THE ANNALS OF "DUNĂREA DE JOS" UNIVERSITY OF GALAȚI
FASCICLE V, TECHOLOLOGIES IN MACHINE BUILDING, ISSN 1221-4566
2007

Algorithm for Economical Characteristics Identification a Machining System

PhD. Eng. Al. Epureanu, Drd. Eng. F. Buruiana, Drd. Eng. S. Ciuntu, Drd. Eng. F. Susac, Drd. Eng. A. Tofan

University of Galati, DITDP – 111, Domneasca Street, 800201 Galati, Romania

ABSTRACT

In the last years, the market evolution is characterized more and more by many changes such as increasing of the new products frequency that are often personalized, large fluctuations of the structure and the commands volume and few fundamental changes in technology. The economical characteristics refer to the relationship between manufacturing and the economical indicators level of the manufacturing system. Technical-economical characteristics identification of a manufacturing system requests developing of an identification algorithm, using knowledge techniques (Knowledge Discovery from Databases and Artificial Neural Network) that should take into account all the current changes of the market. The paper proposes an algorithm for identification of the economical characteristics of a drilling machining system based on the Knowledge Discovery from Databases.

***Key words**: machining system, economical characteristics, unsupervised learning, adaptive control*

1. Introduction

In the last years, the market became a dynamic environment expressed by the continuing changing of the rate between demands and tender that has a direct consequence in big fluctuation of the production. The dynamism of the market is influenced mainly due to increased frequency of new products developing most of these are personalized, large fluctuations of the structure and products commands volume and few fundamental changes in technology. These changes have created an aggressive competition at a global scale and the need to establish a new equilibrium between economy, technology and society became an important demand. For surviving in this new technological environment, the companies need to react to these changes as much as fast and efficient possible from a technical-economical point of view.

As a result, the competition became the most synthetic action of the economical strategies, since it reflects the essential aspects of the importance of the products on the market and completely characterizes the viability of a company. The other economical characteristics cannot be studied separately, being necessary an integrated model of technical- economical characteristics and its parameters identification procedure needs to be set up.

The economical characteristics are destined to define the relationship between manufacturing and the economical indicators level of the manufacturing system so that coupling of the technical and economical characteristics it is the best way to catch all the aspects of the market.

To fit the parameters of one technical-economic model is a process of identification that respects an algorithm designed to use knowledge techniques such Knowledge Discovery from Databases (KDD) and Artificial Neural Network (ANN).

KDD techniques can be updated online so that the technical-economical model parameters can be permanently up-to-date. Knowledge Discovery from Databases is an intelligent tool of mining the knowledge from the databases. It has been applied in the manufacturing for discovering useful and

understandable pattern of the manufacturing data.

An artificial neural network (ANN), often just called a "neural network" (NN), is a mathematical model or computational model based on biological neural networks functioning. It consists of an interconnected group of artificial neurons and processes information using a connexion approach for finding the statistical relations between the input and the output. In most cases, an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase.

In the paper it is proposed an algorithm for economical characteristics identification and it is presented its application for a drilling machining process. The KDD is applied for determines the rules of the drilling machining process that are used in the technical-economical model as input data. KDD consists in identification of the clusters of the process parameters that are connected to the other clusters of the market environment.

2.Technical-economical identification of a system

For carrying-out technical-economical identification of a machining system, an identification algorithm is developed so that its deployment depends on the dynamics, static and the competitiveness characteristics of the system.

For example we use the database of a of a drilling machining holon that has the following parameters: type of the material, hole diameter, drilling speed, number of holes, drilling time, energy consumption, operation cost and waste quantity.

Table 1 – Values of the input parameters of a drilling holon

Input parameters of drilling holon	**Value**
Type of material	9
Hole diameter (mm)	1
Number of holes	2
Drilling speed (mm/s)	3
Number of pieces	4
Drilling time (s)	5
Energy consumption (KW/hour)	6
Operation cost (lei)	7
Waste quantity (Kg)	8

Changing the values of this parameters in time means that these items should be called variables.

Identification algorithm uses as input data a set of monitorized variables, between which it exists an implicit relationship.

2.1. Steps for developing the algorithm

Step 1. Clustering of variables

Clustering is one of the most useful tasks in data mining for discovering groups and identifying interesting distributions and patterns within a set of data. Clustering problem means partitioning into groups of a given data set so that the data points in a cluster are more similar to each other than points in different clusters.

The main concern in the clustering process is to reveal the organization of patterns into groups, which allow to discover the similarities and differences, as well as to derive useful conclusions about them.

For the used example it is consider as independent variables, variables from 3 and 5 columns and the rest of the variables are interdependent..

