

DEEP LEARNING OF COMPLEX INTERCONNECTED PROCESSES FOR BI-LEVEL OPTIMIZATION PROBLEM UNDER UNCERTAINTY

PhD Vatchova B. E., PhD Pavlova K. T., PhD Paunova E. N., D.Sc.Stoilova K.P.
 Institute of Information and Communication Technologies, Sofia 1113, e-mail: boriana.vatchova@gmail.com,
 kristina@hsi.iccs.bas.bg, eli_np@hsi.iccs.bas.bg, k.stoilova@hsi.iccs.bas.bg

Abstract: This research is focus on advanced modelling and design of complex interconnected processes. These processes are characterised by multiple inputs, outputs and state parameters as well as non-linearity, non-stationarity and uncertainty due to environmental disturbances. There are presented two models for deep learning and knowledge extraction using methods of multivalued logical and probable functions and networks model. These models can be applied for complex processes in the area of environment, transportation and complex systems working under uncertainty.

Keywords: COMPLEX PROCESSES, UNCERTAINTY

1. Introduction

Finding a relation between inputs and outputs of complex processes and building an adequate process model is usually the main control objective for these processes. However, the presence of uncertainty as a result of multiple factors, environmental behavior and immeasurable inputs makes the modelling of such processes quite difficult. This is the reason why existing data mining methods have to be supplemented by methods of random functions theory and multi-valued logic. The main advantage of using knowledge based systems for complex interconnected processes is their ability to remove or at least reduce the uncertainty by means of evolvable knowledge extraction based on experimental data in real time.

2. Model of Complex Processes

The complex processes are characterized with two kind of complexity quantitative and qualitative complexity. Quantitative complexity includes large number of inputs, large number of state parameters, large number of outputs. Qualitative complexity includes non-linearity, non-stationary and uncertainty, environmental disturbances and immeasurable of some inputs.

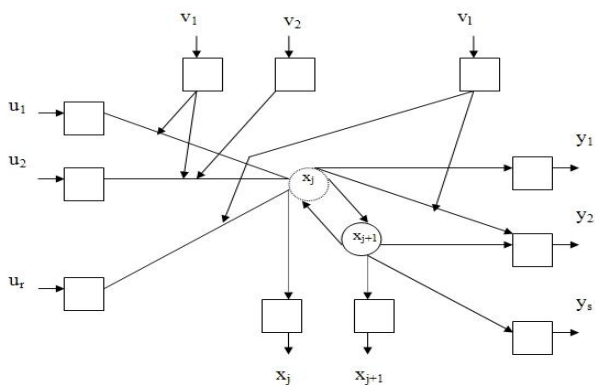


Fig.1 Structural network model of complex systems

where \$u_1, u_r\$ are the inputs
 \$v_1 \dots v_r\$ environmental disturbances
 \$y_1 \dots y_s\$ are the outputs
 \$x_j \dots x_{j+1}\$ are state parameters.

3. Novel modelling of complex processes using Logical Probability Functions (LPF) for evolvable knowledge extraction.

The strength of the proposed approach lies in a new combination of methods from several different fields such as set theory, probability theory and multi-valued logic. The combination of these methods leads to a novel hybrid approach for knowledge extraction in complex processes - the LPF approach - that handles successfully the simultaneous presence of non-linearity and non-stationarity [1].

An example for a LPF of one output \$y_1\$ of the process is presented in Table 1.

\$\langle Ly_1, p\{Ly_1\} \rangle = F\{Lx_1, Lx_2, Lx_3\}\$ where \$Ly_1\$ is the logical output, \$Lx_1, Lx_2, Lx_3\$ are the logical inputs and \$p\{Ly_1\}\$ is the probability of occurrence. A five degree multi-valued logical system \$k=5; A_5 \{a_1, a_2, a_3, a_4, a_5\}\$ is used where the degree of the logical system \$k\$ is 5. Each grouping sequence set \$GLN_r\$ from the multi-valued LPF (MLPF) is a production rule of the knowledge base for the process where \$r=1 \dots 4\$. [2].

