USING AUTOMATIC CLASSIFICATION FOR TECHNOLOGICAL PROCESSES HANDLING

Vitalii Vostotskyi, *IN-43,*
G.L. Litvinenko, *E.L. adviser*

The process of production management represents one of the major problems of mankind. The central problem is the construction of the effective model displaying the structure and properties of information streams and giving enough information for forecasting and management

Traditionally, models have been developed by mathematically formulating phenomenological behavior after the phenomena are adequately understood. If these formulations are amenable to analytical or numerical solution, the results enable robust explanation and prediction of system behavior within the bounds of the models' regions of validity. Data organization can be envisioned as a clustering process that identifies neighbors by their proximity to one another. Both handling large amounts of data and dealing effectively with streams of incoming fresh data require robust self- organizing systems capable of continuous updating themselves in a dynamically adaptive manner. Let us briefly describe the approach for data auto-classifying . The process is initiated by arbitrarily establishing clusters, and then assigning individual data to particular clusters based on the proximity of the data to the clusters' centroids. Successive iterations of the data-assigning process lead to the convergence of a set of clusters capturing similar data according to a pre-specified selectivity criterion. A narrow tolerance limit on the selectivity criterion results in a large number of clusters, each containing a relatively small number of very similar data; a broader tolerance yields fewer, more inclusive clusters with greater internal diversity. Cluster centroids are continually updated to adapt to their member data groups as fresh data are encountered. New clusters are created to accommodate those incoming fresh data which are too far (according to the selectivity criterion) from any existing cluster. Architectural details to construct such self-organizing systems deals with approaches, based on the source data characteristics and technological process complexity. Organizing data in this manner enables us to identify member groups and quantify their similarities, and to predict properties of fresh incoming data by their proximity to historical data whose properties are known.