Incipient Fault Diagnosis of Tennessee Eastman Flowsheet using Signed Directed Graph and Trend Analysis

Mano Ram Maurya, Raghunathan Rengaswamy[1,2] and Venkat Venkatasubramanian[2]
Laboratory for Intelligent Process Systems, School of Chemical Engineering,
Purdue University, West Lafayette, IN 47907, USA.

Abstract
In this article we propose a combined signed directed graph (SDG) and qualitative trend analysis (QTA) framework for incipient fault diagnosis. The SDG is the first level in our framework and provides a possible candidate set of faults based on the incipient response of the process. The search for the actual fault is performed based on a QTA, which uses the temporal evolution of the sensors for further resolution. It is shown that this framework provides fast, reliable and accurate incipient fault diagnosis.
Key Words: Qualitative Trend Analysis, SDG, Incipient Fault Diagnosis, Tennessee Eastman Challenge Problem

1 Introduction

Quick fault detection and diagnosis are key issues for safe and optimal process operation. Quick corrective action can help in minimizing the quality and productivity offsets, and reduce the hazardous consequences in abnormal situations and so on. Further, quick detection and diagnosis allows more time to the operator for counter-measure planning. Hence incipient fault diagnosis (IFD) is an important area of research. Among many models such as fault trees, completely numerical models, probabilistic models etc., signed directed graphs (SDG) have been widely used for IFD. Despite considerable work on SDG, large-scale applications are seldom seen in the literature. Also, till recently, very little attention has been given to automatic development of digraphs and their analysis.
Recently, Maurya *et al.* [5] have made an attempt towards thorough analysis of the algorithms for digraph generation, methodologies for digraph analysis and their applications. It is well known that similar to many other qualitative techniques, SDG-based analysis suffers from generation of numerous spurious solutions due to loss of information while going from quantitative to qualitative domain. Hence integration of SDG-based analysis and some quantitative approach is highly desirable to reduce the number of spurious solutions. Qualitative trend-analysis (QTA), being a data-driven technique, is well-suited for the above task [3, 4]. Further, selection of key variables to be measured plays a key role in IFD and hence its very important to select and use the correct sensors for fault diagnosis (FD). In this paper, we propose a two-level strategy for incipient fault diagnosis. The first level uses SDG for incipient fault diagnosis. The sensors that are used for SDG-based fault diagnosis are chosen using the algorithm proposed by Bhushan and

[1]Current address: CAMP 243, Department of Chemical Engineering, Clarkson University, Potsdam NY 13699-5705 (USA)
[2]Authors to whom all correspondences should be addressed, E-mail: venkat@ecn.purdue.edu, raghu@clarkson.edu

Rengaswamy [1]. QTA is used in the second level to improve resolution. This paper is organized as follows.
In the next section we present an overview of our previous work on SDG development and analysis. We also briefly present SDG-based sensor-location. Integration of SDG and a fuzzy qualitative trend matching based diagnostic approach for IFD is discussed in section 3. Finally, in section 4, the utility of the proposed approach is demonstrated on the Tennessee Eastman case study. Due to space constraint, only minimal essential description of various concepts has been provided.

2 SDG Analysis and SDG-based Sensor Location

In this section, we discuss our previous work on the generation and analysis of SDG followed by SDG-based sensor location.

• **Algorithms for SDG generation:** Algorithms for generating digraphs for systems described by differential equations (DE), algebraic equations (AE) and differential algebraic equations (DAE) of index one have been presented by Maurya *et al.* [5]. Essentially, the explicit causality (for DE systems) or pseudo causality (precedence ordering in AE systems) is captured through directed arcs in a digraph.

• **Methods for SDG analysis:** An exhaustive analysis of the SDGs for various systems has been presented by Maurya *et al.* [5]. It has been shown that initial response of DE systems (and DAE systems with only one perfect matching) can be predicted by using propagation through the shortest path(s) in the SDG and an initial-response table (IRT) can be prepared. This table is used in designing sensor-location. SDG-based IFD is performed by comparison of the measured sign pattern with the sign patterns in the IRT corresponding to various faults.