Between the interdependent variables we can develop the next clusters: 1-4-6, 4-1-2-8, 7-1-2-4-9, 8-2-6.

Step 2 Segmentation of the variables range

Segmentation of the variables range means division of the range in 3 parts. Segmentation can be developing by making the difference between the highest and smallest value of the variation range.

After the segmentation of the first range we have a number of 15 clusters, and we keep only the important ones.

Step 3 Structuring the identification space

Structuring the identification space is made through the identification of the most important segment of variable (1-2-8) and then sharing the variables.

Step 4. Finding a mathematical model for each identification range

For finding a mathematical model for each identification range, we can see that the relationship between the variables is an implicit one, so the mathematical model will be an implicit function.

The family of the mathematical model will be: $f(x,y,z)=0$, that is a polynomial function, with the highest degree 2.

For generating the mathematical model we use the Taylor series for a point in the middle of the domain.

The decomposition of the polynomial in the Taylor series is presented in relation (1) as follows:

$$f(x,y,z)= f(x_0,y_0,z_0)+\left(\frac{x-x_0}{1}\frac{\partial F}{\partial x}\right)_{\substack{x=x_0\\y=y_0\\z=z_0}} +$$

$$+\left(\frac{y-y_0}{1}\frac{\partial F}{\partial y}\right)_{\substack{x=x_0\\y=y_0\\z=z_0}} + \left(\frac{z-z_0}{1}\frac{\partial F}{\partial z}\right)_{\substack{x=x_0\\y=y_0\\z=z_0}} \quad (1)$$

This relation represents the evolution of the function with respect to the variables x, y and z that model the rapid changes of the output according to the input fluctuation. In the relations (2) and (3) it is presented the expression of the f function (2) and for the lack of simplicity, to have a problem with solutions (3 unknown with 3 equations), D=1 (3).

$$f(x,y,z)=Ax+By+Cz+D=0 \quad (2)$$

$$f(x,y,z)= Ax+By+Cz+1=0 \quad (3)$$

If the output is the global value of f(x,y,z) that represents in our model the market characteristics, the identification of the constants of the polynomial, i.e. x, y, z, represents the best solution of one set of parameters that could be applied to a drilling machine holon control.

3. Generation of the data bases

To extract rules and to use knowledge of some previous experiences, a data base was created using the sets of drilling control parameters and the error of the holes obtained after the manufacturing. The data are from the database of the Research Centre Technological Engineering, Galati and represents cases of drilled parts between the years 1983-2005. The number of the selected rows of the data base was 40. To define the correlation between the input parameters and the output parameters of the studied case, a neural network software NNMODEL was in order to find the best model between the input and output parameters. The neural network is based on the principles of the back propagation algorithm (Rumelhart et al., 1986). The ANN model constructs a model based on examples of data with known outputs. A backpropagation network typically comprises three types of neuron layers: an input layer, one or more hidden layers and an output layer each including one or several neurons. As shown in Figure 1, nodes from one layer are connected to all nodes in the following layer, but no lateral connections within any layer, nor feed-back connections are possible. Fifteen input neuron are used, each representing an environmental variable. The output layer comprises one neuron, indicating the presence or absence of a macroinvertebrate taxon.

With the exception of the input neurons, which only connect one input value with its associated weight values, the net input for each neuron is the sum of all input values xn, each multiplied by its weight wjn, and a bias term zj which may be considered as the weight from a supplementary input equalling one.

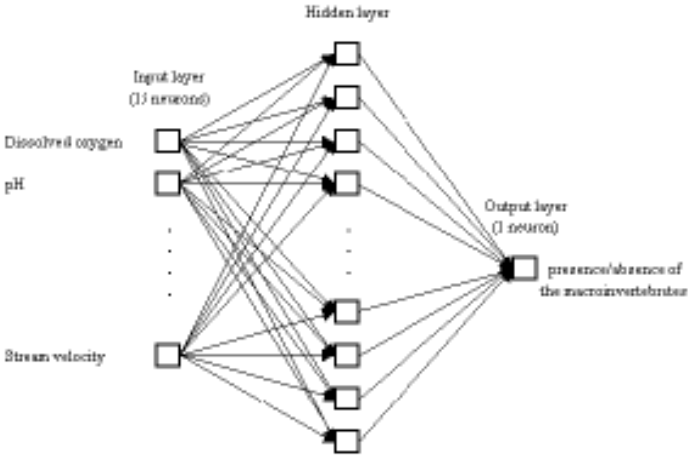

Figure 1. Illustration of a three-layered neural network with one input layer, one hidden layer and one output layer.