Table 1 MLPF \$\langle Ly_1, p\{Ly_1\} \rangle = F\{Lx_1, Lx_2, Lx_3\}\$

Grouping sequence set	\$GLN_1\$	\$GLN_2\$	\$GLN_3\$	\$GLN_4\$	
Frequency of occurrence	\$p\{GLS_1\}\$	\$p\{GLS_2\}\$	\$p\{GLS_3\}\$	\$p\{GLS_4\}\$	
\$Lx_1\$	\$Lx_{11}\$	\$Lx_{21}\$	\$Lx_{31}\$	\$Lx_{41}\$	
\$Lx_2\$	\$Lx_{12}\$	\$Lx_{22}\$	\$Lx_{32}\$	\$Lx_{42}\$	
\$Lx_3\$	\$Lx_{13}\$	\$Lx_{23}\$	\$Lx_{33}\$	\$Lx_{43}\$	
\$Ly_1\$	\$a_1\$	\$p\{Ly_{111}\}\$	\$p\{Ly_{121}\}\$	\$p\{Ly_{131}\}\$	\$p\{Ly_{141}\}\$
	\$a_2\$	\$p\{Ly_{112}\}\$	\$p\{Ly_{122}\}\$	\$p\{Ly_{132}\}\$	\$p\{Ly_{142}\}\$
	\$a_3\$	\$p\{Ly_{113}\}\$	\$p\{Ly_{123}\}\$	\$p\{Ly_{133}\}\$	\$p\{Ly_{143}\}\$
	\$a_4\$	\$p\{Ly_{114}\}\$	\$p\{Ly_{124}\}\$	\$p\{Ly_{134}\}\$	\$p\{Ly_{144}\}\$
	\$a_5\$	\$p\{Ly_{115}\}\$	\$p\{Ly_{125}\}\$	\$p\{Ly_{135}\}\$	\$p\{Ly_{145}\}\$

In this case, MLPF is a knowledge base, whose knowledge is a system of production rules ordered by the columns.

Each column from Table 1 is a production rule or a knowledge base. For example, from column \$j=2\$, for grouping sequence set \$GLN_2\$ with probability of occurrence \$p\{GLS_2\}\$, the associated rule is:

$$\begin{aligned} \text{If } & \langle Lx_1=Lx_{21}, Lx_2=Lx_{22}, Lx_3=Lx_{23} \rangle & (1) \\ \text{Then } & \langle Ly_1=a_1 \text{ with probability } p\{Ly_{121}\} \\ & Ly_1=a_2 \text{ with probability } p\{Ly_{122}\} \\ & Ly_1=a_3 \text{ with probability } p\{Ly_{123}\} \\ & Ly_1=a_4 \text{ with probability } p\{Ly_{124}\} \\ & Ly_1=a_5 \text{ with probability } p\{Ly_{125}\} \rangle, \end{aligned}$$

$$\sum_{i=1}^5 p_{ij} = 1 \text{ where } i=1 \dots 5 \text{ for } j=2.$$

4. Models with evolvable knowledge bases

Using new data sets in real time generates packages of numerical values for inputs and outputs that are updated values of LPFs (B. Vatchova, Derivation and Assessment of Reliability of Knowledge for Multifactor Industrial Processes, PhD Thesis, Bulgarian Academy of Sciences, 2009). [1].

This evolvable knowledge base is a model that is a combination of production rules with the structure 'If logical values of measurable inputs, Then logical values of outputs supplemented by probabilities of occurrence', as shown in Figure 1.

In Figure 1, UGPL_q is a unified grouped package, q is the number of UGPLs, GPL_p is a grouped package, p is number of GPLs, PL_r is an individual package, r is the number of PLs, KB is a knowledge base and L is used, MLPF is a multi-valued LPF and L means 'logical' as a last letter in any acronyms. In this case, the knowledge base evolution is applied periodically by replacing old data with new data [2].

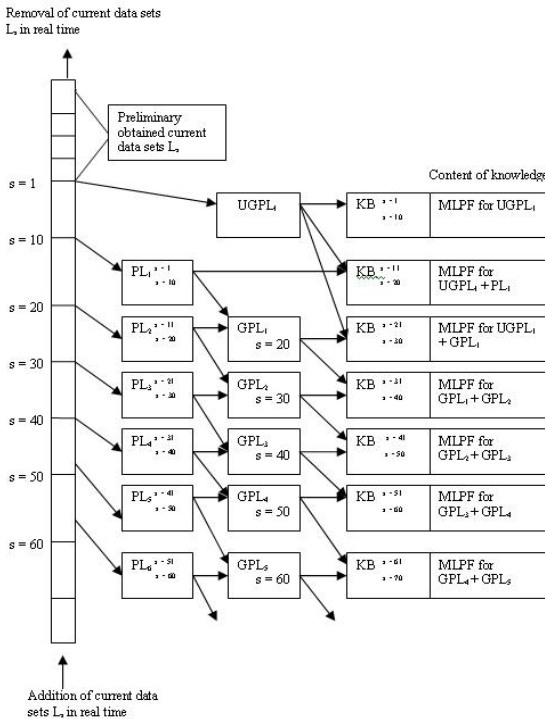


Fig.2. Structure and time relations in models with evolvable knowledge base

5. Models with network structure

The inputs to the models with network structure are logical values of the inputs L_{xij}(T), which form the first layer of the network [3]. The set of dominant data sequence sets GLN_r is the intermediate layer of the network. The set of the outputs L_{yeq} defines the output layer of the network. The connections between the elements of these layers correspond to the relations between the elements of the presented sets and the frequency of occurrence (see Figure 2). [3].

