[P5] Unified application of generative topographic mapping and graphs for monitoring of chemical processes

Matheus Escobar, Hiromasa Kaneko, Kimito Funatsu

Funatsu Laboratory, University of Tokyo, 7-3-1 Hongo, Bunkyo, 113-8654, Tokyo, Japan

Monitoring of chemical processes relies on assessing the relationship between variables. This might be difficult when many variables are available, when the process has non-linear dynamics or when the system is unsupervised, i.e., no reference data indicating normal behavior is available. This work aims for a new approach for unsupervised process monitoring, combining Generative Topographic Mapping (GTM)[1], a non-linear unsupervised tool for dimensionality reduction, and Graph Theory (GT)[2], a mathematical structure that expresses pairwise relations between objects.

GTM is a probabilistic tool that can encapsulate underlying correlations between variables in a 2D latent grid map. For each datum, there is a probability distribution associated that carries valuable information about its state. The natural conclusion that comes from it is that normal data and anomalous data will express different probability patterns in the GTM map. While that is true, these patterns can differ even among their own group, what makes clustering difficult. From this premise, GT is used to establish connections between data with similar characteristics, grouping them in networks where the density of connections within groups determine their overall similarity.

The quality of the network established, however, depends on the criterion used for similarity assessment. From image processing, an index called Structural Similarity Index (SSIM)[3] was used, taking into account standard deviation, mean and covariance of samples for assessing resemblance. Once the network is assembled, clustering can be performed and normal and anomalous data can be identified.

Process monitoring performance was compared with that of GTM alone and Principal Component Analysis (PCA). Initially, artificial data was used for analyzing how the proposed method would respond to six different anomaly scenarios. PCA and GTM could not discriminate normal and anomalous data for all cases, where the proposed method managed to identify normal and anomalous data for all of them. Another case study was also taken into account, using a Vinyl Acetate MATLAB model. Five anomaly patterns embedded in the code were analyzed, where the proposed method discriminated normal and anomalous data well. Such results motivate the use of GTM and GT together as a new approach for process monitoring.

Bibliography:

[1] C. M. Bishop, M. Svensén, and C. K. I. Williams, "GTM: The Generative Topographic Mapping," 10 (1998), pp. 215-234.
[2] F. Harary, *Graph Theory*: Perseus Books, 1994.
[3] W. Zhou, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on,* 13 (2004), pp. 600-612.