2017 IFAC World Congress workshop on:
Learning and Discovery through Process Data Analytics

Speakers: Tongwen Chen, Biao Huang, Sirish L. Shah, Nina Thornhill and Jiandong Wang

Time: Sunday, 9th July 2017; 9:00-17:30

Workshop outline

Process data analytic methods rely on the notion of sensor fusion whereby data from many sensors and alarm tags are combined with process information, such as physical connectivity of process units, to give a holistic picture of health of an integrated plant. The fusion of information from such disparate sources of data is the key step in devising methodologies for a smart strategy for process data analytics

In the context of the application of analytics in the process industry, the objective in this workshop is to introduce participants to tools, techniques and a framework for seamless integration of information from process and alarm databases complemented with process connectivity information. The discovery of information from such diverse and complex data sources can be subsequently used for process and performance monitoring including alarm rationalization, root cause diagnosis of process faults, hazard and operability (Hazop) analysis, safe and optimal process operation. Such multivariate process data analytics involves information extraction from routine process data, that is typically non-categorical (as in numerical process data from sensors), plus categorical (or non-numerical or qualitative and binary) data from Alarm and Event (A&E) logs combined with process connectivity or topology information that can be inferred from the data through causality analysis or as obtained from piping and instrument diagrams of a process. The later refers to the capture of material flow streams in process units as well information flow-paths in the process due to control loops.

Highly interconnected process plants are now common and the analysis of root causes of process abnormality including predictive risk analysis is non-trivial. It is the extraction of information from the fusion of process data, alarm and event data and process connectivity that should form the backbone of a viable process data analytics strategy and this will be the main focus of this workshop.

With the explosion of applications of analytics in diverse areas (such as aircraft engine prognosis, medicine, sports, finance, social sciences and the advertising industry) statistical learning skills are in high demand. The emphasis in this workshop will be on tools and techniques that help in the process of understanding data and discovering information that will lead to predictive monitoring and diagnosis of process faults.

Desired prerequisites for attendees: Basic knowledge of statistics, linear algebra, signal processing, system identification and control.

Target audience: The intended audience for this workshop would be industrial practitioners of control including vendors working in the area of on-line data logging and archiving, graduate students with interests in statistical learning and data science and academics.

Workshop Program
The following topics will be discussed in this workshop. **Each topic will be accompanied by one or more industrial case study to convey the utilitarian value of the learning, discovery and diagnosis from process data.**

- Overview of the broad analytics area with emphasis on its use in the process industry.
- Basic definitions and introduction to supervised and unsupervised learning: simple regression, classification and clustering.
- Data visualization methods; examination and analysis of data in a multivariate framework (in the temporal as well as the spectral domains).
- Data quality assessment: Outlier detection; filtering and data segmentation.
- Elements of statistical inference and learning including Bayesian methods.
- Multivariate methods for data analysis: SVD, PCA, PLS, SVR.
- Case studies on nearest neighbour methods for multivariate detection and diagnosis of transient disturbances.
- Alarm data analysis: Detection and removal of nuisance alarms; root-cause analysis of alarms and alarm floods.
- Data-based causality analysis for identification of process topology.
- Future areas to explore in the use of statistical learning, data science and analytics for improved process operation.

Workshop speakers:
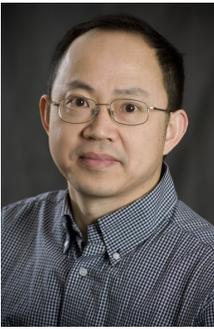
Tongwen Chen

Tongwen Chen is presently a Professor of Electrical and Computer Engineering at the University of Alberta, Edmonton, Canada. His research interests include computer and network based control systems, event triggered control, process safety and alarm systems, and their applications to the process and power industries. He is a Fellow of IEEE, IFAC, as well as Canadian Academy of Engineering.

