

Chapter 2

Examples of Applications for Connectivity and Causality Analysis

Abstract Connectivity and causality have a lot of potential applications, among which we focus on analysis and design of large-scale complex industrial processes. A direct application by establishing connectivity and causality is to build a topological model before parameter identification for complex industrial processes that are usually multi-input, multi-output systems with many internal closed loops. In abnormal situation management, process topology can be employed for root cause analysis, risk analysis, and consequential alarm identification using the information of fault propagation. These potential applications include both off-line analysis and on-line diagnosis. In addition, process topology can eventually be used in design of control structures because process topology determines the natural structure of the distributed plant-wide control.

Keywords System identification · Closed loop · Vector autoregressive models · Root cause analysis · Hazard and operability study · Consequential alarms · Abnormal situation · Plant-wide control structures · Process flow diagrams · Sensor location

Among various applications of connectivity and causality, we will focus on analysis and design of large-scale complex industrial processes. Existing studies have shown great potential of applying connectivity and causality analysis to such cases; the illustration of these applications will also highlight different approaches for causality identification and analysis.

2.1 Topology Modeling and Closed-Loop Identification

A direct application by establishing connectivity and causality is to build a topology model for a complex industrial process. Given process data, system identification is the typical black-box modeling approach. If a known system structure is assumed, there are plenty of methods to estimate parameters. However, in multivariate cases,

structure identification is more important and should be performed before parameter estimation. In particular, what we mean by structure here is not only the orders of the local linear models, but also the linkage between variables.

When there are many process variables, it is unwise to separate them into inputs and outputs, because each of them reveals some information of the system and can thus be regarded as a state variable. As a result, traditional quantitative system identification techniques do not work well, and one way is to assume the topology before estimating the orders and parameters of each closed-loop (bidirectional) path model, namely, the local linear model between every two variables. However, there are too many combinations according to the existence of a link between every two variables. Thus, it is more reasonable to assume that each pair of variables are linked together and then estimate the orders and parameters of the path model; if the results show that the link is too weak, then such a link can be removed [9, 10]. This only requires a single-input-single-output (SISO) framework to deal with every path. However for a rigorous analysis, a multiple-input-multiple-output (MIMO) framework, such as a vector autoregressive (VAR) model, should be considered because every variable in a multivariate system may influence as well as be influenced by more than one variable. The above idea suffers from a high computational burden, yet if the topology is known a priori, then the computational burden can be lowered significantly. For this purpose, topology modeling based on process connectivity capturing or process data analytics would help.

2.2 Root Cause Analysis

When the system encounters an abnormal situation, there must be one or more elements showing abnormal symptoms or measurements. If there is only one abnormal element, then this is a local fault in most cases and one should then look into the commensurate part to figure out the problem. If there are multiple abnormal elements, we should be aware that this could be due to some interaction that results in propagation of the source fault. For example, in a pipe network, if an upstream valve is partially blocked, then there will be a series of abnormal events downstream, e.g., reduction of flow rate, decrease of liquid level, and even dry-out of a vessel. When an operator finds that there is something wrong in such a process, there may exist multiple abnormal symptoms; to resolve this situation, the operator should not just tune the valves associated with the vessel for example, because this may make the situation even worse; instead he or she should find the root cause promptly and eliminate it. Once the root cause is resolved, all the other issues disappear accordingly.

Given the topology, or connectivity/causality to be specific, a backward traversal along the paths can be performed to find the root cause, namely, the original abnormal element that causes all the other abnormal elements [17]. What we assume here is that the fault should propagate along the established paths; this is the case most of the times. Among other events the abnormal situations considered in the examples

here and generally include cases such as deviation from normal values, oscillations, sensor or actuator malfunction, process or equipment failure, and misoperation.

Take oscillating variables as an example, which is a typical plant-wide disturbance. By using some data-driven methods, oscillating variables can be identified, which are also called efficient nodes in the terminology of signed digraphs because they are the nodes that should be studied. Jiang et al. [11] used a control loop digraph to describe the topology of control loops and, by examining the domain of influence of each control loop were able to find a ranked list of root cause variables to be those that are able to reach all the other oscillating variables along the paths. For a survey of this application, please refer to [5]. Similar work has also been reported in the early study of [4].

2.3 Risk Analysis: HAZOP

Risk analysis is a way to examine a process to identify and evaluate problems that may represent risks. As a representative qualitative approach, hazard and operability study (HAZOP) is frequently applied to planned or existing processes in a structured and systematic way. This task is carried out based on guide words by a series of team meetings. If the topology is available, then this procedure can be relatively straight forward and clear. There are several other studies that use signed digraphs or other graph models for HAZOP study [14–16, 19]. In [20], HAZOP is considered as one of the two main areas of the signed digraph technology (fault diagnosis as another one, as mentioned in the previous section) by using the inference engine essentially based on the search of process topology. Different from root cause analysis, such search is a forward search to find the resulting consequence while the former is a backward search to find the root. The purpose of HAZOP analysis is to find all possible consequences of any assumed faults. But if one wants to estimate the probability of events, quantitative information needs to be incorporated. With such a scheme one can obtain a computer-aided HAZOP analysis.

2.4 Consequential Alarm Identification

Alarm management is an emerging area in the process control community [8]. For monitoring of complex industrial processes, a lot of alarms tags are configured for all kinds of variables. For example, a process variable can trigger high/low alarms to reflect its states. During abnormal situations, alarms should be raised to remind operators to take actions. Ideally, one abnormal situation should trigger one and only one alarm; however, because of redundancy, interactions and correlations between variables even a single abnormal event will result in the annunciation of many alarms. In addition, since a fault can propagate throughout the process, alarms also show up in a specific order. This list of consecutive alarms may be dependent; thus we call

them consequential alarms [13]. Consequential alarms over a very short period of time often lead to be construed as alarm floods, leading to a dangerous situation as the console operators or engineers may not be able to identify true root causes. For this case, process topology would be of great help in describing the relationship between alarm tags.