• **SDG-based sensor-location:** Bhushan and Rengaswamy [1] have presented an algorithm for SDG-based sensor-network design. The IRT is used to formulate the optimization (selection of minimum number of sensors to guarantee observability and to ensure maximum fault resolution) problem. Fault resolution is maximized by selecting at least one sensor from the sets of discriminating sensors (identified by calculating the symmetric difference between the sets of affected sensors by two different faults) corresponding to all fault pairs.

3 Integrated Framework for Incipient Fault Diagnosis

The proposed (novel) integrated framework is based on a two-step architecture. In typical large-scale plants, the number of possible faults (including external disturbances, parametric faults, control loop malfunctions etc.) may be as large as thousands. In a data-driven approach like QTA, the similarity index between the measured signal and the signal corresponding to an assumed fault is between 0 and 1 [3, 4]. As the number of possible faults increases, the separation between similarity indices corresponding to various fault scenarios decreases and decision making (based upon comparison with a threshold) becomes difficult. Also, computational complexity increases linearly with respect to the number of fault scenarios stored in the fault-signature database. SDG-based incipient fault diagnosis is a viable alternative to reduce the size of the set of candidate faults for a given scenario to be diagnosed. This forms the first step of the framework.

The SDG-based analysis is robust in the qualitative linear regime (the region in which the sign of the arcs in the SDG do not change) and hence is used in the first level to enhance completeness.
Thus the first step of the framework uses SDG-based IFD to reduce the size of the set of candidate faults. The second step uses QTA-based fuzzy trend matching between the signal to be diagnosed and the signals corresponding to the reduced set of candidate faults. The essential idea is to allow 0/1 as well as fractional matching index between two primitives. An interested reader is referred to the articles by Dash *et al.* [3, 4]. It is expected that the number of faults in the reduced set would be much less as compared to the set of all the fault scenarios considered in the database. Thus the separation between the concerned similarity indices would increase and computational complexity would decrease. The integrated framework is shown in Figure 1. A case study is presented in the next section to show the application of the SDG-QTA framework.

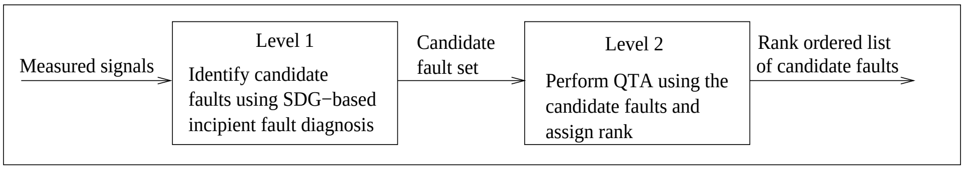

Figure 1: SDG-QTA based integrated framework for incipient FD

4 Case Study: Fault Diagnosis of Tennessee Eastman Flowsheet

The Tennessee Eastman problem has been considered as a benchmark problem for a number of applications such as MPC, sensor-location, causal-map based FD etc. Only FD-relevant description is given here. A detailed description is presented by Bhushan and Rengaswamy [2]. There are 33 fault nodes (16 bidirectional faults and 1 unidirectional fault) in this case study (including three control loops). These faults have been categorized by their sign ('+' or '-') and type (such as sensor bias, control valve bias *etc.*). Total number of nodes in the SDG is 165. Only single fault scenarios have been considered in this case study. The SDG-based sensor location is used to identify placement of sensors. Based upon the sensor-network thus obtained, 14 process variables (explicit notation/description not provided) are measured. Sampling rate is 100 samples per hour. SDG-QTA based IFD is presented next.

4.1 SDG-QTA based incipient fault diagnosis

As mentioned in section 2, an initial response table is prepared. The sign of the initial evolution of measured variables is compared with sign pattern in the initial response table to find the candidate fault set. Thus SDG-based IFD can be performed as soon as non-zero signs are detected. During the second step, QTA is performed explicitly on the candidate fault set (generated by SDG-based IFD) only and hence correct measurement of the sign of initial evolution is extremely important. As the signals evolve with time, the qualitative trends (represented as a string of primitives) of measurements are continuously compared with the signatures of the candidate faults and the candidate faults are rank ordered in decreasing order of similarity measure between the measured signals and the fault signatures stored in the database. The most probable fault is the fault with highest similarity measure.