Before training, the values of the weights and biases are initially set to small random numbers. Subsequently, a set of input/output vector pairs is presented to the network. For each input vector, the output vector is calculated by the neural network model, and an error term is calculated for the outputs of all hidden and output neurons, by comparing the calculated output vector and the actual output vector (the "target"). Using this error term, the weights and biases are updated in order to decrease the error, so future outputs are more likely to be correct. This procedure is repeated until the errors become small enough or a predefined maximum number of iterations is reached. This iterative process is termed "training". After the training, the ANN can be validated using independent data.

Application of the neural network consists in creating the data base, creating the neural network model, that establish which variables are set as input variables and which one are output variables. In the present case, all the control parameters (cluster 1-2-8) are input parameters and the output is the error of the diameter of the drilled part.

Initiation and training are the next steps of getting a dynamic model that

correlate the input with output variables. The best model that this trained network was provided is presented in (4):

$$Q= \sum F(x_i,y_i,z_i)^2 =\sum(error_i)^2 \ , \qquad (4)$$

where Q is :

$$Q= \sum(A\ x_i+By_i+Cz_i+1)^{2.} \ , \qquad (5)$$

Identification of x, y and z variables that gives the best solution in order to have the equality between the two terms of the equation 4 is an optimisation problem and offer the solution between input and output.

The output parameters selected in the neural network data base, the dimensional error of the part, is a global indicator of the quality of the production. If the errors are exceeding to the tolerances range, the parts are rejected. The output could be a combination of the quality and cost functions, even the rate of the currency at the day of the manufacturing. The output should spot all the aspects of the market that is a dynamic environment.

4. Identification procedure and results

Generally, an optimisation problem has established the cost function that means the target of the optimisation and a set of restriction that the solution should respect. In the present paper, the cost function is to minimize the difference between the sum of the power two of the function F and power two of the sum of the errors imposed by correspondence with the same value of F.

For finding the solution of x, y, z, the derivatives of Q function are calculated with respect to the A, B, C.

$$\frac{\partial Q}{\partial A}=0 \qquad (6)$$

$$\frac{\partial Q}{\partial B}=0 \qquad (7)$$

$$\frac{\partial Q}{\partial C}=0 \qquad (8).$$

Solving the above-presented differential system, 3 equation system with 3 unknowns will be produced A, B, C

$$\begin{cases} A\Sigma x_i^2 + B\Sigma x_i y_i + C\Sigma x_i z_i = -\Sigma x_i \\ A\Sigma x_i y_i + B\Sigma y_i^2 + C\Sigma y_i z_i = -\Sigma y_i \\ A\Sigma x_i z_i + B\Sigma y_i z_i + C\Sigma z_i^2 = -\Sigma z_i \end{cases} \qquad (9)$$

After solving the system we will have: A= 0.3123, B= 0.0260, C= 0.373. By replacing the values of the coefficients in

$$Ax+By+Cz+1=0, \qquad (10)$$

there are calculated x', y', z' as follows:

$$x'=\frac{0.026y+0.373z+1}{0.3123} \qquad (11)$$

$$y'=\frac{0.3123x+0.373z+1}{0.026} \qquad (12)$$

$$z'=\frac{0.3123x+0.026y+0.373}{0.373}. \qquad (13)$$

Considering that the exactly values of x, y and z values cannot be reached, there are defined the following errors (14), (15) and (16):

$$\Delta x = x' - x \qquad (14)$$

$$\Delta y = y' - y \qquad (15)$$

$$\Delta z = z' - z . \qquad (16)$$

For given values of the errors, the variables x, y and z can be calculated. The errors represent the precision class in which the manufacturing is carried out. These errors represent in fact the fluctuation of the conditions for the manufacturing. This fluctuation can have many reasons but its main role is to adjust the manufacturing process in each day according to the market demands.

5. Conclusions

The developed algorithm allows the identification of the variables of one model that represents the relation between the output and the input model. This relation represents a technical-economical model that can control one manufacturing process without requesting experiments and based on the extraction of the knowledge from the previous experience.

It can be finding the values of the model that optimise the relation between the manufacturing process and the market demands. The algorithm includes methods of analytical modelling as well as intelligent artificial methods.

The resulted algorithm is perfectible and represents one start in judging manufacturing process as dynamic events that can be solved with the methods from economic science.

6. Acknowledgements

The work was carried out within the research contract CEEX 22/2005 financed by the Romanian Ministry of Education and Research in the framework of the Research of Excellence Program.

Bibliography

1. E. B. Fowlkes & C. L. Mallows (September 1983). - *A Method for Comparing Two Hierachical Clusterings*, Journal of the American Statistical Association 78 (383): 553–584.

