

Qualitative Trend Clustering of Process Data for Fault Diagnosis *

Zhou Bo and Ye Hao

Abstract—In this paper, a qualitative trend clustering (QTC) algorithm is developed for fault diagnosis of industrial process. The clustering procedure mainly involves three steps. Firstly, the process data are segmented into consecutive episodes to which qualitative primitives will be assigned. Secondly, Smith Waterman algorithm for local sequence alignment in bioinformatics is utilized to derive the similarity measure of any two qualitative representations. Thirdly, pattern clustering analysis of process data is carried out to discriminate normal and abnormal conditions. The advantages of the proposed method include requiring less prior knowledge, being more robust to process noise and variation of signal characteristics. The application of QTC in industrial processes is illustrated on a real blowing down and recovery process of blast furnace iron making process, which shows its potentials in fault diagnosis and process monitoring tasks.

I. INTRODUCTION

The collected historical operational data remain a useful resource for process monitoring. Through data mining of process history data, more insights into process faults and process behaviors and can be derived. Such kinds of process information will greatly fascinate process monitoring and improve the performance of fault detection and diagnosis of plants to a large extent [1].

Clustering analysis is a kind of unsupervised classification, which classes observations into groups with the aid of similarity measure between data patterns [2, 3]. It has been widely adopted to analyze the underlying data structure of process data, and further perform process state monitoring and fault detection. Some related literatures are reviewed in the following. The authors in [4] utilize the fuzzy c-means clustering and principal component analysis (PCA) to study the operational data collected from a refinery fluid catalytic cracking process, revealing that different zones in reduced data space correspond to different product grades. In order to handle temporal signals in a more effective manner, a dynamic PCA-based similarity measure is proposed to compare different process modes and transitions, whose effectiveness is proven by the superior clustering performance on the multivariate data sets of agile chemical plants [5]. Similarly, the authors in [6] present a nonhierarchical framework consisting of a so called k-PCA models algorithm and a moving window technique to distinguish different process states including steady states, transitions and faults. Apart from continuous process, applications of clustering analysis are also found in batch process. An example is the work of [7]

where principal scores and Q-statistics of multi-way PCA models built from the historical process data of a sequencing batch reactor (SBR) processes are fed to the learning algorithm for multivariable data analysis (LAMDA) for further clustering. This procedure produces an efficient discrimination between normal and abnormal conditions as well as a clear differentiation of the operational changes. Clustering analysis has been served as an alternative tool for fault diagnosis of chemical processes. The works of [8] and [9] are the representatives which identify locations and causes of faults based on pattern recognition of features of process data.

However, most of the abovementioned works try to extract quantitative features (such as principal scores, Q-statistics and the direction of the principal component, etc.) from original data sets. In spite of carrying detailed information about the process being investigated, these kind of quantitative features are usually more vulnerable to disturbances and run-to-run differences of the same process. Conversely, as highlighted by [1], qualitative trend proves to be a succinct qualitative descriptor that can explain important events in the process from a higher level perspective. So the aim of this paper is to study the qualitative trend clustering (QTC) of process data for the purpose of diagnosing faults in a process.

The QTC based fault diagnosis method proposed in this paper belongs to a process data based fault diagnosis method, i.e. qualitative trend analysis (QTA), but QTC performs the fault diagnosis in a different way from most of the applications of QTA. Existing results on the application of QTA to fault diagnosis mainly fall into two categories: (1) identifying faults through matching the currently observed trends with the trend patterns stored in previously constructed fault knowledge base [10-12]; (2) inferring the nature of faults based on the interpretation of extracted trends directly with the aid of expertise knowledge or specific rules about the process [13-15]. In such QTA based fault diagnosis methods, the prior knowledge such as fault database and expertise knowledge is indispensable, but usually difficult to gather in practice. However, in QTC, fault diagnosis is performed through clustering samples based on qualitative trend features, which doesn't need that priori knowledge. So far, only a few works related to QTC have been reported [16, 17]. In this paper, the qualitative trends are firstly extracted from time series by interval-halving framework in [18]. Then a well-known local sequence alignment algorithm in bioinformatics [19] is adopted to derive a similarity measure between two primitive sequences. At last, a final hierarchical cluster solution can be given and provides meaningful classification of various faults. The main advantages of this work can be summarized as (1) QTC based fault diagnosis is performed through unsupervised classification of process data, which needs less prior knowledge compared with other QTA based methods; (2)

*Research supported by National Nature Science Foundation under Grant 61290324.