4.2 Results

All the single fault scenarios (including malfunction inside control loops) have been tested. The results (of 27 fault scenarios[3]) are presented in Table 1. The faults {16-18, 25-27} and {19-21, 28-30} are related to control loops 1 and 2, respectively. The second column of Table 1 shows the set of candidate faults generated by SDG-based IFD. The remaining columns list the top two faults with highest similarity measures at 4 different times as the process evolves. Figure 2 shows that similarity measure between two similar fault scenarios is high (fairly above 0.5). The similarity measures shown in Table 1 and Figure 2 are the minimum of the similarity measures among 14 signals. Thus the highest similarity measure (for the most probable fault) need not be very close to one. Important results from Table 1 are-

1. The size of candidate fault sets is about 20% of the total number of fault scenarios. Thus computational complexity of level 2 is reduced considerably (5 times faster).
2. The similarity measure of the second most probable fault is well below the highest similarity measure. Thus a better similarity measure spread is achieved.
3. For the introduced fault scenarios 6, 7, 14 and 15, the SDG-based candidate fault set (column 2) does not include the actual fault. Thus four false negatives and false positives are generated as a result of incorrect measurement of the initial response due to large amount of noise. The true signal gets masked by the large noise during initial evolution. Another reason is that the system might not exhibit the initial response for a large enough time period for it to be detected. However, this is not a serious concern due to the very small sampling rates that can be achieved in the DCS nowadays. Thus as long as noise level is low, SDG-based IFD approach should be generally applicable. Ofcourse, it is assumed that a good model (that captures the initial response well in the absence of noise) is used to develop the SDG to start with. Another important observation in this case study is the following. A signal with low signal to noise ratio (SNR) is not useful for QTA-based FDD as useful features get masked by the noise. As the SNR becomes very high, a large number of primitives are identified to represent the data which might be undesirable. One way to overcome this problem would be to let the noise level in the QTA algorithm be a tunable parameter. However, in general, the QTA is able to capture the important features of the sensor trends to provide a high resolution of faults in the proposed integrated framework.

5 Conclusions

Based on SDG and QTA, a two level framework has been presented for incipient fault diagnosis. The application of the integrated framework has been shown through Tennessee Eastman case study. The single fault scenarios have been tested exhaustively. Correct fault diagnosis is achieved for 85% of the cases. The proposed framework combines the completeness property of SDG with the resolution property of the QTA to provide a promising framework for incipient fault diagnosis.

[3]Fault types are: *s*- sensor bias, *set*- set point change, *v*- control valve bias, *o*- other faults.

