2): 503–512.

Yang, F., Shah, S.L. and Xiao, D. (2010a). SDG (Signed Directed Graph) based process description and fault propagation analysis for a tailings pumping process, *Proceedings of the 13th IFAC Symposium on Automation in Mining, Mineral and Metal Processing, Cape Town, South Africa*, pp. 50–55.

Yang, F., Xiao, D. and Shah, S.L. (2010b). Qualitative fault detection and hazard analysis based on signed directed graphs for large-scale complex systems, in W. Zhang (Ed.), *Fault Detection*, IN-TECH, Vukovar, pp. 15–50.

Yim, S. Y., Ananthakumar, H.G., Benabbas, L., Horch, A., Drath, R. and Thornhill, N.F. (2006). Using process topology in plant-wide control loop performance assessment, *Computers and Chemical Engineering* **31**(2): 86–99.



Fan Yang earned a B.Eng. degree in 2002 and a Ph.D. degree in 2008, both from Tsinghua University. From 2009 to 2011 he worked as a visiting professor and PDF at the University of Alberta. Currently he is a lecturer at Tsinghua University. He obtained the Young Research Paper Award from the IEEE CSS Beijing Chapter in 2006. His research interests include topology modeling of large-scale processes, abnormal events monitoring, process hazard analysis, and smart

alarm management.



Sirish L. Shah has been with the University of Alberta since 1978, where he currently holds the NSERC-Matrikon-Suncor-iCORE Senior Industrial Research Chair in Computer Process Control. He was the recipient of the Albright & Wilson Americas Award of the Canadian Society for Chemical Engineering (CSCHE) in recognition of distinguished contributions to chemical engineering in 1989, the Killam Professor in 2003, and the D.G. Fisher Award of the CSCHE

for significant contributions in the field of systems and control. The main area of his current research is fault detection and diagnosis, alarm design and analysis, process and performance monitoring, system identification as well as design and implementation of soft sensors.



Deyun Xiao is a professor in the Department of Automation at Tsinghua University. Since his graduation from Tsinghua University in 1970, he has been with this department as a faculty member. His research interests cover a wide range of topics related to the area of control science and engineering, including fault diagnosis and hazard assessment, system identification and parameter estimation, multi-sensor data fusion, intelligent measurement and control, intelligent sensing techniques, computer control systems, real-time database systems, industrial monitoring software, enterprise comprehensive automation systems, and intelligent transportation systems.

Received: 18 January 2011

Revised: 20 July 2011