Zhou Bo and Ye Hao are with Department of Automation, Tsinghua University, Beijing, China. Ye Hao is the corresponding author with phone: +86-01062790497. (e-mail: zhou-b11@mails.tsinghua.edu.cn, haoye@tsinghua.edu.cn)

compared with quantitative feature based clustering, QTC utilizes the qualitative trend feature which is more robust to process noise and variation of signal characteristics; (3) Adopting the local sequence alignment algorithm [19] in the trend similarity measure step makes the similarity estimation more suitable for the case in which the lengths of signals differ largely or the ends of trends are not accurately identified.

The rests of this paper are organized as follows. Section 2 briefly introduces the basic concepts of QTA. Section 3 deals with QTC by utilizing the Smith Waterman algorithm. In Section 4, an application example for classification of data of the blowing down and recovery process in a real blast furnace is given to illustrate the clustering procedure, which shows the effectiveness of proposed scheme. Section 5 finishes with conclusions.

II. A BRIEF INTRODUCTION TO QTA

A. Trend Definition and Extraction

A qualitative trend representation is a consecutive sequence of episodes during which an alphabetic symbol called primitive is used to characterize the temporal behaviors [20] [21]. The authors in [20] build a formal framework capable of transforming time series of process variables into successive trend segments on the basis of triangular episode. This temporal episode is characterized by a graphic triangulation where the signal's first derivative and second derivative remain constant. As shown in Fig. 1 [21], the primitives in episodes are encoded into seven letters from A to G. To meet the demands of various practical applications, many other researchers have expanded the sets of primitives to incorporate other information, such as the magnitude and duration of episode into the representation. Examples include [22] and [15], to name a few.

On the other hand, many works have been devoted to the development of trend extraction algorithms. Owing to its easy implementation and robustness to noise, the polynomial fit based method [11] has gained its popularity in many previous studies. Since trend extraction is not the main issue in this paper, we recommend interested readers to refer to the works of [23-25] for trend extraction algorithms. In this paper, the interval halving framework proposed by Dash et al. [18] is employed to identify process trends automatically due to its little requirement of prior knowledge and good performance.

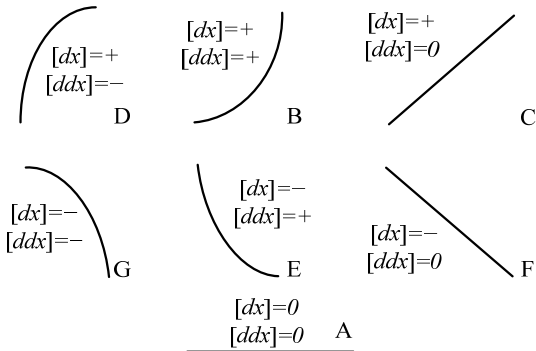


Figure 1. Seven primitives

B. Trend Similarity Measure

Trend similarity measure is also an important part of QTA which defines the closeness of two trends, which can be utilized to compare qualitative representations of sensor data. The authors in [10] and [11] propose similarity indices based on fuzzy trend matching. The similarity between different primitives is quantified by a similarity matrix in a fuzzy manner. Then the similarity measure is built on the sum of the similarities of the two primitive sequences. In [26], the dissimilarity measure based on matching degrees of shape, duration and magnitude between the dictionary and real-time trends is defined. Dynamic time warping (DTW) is another similarity measure that calculates the minimal distance between two sequences by aligning them in the time axis [27, 28]. Qualitative sequence similarity index (QSSI) recently proposed in [29] is a new similarity criterion based on Needleman Wunsch algorithm.