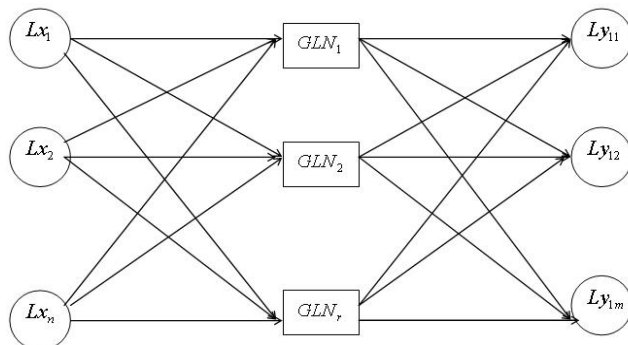


Fig.3. Models with network structure

The following relations are introduced here: R_{LXGLNr} is the relation between sets of the inputs L_{xij}(T) and dominant grouping sequence sets GLN_r; R*_{LXGLNr} is the relation between the frequency of occurrence of elements of the sets; R_{GLNrLy} is the relation

between the logical values of the elements of intermediate layer and output layer; R*_{GLNrLy} is the relation between the frequency of occurrence of the elements between the intermediate and the output layer.

Using the deep model with network structure, logical values and probability of occurrences of the outputs are calculated for each combination of logical values of measurable inputs [3].

$$Ly = R_{GLNrLy} \times R_{LXGLNr} \times LX \tag{2}$$

$$p\{Ly\} = R^*_{GLNrLy} \times R^*_{LXGLNr} \times p\{LX\} \tag{3}$$

The LPF network models are similar to neural network models. The difference between these two network models is that the proposed approach with embedded LPF doesn't have to be trained like neural networks and the elements are with computational and logical operations.

The research is multidisciplinary as it considers a wide range of real interconnected complex processes in the areas of transportation and the environment. These processes often operate under uncertainty due to environmental disturbances, vague or incomplete data.

The interdisciplinary aspect of the research is also obvious because of the application of several scientific areas such as computing, engineering and mathematics. The method for knowledge derivation will use concepts from set theory, multi-valued logic, probability theory and statistical method

Conclusions

The model with network structure is better for non-stationary processes than the model with updatable knowledge base because of its capability to interpolate new data.

The proposed methodology could be used for a wide range of real interconnected complex processes in the areas of transportation and the environment. [4],[5]. These processes often operate under uncertainty due to environmental disturbances, vague or incomplete data.

This work has been partly supported by project H12/8, 14.07.2017 of the Bulgarian National Science fund: Integrated bi-level optimization in information service for portfolio optimization contract DH12/10, 20.12.2017.

References

1. Vatchova B. Derivation and reliability evaluation of knowledge for multi-factor technological processes, Bulgarian Academy of Sciences, Sofia, 2009.
2. Gegov E.A., B.Vatchova, E.D.Gegov, Multi-valued Method for Knowledge Extraction and Updating in Real Time, Varna, IEEE, 2008, vol.2, p. 17-6 17-8.
3. Vatchova B., Gegov A.. Production rule and network structure models for knowledge extraction from complex processes under uncertainty. Chapter Recent Contributions in Intelligent Systems. Series Studies in Computational Intelligence. Editors Sgurev et al.. 657, Springer, 2017, ISBN:978-3-319-41437-9, DOI:10.1007/978-3-319-41438-6_20, 379-390. SJR:0.19
4. Stoilova, K., T. Stoilov. Hierarchical optimization for fast resource allocation. Chapter 2 in book "Time Management" Edited by Todor Stoilov, ISBN 978-953-51-0335-6, InTech publisher, 2012, p.31-46, http://www.intechopen.com/books/time-management
5. Ivanov VI. Monitoring of urban road traffic. International conference automatics and informatics Bulgaria, Sofia, October 4 - 6, 2017, 2017, ISSN:1313-1850, 135-141.