Table 1: SDG-QTA based incipient FD

Actual fault (no.)	SDG-based fault set (Incipient FD)	QTA-based FD (top two fault with highest similarity)			
		3 hrs	6 hrs	9 hrs	12 hrs
1 [o, +]	1 4 8 14 16	1 (0.49), 4 (0.41)	1 (0.73), 4 (0.33)	1 (0.78), 4 (0.22)	1 (0.81), 4 (0.32)
2 [o, +]	2 8 14 18	2 (0.81), 18 (0.31)	2 (0.85), 18 (0.36)	2 (0.91), 18 (0.31)	2 (0.72), 18 (0.38)
3 [o, +]	3 8 14 16	3 (0.75), 8 (0.02)	3 (0.83), 8 (0.35)	3 (0.81), 8 (0.30)	3 (0.81), 8 (0.26)
4 [o, +]	1 4 8 14 16	4 (0.72), 14 (0.36)	4 (0.69), 14 (0.56)	4 (0.79), 14 (0.36)	4 (0.85), 14 (0.41)
5 [o, +]	1 4 5 6 8 27	5 (0.83), 27 (0.44)	5 (0.64), 27 (0.53)	5 (0.75), 27 (0.40)	5 (0.74), 27 (0.51)
6 [o, +]	8 27	27 (0.14), 8 (0.00)	27 (0.23), 8 (0.04)	27 (0.38), 8 (0.11)	27 (0.27), 8 (0.09)
7 [o, +]	3 8 14 19 25	8 (0.13), 25 (0.10)	8 (0.30), 25 (0.15)	8 (0.22), 25 (0.20)	8 (0.30), 25 (0.19)
8 [o, -]	7 8 14 27	8 (0.83), 14 (0.40)	8 (0.89), 14 (0.27)	8 (0.68), 14 (0.34)	8 (0.65), 14 (0.31)
9 [o, -]	6 8 9 12 25	9 (0.78), 12 (0.57)	9 (0.69), 12 (0.60)	9 (0.75), 12 (0.59)	9 (0.70), 12 (0.48)
10 [o, -]	6 8 10 27	10 (0.89), 27 (0.38)	10 (0.75), 27 (0.31)	10 (0.75), 27 (0.35)	10 (0.79), 27 (0.35)
11 [o, -]	6 8 11 25	11 (0.81), 8 (0.19)	11 (0.75), 8 (0.16)	11 (0.73), 8 (0.16)	11 (0.80), 8 (0.19)
12 [o, -]	8 9 12 18	12 (0.81), 9 (0.47)	12 (0.78), 9 (0.54)	12 (0.80), 9 (0.56)	12 (0.82), 9 (0.50)
13 [o, -]	8 9 12 13 14 18	13 (0.77), 8 (0.33)	13 (0.73), 8 (0.58)	13 (0.76), 8 (0.50)	13 (0.72), 8 (0.66)
14 [o, -]	8 18 30	8 (0.38), 30 (0.30)	8 (0.29), 30 (0.28)	8 (0.39), 30 (0.35)	8 (0.39), 30 (0.36)
15 [o, -]	6 7 8 16	8 (0.60), 7 (0.08)	8 (0.71), 7 (0.10)	8 (0.51), 7 (0.22)	8 (0.64), 7 (0.24)
16 [s, +]	8 14 16	16 (0.72), 14 (0.06)	16 (0.78), 14 (0.08)	16 (0.58), 14 (0.08)	16 (0.74), 14 (0.12)
17 [set, +]	6 8 17	17 (0.78), 8 (0.59)	17 (0.90), 8 (0.64)	17 (0.73), 8 (0.58)	17 (0.68), 8 (0.53)
18 [v, +]	6 8 18	18 (0.70), 8 (0.38)	18 (0.84), 8 (0.49)	18 (0.78), 8 (0.38)	18 (0.78), 8 (0.44)
19 [s, +]	3 8 14 16 19	19 (0.86), 16 (0.35)	19 (0.67), 16 (0.51)	19 (0.75), 16 (0.36)	19 (0.77), 16 (0.44)
20 [set, +]	8 11 20 25	20 (0.94), 25 (0.40)	20 (0.91), 25 (0.61)	20 (0.80), 25 (0.54)	20 (0.78), 25 (0.53)
21 [v, +]	3 8 16 21	21 (0.82), 16 (0.14)	21 (0.72), 16 (0.25)	21 (0.81), 16 (0.28)	21 (0.81), 16 (0.39)
25 [s, -]	6 8 25	25 (0.81), 8 (0.59)	25 (0.87), 8 (0.59)	25 (0.81), 8 (0.63)	25 (0.63), 8 (0.54)
26 [set, -]	8 14 26	26 (0.86), 8 (0.06)	26 (0.78), 8 (0.08)	26 (0.58), 8 (0.08)	26 (0.70), 8 (0.10)
27 [v, -]	8 14 27	27 (0.79), 8 (0.11)	27 (0.83), 8 (0.10)	27 (0.74), 8 (0.10)	27 (0.81), 8 (0.10)
28 [s, -]	6 8 11 25 28	28 (0.86), 25 (0.40)	28 (0.75), 25 (0.59)	28 (0.75), 25 (0.54)	28 (0.80), 25 (0.53)
29 [set, -]	3 8 16 29	29 (0.91), 16 (0.35)	29 (0.74), 16 (0.61)	29 (0.77), 16 (0.37)	29 (0.67), 16 (0.44)
30 [v, -]	8 11 25 30	30 (0.77), 8 (0.18)	30 (0.73), 8 (0.32)	30 (0.86), 8 (0.33)	30 (0.81), 8 (0.40)