It is worth to mention that the above similarity measures are defined on the basis of global alignment. That is, they take all the elements of the sequences into consideration in the definition. However, in some real-world industrial situations, the lengths of the primitive sequences of two trends may differ to a large degree, which brings challenges in the implementation of the above algorithms. There also exist situations where the starts and ends of two trends are not accurately identified or only subparts of the trends are similar, then a global end to end alignment of two sequences may not deliver satisfactory results. These are the main motivations of this paper for the utilization of Smith Waterman algorithm for local sequence alignment in defining the similarity measure of two qualitative trend representations.

III. QTC WITH SMITH WATERMAN ALGORITHM

A. Smith Waterman Algorithm

In this section, Smith Waterman Algorithm used in bioinformatics for identifying locally similar subsequences will be briefly introduced first. Given two character sequences $PS_1 = \{P_1, P_2, \dots, P_M\}$ and $PS_2 = \{\hat{P}_1, \hat{P}_2, \dots, \hat{P}_N\}$, Smith Waterman algorithm [19] tries to find a pair of subsegments with the maximal similarity from the two sequences. Define a scoring matrix $H \in R^{(M+1) \times (N+1)}$, where $H_{i,j}$ denotes the maximal similarity score of two subsequences of PS_1 and PS_2 ending in P_i and \hat{P}_j , respectively. Firstly, matrix H is initialized with [19]

$$\begin{aligned} H_{i,0} &= 0, 0 \leq i \leq M \\ H_{0,j} &= 0, 0 \leq j \leq N. \end{aligned} \quad (1)$$

The rest elements in H can be calculated through dynamic programming with the following rule [19]

$$H_{i,j} = \max \left\{ H_{i-1,j-1} + s(P_i, \hat{P}_j), H_{i-k,j} + W_k, H_{i,j-l} + W_l, 0 \right\} \quad (2)$$

where $s(P_i, \hat{P}_j)$ is the local similarity score given between the i -th element in PS_1 and j -th element in PS_2 , W_i (i.e.,

W_k and W_l in (2)) denotes the gap penalty weight given to insertions/deletions in the gap.

When the scoring matrix H is built, the next procedure is locating the pair of subsegments of maximal similarity with the following steps. Firstly, find the highest score in H . Assume its position is (i, j) . Then go back to the corresponding position leading to this highest score. The direction of this movement produces the associate alignment, which can be described as follows [19].

- (1). If $(i-1, j-1)$ leads to (i, j) , then P_i and \hat{P}_j are aligned.
- (2). If $(i-1, j)$ leads to (i, j) , then there is an insertion and P_i is aligned with a gap.
- (3). If $(i, j-1)$ leads to (i, j) , then there is a deletion and \hat{P}_j is aligned with a gap.

Such a traceback procedure is sequentially repeated until an element of H with zero value is reached, identifying the most similar regions between the two sequences as well as the corresponding alignments.

B. Trend Similarity with Smith Waterman Algorithm

In this subsection, the Smith Waterman algorithm [19] is utilized to estimate the similarity between two trend representations. Given two signals s_1, s_2 as well as the corresponding primitive sequences $PS_1 = \{P_1, P_2, \dots, P_M\}$ with primitive P_i spans over $[t_i, t_{i+1}]$ and $PS_2 = \{\hat{P}_1, \hat{P}_2, \dots, \hat{P}_N\}$ with primitive \hat{P}_i over $[\hat{t}_i, \hat{t}_{i+1}]$. Suppose that the pair of segments identified by the Smith Waterman algorithm are $\{P_{M_1}, P_{M_1+1}, \dots, P_{M_2}\}$ and $\{\hat{P}_{N_1}, \hat{P}_{N_1+1}, \dots, \hat{P}_{N_2}\}$ respectively. It can be concluded that these two segments are the most similar subsequences from the original two sequences. That is, primitive sequence $\{P_{M_1}, P_{M_1+1}, \dots, P_{M_2}\}$ is qualitatively similar to $\{\hat{P}_{N_1}, \hat{P}_{N_1+1}, \dots, \hat{P}_{N_2}\}$. Based on these facts, the similarity measure between the two different qualitative representations can be defined as

