TIME-SERIES DATAMINING FOR CONTINUOUS BIOPROCESS ANALYSIS

Yang Yang, Department of Biochemical Engineering, University College London, UK
y.yang@ucl.ac.uk
Suzanne S. Farid, Department of Biochemical Engineering, University College London, UK

Key Words: continuous chromatography, time-series datamining, anomaly detection, temporal forecasting

Continuous bioprocessing technologies are attractive to biopharmaceutical manufacturers given their potential to offer cost and quality advantages. Compared to batch processes, continuous bioprocesses requires more automation and sensors and thus generate more data. A key challenge for real-time process monitoring and control is how best to combine and transform all data sources so as to create a process fingerprint for a continuous bioprocess. This work introduces a time-series datamining technique to analyze historical continuous chromatography records generated by the BioSMB™ chromatography system for pattern recognition and anomaly detection. A dynamic time warping (DTW) algorithm combined with a K-means clustering method was applied to identify the motif patterns of various sensors so as to link the patterns with different process settings and establish process fingerprints. Case studies will be presented demonstrating how these advanced dynamic multivariate data analysis techniques can be used to rapidly detect anomaly patterns in continuous chromatography runs as well as their root causes. This work demonstrates the feasibility of real-time monitoring of continuous bioprocesses using time-series data mining methods.