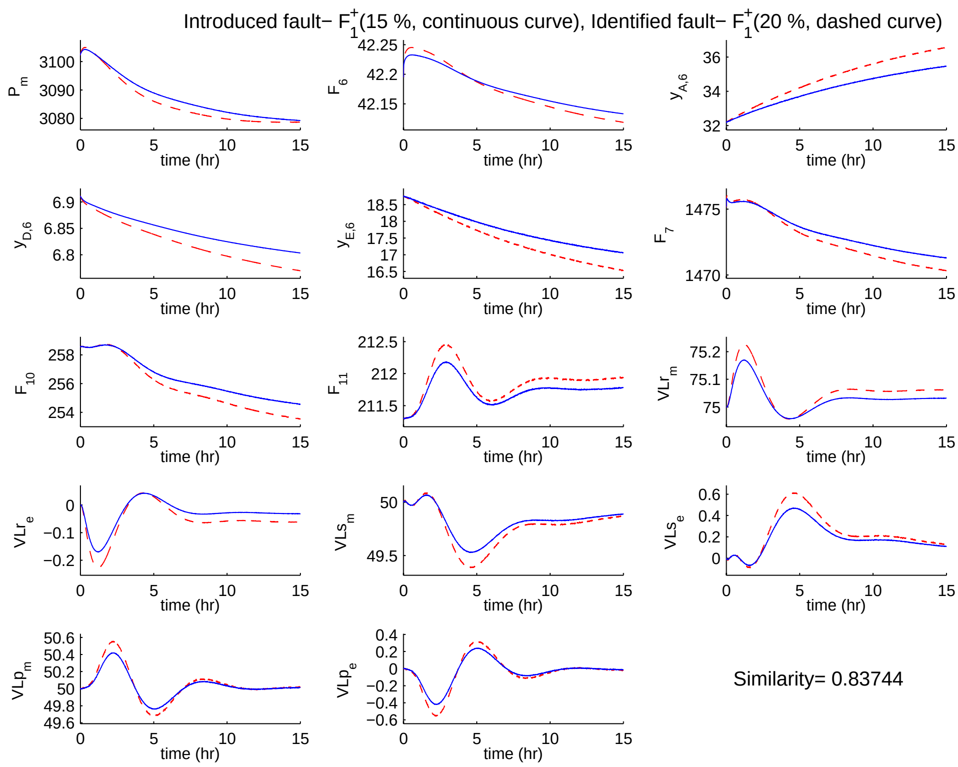

Figure 2: Sensor trends for simulated fault 1 and identified fault 1

References

[1] M. Bhushan and R. Rengaswamy. Design of sensor location based on various fault diagnostic observability anf reliability criteria. *Comput. Chem. Engg.*, 24:735–741, 2000.

[2] M. Bhushan and R. Rengaswamy. Comprehensive design of sensor network for chemical plants based on reliability criteria: II. Applications. In print *Ind. Engng chem. Res.*, 2002.

[3] S. Dash, R. Rengaswamy, and V. Venkatasubramanian. Fuzzy-logic based trend classification for fault diagnosis of chemical processes. Submitted to *Comput. Chem. Engng.*, 2001.

[4] S. Dash, R. Rengaswamy, and V. Venkatasubramanian. A novel interval-halving framework for automated identification of process trends. Submitted to *Comput. Chem. Engng.*, 2001.

[5] M.R. Maurya, R. Rengaswamy, and V. Venkatasubramanian. Systematic development and application of digraphs for process diagnosis and hazards analysis. In 4^{th} *IFAC Workshop on On-Line Fault Detection & Supervision in the Chemical Process Industries (CHEMFAS-4), Korea, June 7-8*, 2001.