$$S(PS_1, PS_2) = \min \left\{ \frac{(t_{M_2+1} - t_{M_1})}{(t_{M+1} - t_1)}, \frac{(\hat{t}_{N_2+1} - \hat{t}_{N_1})}{(\hat{t}_{N+1} - \hat{t}_1)} \right\} \quad (3)$$

where $(t_{M_2+1} - t_{M_1}) / (t_{M+1} - t_1)$ denotes the ratio of duration of aligned primitive sequence $\{P_{M_1}, P_{M_1+1}, \dots, P_{M_2}\}$ to that of the comparing primitive sequence $\{P_1, P_2, \dots, P_M\}$, so is $(\hat{t}_{N_2+1} - \hat{t}_{N_1}) / (\hat{t}_{N+1} - \hat{t}_1)$. It can be seen that $S(PS_1, PS_2)$ is between 0 and 1, where 0 means complete dissimilarity and 1 means complete similarity. A graphic illustration for this similarity measure of qualitative representations is given in Fig. 2.

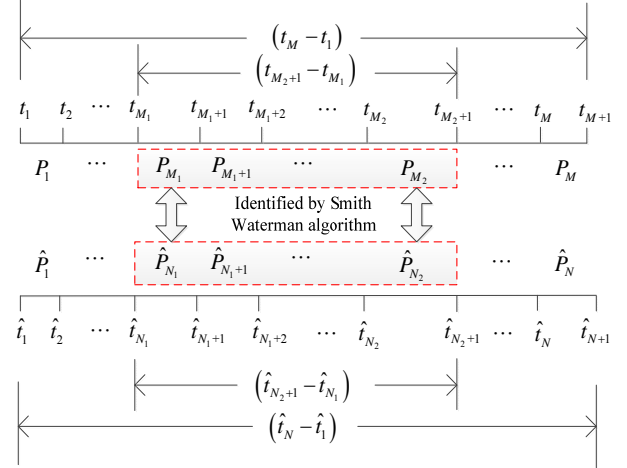


Figure 2. Similarity measure based on the local alignment with Smith Waterman algorithm

C. Hierarchical Clustering with Qualitative Trends

In hierarchical clustering, samples are merged to form groups based on the distance between the samples and the groups continue to merge based on the chosen linkage criterion until they end in one category including all the observations [2, 3]. The distance between two signals can be determined from the similarity measure of the corresponding qualitative representations as

$$d(s_1, s_2) = 1 - S(PS_1, PS_2) \quad (4)$$

Distance metric has a significant impact on the final hierarchy structure. The patterns discovered by this qualitative trend clustering will be considered to exhibit similar qualitative evolution in the time axis. In summary, the procedures of QTC can be described as:

- (1) Use the primitive language listed in Fig. 1 [21] to describe the seven types of the trends.
- (2) Segment the signals and identify their qualitative primitive sequences by utilizing the trend extraction methods in interval halving framework of Dash et al. (2004) (Please refer to Sec. 2.1).
- (3) Compute the similarity/distance between each pair of the qualitative representations of the signals using the Smith Waterman algorithm (Please refer to Sec. 3.2).
- (4) Carry out cluster analysis of the process data with agglomerative hierarchical clustering algorithm [2, 3].

IV. ILLUSTRATING EXAMPLE

A. Blast Furnace and Collected Data

In this paper, an application of QTC to the historical data of a real blast furnace will be given to verify the proposed method. Blast furnace is a high temperature counter current reactor where iron ore, coke and flux are charged into the top and the preheated blast air carrying the pulverized coal is blown into the bottom of furnace through tuyeres. The ferrous and ferric oxide which is in the form of sinter and pellets is then reduced to liquid pig iron by the reducing gas of CO and H₂. The melting liquid including iron and slag is periodically removed from furnace hearth through taphole and exhaust gases exit from the top section of furnace [30].

The data invested here belong to some blowing down and recovery processes of a blast furnace (with volume of 2000 m³) of Guangxi Liuzhou Iron and Steel Company located in Guangxi, China. The duration of each sample is listed in Table 1. Here, one sample means a complete blowing down and recovery process in the iron-making process. There are 29 variables monitored by the distributed control system at a sample time of 10 seconds which are summarized in Table 2. The blowing down process may be carried out due to several reasons, such as equipment overhaul, equipment failure, severe abnormal conditions happening in the process, etc. However, these samples are stored in the historical database without any records describing what happened, except sample No. 29, which is caused by an accident according to the accident report. Here we try to use the proposed QTC method to classify them in an unsupervised way, so that the different patterns can be revealed and the abnormal operations and normal operations can be discriminated.