For on-line analysis, this is similar to the root cause analysis because the most important task is to find the root cause of all related or consequential alarms. If the root cause is resolved, then all the alarms can be removed. For off-line analysis, more can be done, for example, to obtain and analyze alarm sequences of typical abnormal situations [3, 12]. When an alarm flood occurs, the alarm sequence is recorded and compared to recorded known sequences to find the most possible root causes so that the previously known and successful alarm mitigation solution can be retrieved immediately; this approach can also be adapted for on-line applications.

2.5 Plant-Wide Control Structure Design

Connectivity and causality reflect the essential nature of a process, so the above applications are all based on a given topology and are aimed at analysis. Such topology can eventually be used in design of control structures because process topology determines the natural structure of the distributed plant-wide control. There are a few studies in this area: Alabi used process flow diagrams (PFDs) in degrees of freedom (DOF) analysis [1]; Cameron and Hangos discussed observability and controllability studies based on structural information [6]; and Hangos and Tuza used graph-theoretical models in optimal control structure selection [7]. These applications of topology can serve as a precursor for other complex and quantitative applications.

Another application of process topology is its use in sensor location. For example, graph models have been used to design feasible and optimal sensor location strategies according to fault detectability and identifiability criteria [2, 18].

2.6 Chapter Summary

Several potential applications of process topology/connectivity and causality have been introduced in this chapter, including modeling, analysis and control structure design. These are just a few among many applications that are likely to be pursued further in the future. It should be noted that qualitative topology has to be incorporated with quantitative information before a comprehensive application, and the topology should be adapted to different application requirements.

To develop a formal framework for these applications, we first need to formalize the description of topology; this is the objective of the next Chap. 3.

References

1. Alabi DB (2010) Automated analysis of control degree of freedom. Master's thesis, Department of chemical engineering and chemical technology, Imperial College London, London
2. Bhushan M, Rengaswamy R (2002) Comprehensive design of a sensor network for chemical plants based on various diagnosability and reliability criteria. 1. framework. *Ind Eng Chem Res* 41:1826–39
3. Cheng Y, Izadi I, Chen T (2013) Pattern matching of alarm flood sequences by a modified smith-waterman algorithm. *Chem Eng Res Des* 91(6):1085–94
4. Chiang LH, Braatz RD (2003) Process monitoring using causal map and multivariate statistics: fault detection and identification. *Chemometr Intell Lab Syst* 65:159–78
5. Duan P, Shah SL, Chen T, Yang F (2014) Methods for root cause diagnosis of plant-wide oscillations. *AIChE J*. doi:[10.1002/aic.14391](https://doi.org/10.1002/aic.14391)
6. Hangos K, Cameron I (2001) Process modelling and model analysis. Academic Press, London
7. Hangos KM, Tuza Z (2001) Optimal control structure selection for process systems. *Comput Chem Eng* 25(11–12):1521–1536
8. Izadi I, Shah SL, Shook D, Chen T (2009) An introduction to alarm analysis and design. In: Proceedings of 7th IFAC symposium on fault detection, supervision and safety of technical processes, Barcelona, Spain, pp 645–650
9. Jiang B, Yang F, Jiang Y, Huang D (2012) An extended AUDI algorithm for simultaneous identification of forward and backward paths in closed-loop systems. In: Proceedings of 2012 international symposium on advanced control of chemical processes, Singapore, pp 396–401.
10. Jiang B, Yang F, Huang D, Wang W (2013) Extended-AUDI method for simultaneous determination of causality and models from process data. In: Proceedings of 2013 American control conference, Washington, pp 2491–2496.
11. Jiang H, Patwardhan R, Shah SL (2009) Root cause diagnosis of plant-wide oscillations using the concept of adjacency matrix. *J Process Control* 19(8):1347–54
12. Nishiguchi J, Takai T (2010) Ipl2 and 3 performance improvement method for process safety using event correlation analysis. *Comput Chem Eng* 34(12):2007–13
13. Rothenberg D (2009) Alarm Manag Process Control. Momentem Press, New York
14. Srinivasan R, Venkatasubramanian V (1998a) Automating hazop analysis of batch chemical plants: part i. the knowledge representation framework. *Comput Chem Eng* 22(9):1345–5
15. Srinivasan R, Venkatasubramanian V (1998b) Automating hazop analysis of batch chemical plants: part ii. algorithms and application. *Comput Chem Eng* 22(9):1357–70
16. Venkatasubramanian V, Zhao J, Viswanatha S (2000) Intelligent system of hazop analysis of complex process plants. *Comput Chem Eng* 24:2291–302
17. Yang F, Xiao D (2012) Progress in root cause and fault propagation analysis of large scale industrial processes. *J Control Sci Eng* 2012:10 (Article ID 478373)
18. Yang F, Xiao D, Shah SL (2009) Optimal sensor location design for reliable fault detection in presence of false alarms. *Sensors* 9(11):8579–92
19. Yim SY, Ananthakumar HG, Benabbas L, Horch A, Drath R, Thornhill NF (2006) Using process topology in plant-wide control loop performance assessment. *Comput Chem Eng* 31(2):86–99
20. Zhang Z, Wu C, Zhang B, Xia T, Li A (2005) Sdg multiple fault diagnosis by real-time inverse inference. *Reliab Eng Syst Saf* 87:173–89