2. Clatworthy, J., Buick, D., Hankins, M., Weinman, J., & Horne, R. (2005). *The use and reporting of cluster analysis in health psychology,* A review. British Journal of Health Psychology 10: 329-358.

3. Cole, A. J. & Wishart, D. (1970) - *An improved algorithm for the Jardine-Sibson method of generating overlapping clusters*, The Computer Journal 13(2):156-163.

4. Ester, M., Kriegel, H.P., Sander, J., and Xu, X. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, Portland, Oregon, USA: AAAI Press, pp. 226–231.

5. Heyer, L.J., Kruglyak, S. and Yooseph, S., Exploring Expression Data: Identification and Analysis of Coexpressed Genes, Genome Research 9:1106-1115.

6. S. Kotsiantis, P. Pintelas, Recent Advances in Clustering: A Brief Survey, WSEAS Transactions on Information Science and Applications, Vol 1, No 1 (73-81), 2004.

7. Huang, Z. (1998). Extensions to the K-means Algorithm for Clustering Large Datasets with Categorical Values. Data Mining and Knowledge Discovery, 2, p. 283-304.

8. Jardine, N. & Sibson, R. (1968). The construction of hierarchic and non-hierarchic classifications. The Computer Journal 11:177.

9. The on-line textbook: Information Theory, Inference, and Learning Algorithms, by David J.C. MacKay includes chapters on k-means clustering, soft k-means clustering, and derivations including the E-M algorithm and the variational view of the E-M algorithm. 11. MacQueen, J. B. (1967). Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1:281-297

10. Ng, R.T. and Han, J. 1994. Efficient and effective clustering methods for spatial data mining. Proceedings of the 20th VLDB Conference, Santiago, Chile, pp. 144–155.

11. Prinzie A., D. Van den Poel (2006), Incorporating sequential information into traditional classification models by using an element/position-sensitive SAM. Decision Support Systems 42 (2): 508-526.

12. Romesburg, H. Clarles, Cluster Analysis for Researchers, 2004, 340 pp. ISBN 1-4116-0617-5 or publisher, reprint of 1990 edition published by Krieger Pub. Co.. A Japanese language translation is available from Uchida Rokakuho Publishing Co., Ltd., Tokyo, Japan.

13. Zhang, T., Ramakrishnan, R., and Livny, M. 1996. BIRCH: An efficient data clustering method for very large databases. Proceedings of ACM SIGMOD Conference, Montreal, Canada, pp. 103–114.

14. Stan Salvador and Philip Chan, Determining the Number of Clusters/Segments in Hierarchical Clustering/Segmentation Algorithms, Proc. 16th IEEE Intl. Conf. on Tools with AI, pp. 576-584, 2004.

15. Can, F., Ozkarahan, E. A. (1990) "Concepts and effectiveness of the cover coefficient-based clustering methodology for text databases." ACM Transactions on Database Systems. 15 (4) 483-517.

16. V.N. Gaitonde, S.R. Karnik and J. Paulo Davim, Taguchi multiple-performance characteristics optimization in drilling of medium density fibreboard (MDF) to minimize delamination using utility concept, Journal of Materials Processing Technology, Volume 196, Issues 1-3, 21 January 2008, Pages 73-78..

Algoritm pentru identificarea caracteristicilor economice ale unui sistem de productie

Rezumat:

A lot of changes have appeared on the market. These are the increasing frequency of introduction the new products (that are frequently personalized), large fluctuations of the structure and production volume and some concept changes in the manufacturing tehnology. The economical characteristics is refered to the relationship between the manufacturing and the level of the economical coefficients of the manufacturing system. To identify the tehnico-economical characteristics of a manufacturing system it is necessary to develop an algorithm using artificial knowledge tehniques (Knowledge Discovery from Databasis, Artificial Neural Network, Genetic Algorithm).

Algorithmus für die Wirtschaftliche Caracteristics Identifizierung für ein Fertigungssystem

Zusammenfassung:

Viele Änderungen sind auf dem Markt erschienen. Diese sind die zunehmende Frequenz der Einführung die neuen Produkte (die oft personifiziert werden), große Schwankungen der Struktur und des Produktionsvolumens und einiger Konzeptänderungen in der Herstellung tehnology. Die wirtschaftlichen Eigenschaften werden auf die Beziehung zwischen der Herstellung und dem Niveau der wirtschaftlichen Koeffizienten des Produktionssystems verwiesen. Um die tehnico-wirtschaftlichen Eigenschaften eines Produktionssystems zu identifizieren, ist es notwendig, einen Algorithmus zu entwickeln, der künstliche Kenntnisse tehniques verwendet (Kenntnisse-Entdeckung von Databasis, Künstlichem Nervennetz, Genetischem Algorithmus).