TABLE I. THE DURATION OF EACH SAMPLE

1	1:49:10	11	3:27:51	21	2:57:35	31	3:58:05
2	28:10:05	12	18:30:58	22	1:12:23	32	6:30:51
3	1:54:04	13	2:29:10	23	3:21:04	33	2:09:52
4	0:58:50	14	0:52:37	24	3:31:08	34	2:14:09
5	4:21:58	15	0:28:30	25	20:36:28	35	4:44:16
6	2:35:46	16	3:04:22	26	2:28:30	36	4:36:56
7	2:05:53	17	29:41:21	27	4:51:37	37	3:44:15
8	2:01:00	18	6:38:39	28	2:55:10	38	2:44:33
9	2:16:14	19	3:39:25	29	25:24:02	39	2:17:53
10	2:32:34	20	2:07:25	30	2:39:17		

TABLE II. THE VARIABLES STORED IN DATABASE

1	rate of enriched oxygen	11	bosh gas index	21	temperature of hot blast
2	permeability index	12	theoretical flame temperature	22	top temperature 1
3	CO	13	top pressure 1	23	top temperature 2
4	H ₂	14	top pressure 2	24	top temperature 3
5	CO ₂	15	top pressure 3	25	top temperature 4
6	standard air speed	16	pressure of enriched oxygen	26	resistance index
7	flow rate of oxygen enrichment	17	pressure of cold blast	27	blast humidity
8	flow rate of cold blast	18	total press drop	28	set value of coal injection
9	kinetic energy of blast	19	pressure of hot blast	29	actual value of coal injection
10	bosh gas volume	20	actual air speed		

TABLE III. THE SIMILARITY SCORE MATRIX OF SEVEN PRIMITIVES

primitive	A	B	C	D	E	F	G
A	1	-0.5	0	-0.5	-0.5	0	-0.5
B	-0.5	1	0.5	0.25	-1	-1	-1
C	0	0.5	1	0.5	-1	-1	-1
D	-0.5	0.25	0.5	1	-1	-1	-1
E	-0.5	-1	-1	-1	1	0.5	0.25
F	0	-1	-1	-1	0.5	1	0.5
G	-0.5	-1	-1	-1	0.25	0.5	1

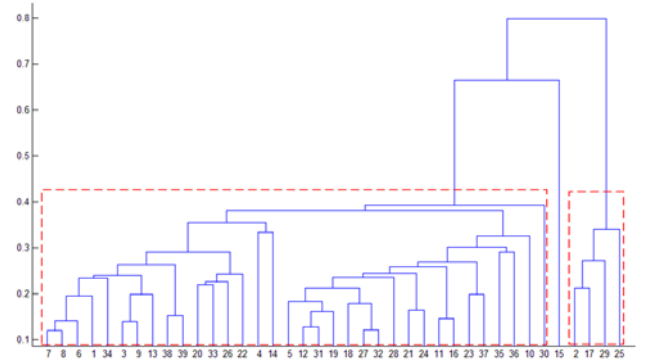


Figure 3. The hierarchical cluster tree of the data of the blowing down and recovery process of blast furnace

B. Clustering Results and Discussion

In this example, distance between each pair of samples is defined as the average of the corresponding distance between the 29 variables. The unweighted average distance is chosen as the linkage criterion in constructing the agglomerative hierarchical cluster tree. The local similarity score $s(P_i, \hat{P}_j)$ is given in Table 3 and the gap penalty weight is chosen as $W_i = -0.5i$. The final results are depicted in Fig. 3.