Biao Huang

Biao Huang received the B.Sc. and M.Sc. degrees in automatic control from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 1983 and 1986, respectively, and the Ph.D. degree in process control from the University of Alberta, Edmonton, AB, Canada, in 1997. He joined the University of Alberta, in 1997, as an Assistant Professor with the Department of Chemical and Materials Engineering, where he is currently a Professor. He is the Industrial Research Chair in Control of Oil Sands Processes with Natural Sciences and Engineering Research Council of Canada, and the Industry Chair in Process Control with Alberta Innovates Technology Futures. He has applied his expertise extensively in industrial practice particularly in oil sands industry. His current research interests include process control, system identification, control performance assessment, Bayesian methods, and state estimation. He is a fellow of the Canadian Academy of Engineering and the Chemical Institute of Canada.

Sirish Shah

Sirish L. Shah is Emeritus Professor at the University of Alberta where he held the NSERC-Matrikon-Suncor-iCORE Senior Industrial Research Chair in Computer Process Control from 2000 to 2012.
The main area of his current research is process and performance monitoring, system identification and design, analysis and rationalization of alarm systems. He has co-authored three books, the first titled "Performance Assessment of Control Loops: Theory and Applications", a second book titled book titled "Diagnosis of Process Nonlinearities and Valve Stiction: Data Driven Approaches" and a more recent brief monograph titled, "Capturing Connectivity and Causality in Complex Industrial Processes". He is a fellow of the Canadian Academy of Engineering and the Chemical Institute of Canada.

Nina Thornhill

Nina Thornhill holds the ABB Chair of Process Automation at Imperial College London. She and her research group work at the interface between industry and academia and is interested in finding ways in which industrial operations can be controlled, operated, maintained and improved through handling of process information, data integration and information extraction. Her research interests include optimization and industrial data analysis with applications in oil and gas, chemicals, and electricity transmission.

She received a BA in Physics from Oxford University, MSc in Control Systems from Imperial College and PhD from University College London. She worked at ICI, British Aerospace and in the Department of Electronic and Electrical Engineering at University College London before moving to the Department of Chemical Engineering at Imperial College in 2007.

Jiandong Wang

Jiandong Wang is presently a Professor with the Department of Industrial Engineering and Management in the College of Engineering, Peking University, China. He received the B.E. degree in automatic control from Beijing University of Chemical Technology, Beijing, China, in 1997, and the M.Sc and Ph.D. degrees in Electrical and Computer Engineering from the University of Alberta, Canada, in 2003 and 2007, respectively. From 1997 to 2001, he was a Control Engineer with the Beijing Tsinghua Energy Simulation Company, Beijing, China. From February 2006 to August 2006, he was a Visiting Scholar at the Department of System Design Engineering at the Keio University, Japan.

His research interests include system identification, industrial alarm systems, control loop performance monitoring and optimization, and their applications to power, chemical and petrochemical industrial processes. Dr. Wang has been served as an Associate Editor for Systems and Control Letters.

Workshop time schedule:

Speaker+time	Topics to be discussed
Sirish Shah 9:00 to 10:30	• Overview of the broad analytics area with emphasis on its use in the process industry; • Basic definitions and introduction to supervised and unsupervised learning: simple regression, classification and clustering; • Data visualization methods; • Data quality assessment: Outlier detection; filtering and data segmentation.
10:30 to 10:45	Coffee break
Biao Huang 10:45 to 12:15	Elements of statistical inference and learning: • Bayesian methods. • Soft sensor analytics. Industrial case studies
12:15 to 13:15	Lunch break
Nina Thornhill 13:15 to 14:30	Multivariate methods for data analysis: SVD, PCA, PLS, SVR. Case studies on nearest neighbour methods for multivariate detection and diagnosis of transient disturbances.
Jiandong Wang 14:30 to 15:45	Alarm data analysis on detection and removal of nuisance alarms. Case studies with applications in the power generation plants.
15:45 to 16:00	Coffee break
Tongwen Chen 16:00 to 17:15	Data-based causality analysis for identification of process topology, Granger causality, transfer entropy, and case studies for root cause analysis and fault diagnosis.
17:15 to 17:30	Closing remarks and general discussion