It can be seen from Fig. 3 that the group of 39 samples can be roughly classified into 3 categories. A further investigation into each category shows that: (1) the first category consisting of 34 samples corresponds to the standard blowing down and recovery process, i.e., the process behavior is normal and no accidents have been reported during the operation; (2) the second category has only one sample (i.e., No. 15). This sample is special because its time duration is the shortest in the group (about half an hour). It belongs to the slow-wind blowing operation and the pressure of hot blast has not been reduced to zero which is different from the rest samples; (3) the third category consists of 4 samples (i.e., No.2, No.17, No.29 and No.25). The sample No.29 turns out to be an accident caused by the cool of the blast furnace, which has been recorded and analyzed by the operator before our work. However, the other three samples also turn out to be in abnormal condition. There exist obvious signs of unstable and unsmooth furnace condition during the recovery process and the variables of the processes show a clear deviation from normal operation.

The above analyses show that the clustering results can be used in the diagnosis of faults and abnormal states of blast furnace iron making process. The methods can be also applied to the data mining of the historical data sets, helping

to build the fault data/knowledge base which is of crucial importance to fault detection and diagnosis tasks. Furthermore, with more interpretations of clustering results, more deep knowledge and understanding about control strategy and fault diagnosis of blast furnace iron making process will be gained.

V. CONCLUSION

Qualitative trend analysis is an efficient tool to extract the qualitative characteristics from signals. In this paper, hierarchical clustering of process data is developed based on trend representations of variables. In order to evaluate the similarity between the primitive sequences more accurately, the Smith Waterman algorithm, a local sequence alignment algorithm in bioinformatics, is adopted. The clustering results can be used for the purpose of fault diagnosis directly and they can be applied in the construction of knowledge-base. Further research will focus on defining more rational similarity measure in a multivariable and multi-scale environment.

REFERENCES

- [1] V. Venkatasubramanian, R. Rengaswamy, S. Kavuri and K. Yin, "A review of process fault detection and diagnosis Part III: Process history based methods", *Computers & Chemical Engineering*, 27(3), pp. 327-346, 2003.
- [2] A.K. Jain, M.N. Murty and P.J. Flynn, "Data clustering: a review", *ACM Computing Surveys*, 31(3), pp. 264-323, 1999.
- [3] T. Fu, "A review on time series data mining", *Engineering Applications of Artificial Intelligence*, 24(1), pp. 164-181, 2011.
- [4] Y.M. Sebzalli and X.Z. Wang, "Knowledge discovery from process operational data using PCA and fuzzy clustering", *Engineering Applications of Artificial Intelligence*, 14(5), pp. 607-616, 2001.
- [5] R. Srinivasan, C. Wang, W.K. Ho and K.W. Lim, "Dynamic principal component analysis based methodology for clustering process states in agile chemical plants", *Industrial & engineering chemistry research*, 43(9), pp. 2123-2139, 2004.
- [6] S. Beaver, A. Palazoglu and J.A. Romagnoli, "Cluster analysis for autocorrelated and cyclic chemical process data", *Industrial & Engineering Chemistry Research*, 46(11), pp. 3610-3622, 2007.
- [7] K. Villez, M. Ruiz, G. Sin, J. Colomer, C. Rosen and P.A. Vanrolleghem, "Combining multiway principal component analysis (MPCA) and clustering for efficient data mining of historical data sets of SBR processes", *Water Science & Technology*, 57(10), pp. 1659-1666, 2008.
- [8] F. Akbaryan and P.R. Bishnoi, "Fault diagnosis of multivariate systems using pattern recognition and multisensor data analysis technique", *Computers & Chemical Engineering*, 25(9), pp. 1313-1339, 2001.
- [9] L.H. Chiang, M.E. Kotanchek and A.K. Kordon, "Fault diagnosis based on Fisher discriminant analysis and support vector machines", *Computers & Chemical Engineering*, 28(8), pp. 1389-1401, 2004.
- [10] S. Dash, R. Rengaswamy and V. Venkatasubramanian, "Fuzzy-logic based trend classification for fault diagnosis of chemical processes", *Computers & Chemical Engineering*, 27(3), pp. 347-362, 2003.
- [11] M.R. Maurya, R. Rengaswamy and V. Venkatasubramanian, "Fault diagnosis using dynamic trend analysis: A review and recent developments", *Engineering Applications of Artificial Intelligence*, 20(2), pp. 133-146, 2007.
- [12] F.I. Gamero, J. Meléndez and J. Colomer, "Process diagnosis based on qualitative trend similarities using a sequence matching algorithm", *Journal of Process Control*, 24(9), pp. 1412-1424, 2014.
- [13] F.I. Gamero, J. Colomer, J. Meléndez and P. Warren, "Predicting aerodynamic instabilities in a blast furnace", *Engineering Applications of Artificial Intelligence*, 19(1), pp. 103-111, 2006.
- [14] K. Villez, C. Rosén, F. Anctil and P.A. Vanrolleghem, "Qualitative Representation of Trends (QRT): Extended method for identification of consecutive inflection points", *Computers & Chemical Engineering*, 48, pp. 187-199, 2013.
- [15] K. Villez, V. Venkatasubramanian and R. Rengaswamy, "Generalized shape constrained spline fitting for qualitative analysis of trends", *Computers & Chemical Engineering*, 58, pp. 116-134, 2013.
- [16] El.S. Ferchichi, S. Zidi, K. Laabidi, M. Ksouri and S. Maouche, "Meanshift Clustering based trend analysis distance for fault diagnosis", 2013 IEEE International Conference on Systems, Man, and Cybernetics, pp. 1877-1882, 2013.
- [17] B. Balaskó, S. Nemeth and J. Abonyi, "Time-series similarity application to qualitative process trend analysis", *Acta Agraria Kaposváriensis*, 2(11), pp. 197-209, 2007.
- [18] S. Dash, M.R. Maurya and V. Venkatasubramanian, "A novel interval-halving framework for automated identification of process trends", *AIChE journal*, 50(1), pp. 149-162, 2004.
- [19] T.F. Smith and M.S. Waterman, "Identification of common molecular subsequences", *Journal of Molecular Biology*, 147(1), pp. 195-197, 1981.
- [20] J. Cheung and G. Stephanopoulos, "Representation of process trends - Part I. A formal representation framework", *Computers & Chemical Engineering*, 14(4), pp. 495-510, 1990.
- [21] Janusz, M.E. and V. Venkatasubramanian, "Automatic generation of qualitative descriptions of process trends for fault detection and diagnosis", *Engineering Applications of Artificial Intelligence*, 4(5), pp. 329-339, 1991.
- [22] J.C. Wong, K.A. McDonald and A. Palazoglu, "Classification of process trends based on fuzzified symbolic representation and hidden markov models", *Journal of Process Control*, 8(5), pp. 395-408, 1998.
- [23] E. Keogh, S. Chu and M. Pazzani, "An online algorithm for segmenting time series", *Proceedings IEEE International Conference on Data Mining*, pp. 289-296, 2001.
- [24] S. Charbonnier, C. Garcia-Beltan, C. Cadet and S. Gentil, "Trends extraction and analysis for complex system monitoring and decision support", *Engineering Applications of Artificial Intelligence*, 18(1), pp. 21-36, 2005.
- [25] S. Charbonnier and F. Portet, "A self-tuning adaptive trend extraction method for process monitoring and diagnosis", *Journal of Process Control*, 22(6), pp. 1127-1138, 2012.
- [26] A. Sundarraman and R. Srinivasan, "Monitoring transitions in chemical plants using enhanced trend analysis", *Computers & Chemical Engineering*, 27(10), pp. 1455-1472, 2003.
- [27] J. Colomer, J. Melendez and F.I. Gamero, "Qualitative representation of process trends for situation assessment based on cases". 15th IFAC World Congress, 2002.
- [28] B. Balasko, Z. Banko and J. Abonyi, "Analyzing trends by symbolic episode representation and sequence alignment", *Proceedings of the 15th Mediterranean Conference on Control & Automation*, 2007.
- [29] F.I. Gamero, J. Melendez and J. Colomer, "Qssi: a new similarity index for qualitative time series Application to classify voltage sags", *Applied Artificial Intelligence*, 25(2), pp. 141-162, 2011.
- [30] M. Geerdes, H. Toxopeus, C. Van Der Vliet, *Modern blast furnace ironmaking: an introduction*, IOS Press, 2009.