

Polytopic Representation of Parameter Varying Neural Models for Loading Systems

T. Vadvári, P. Várlaki

Széchenyi István University, System Theory Laboratory
Egyetem tér 1, H-9026 Győr, Hungary
E-mail: varlaki@sze.hu

Abstract: In the paper a heuristic approach is introduced aimed first of all to model supply chains which behavior may depend on many parameters and their analytic description is many times problematic. The proposed approach is based on identifying local vertex models in the parameter space in form of multilayer perceptrons (MLPs). The parameter varying neural system might then be modeled as the convex combination of the identified vertex systems. Depending on the dimension of the parameter space the number of identified vertex models might be large, therefore their reduction is crucial. In order to achieve this goal, first the vertex systems are transformed into HOSVD based polytopic representation followed by complexity reduction.

Keywords: MLP, LPV, modeling, logistics, HOSVD

1. Introduction

In supply chains one can observe that the information demand varies between its certain stages (see 1), i.e. the effective demand forecasting may be crucial from their performance point of view [11][8]. Usually the first stage is represented by the supplier and the last one by the customer. In-between there may be numerous stages, e.g. warehouse, retailer, etc. [14], which behavior might depend on many parameters thus the input-output characteristics of the whole system may reflect strong non-linearity. Each stage may further be divided into sub-stages making the modelling task more difficult. For designing control systems of such logistical processes the service strategy and operational algorithm of the given system must be known. Service strategy is the set of procedures, rules that determine the direction, feature and measure of state transformation of the system and its sub-systems for all possible situations and conditions [2].

Various models have already been developed for supply chains [13][9]. Most of these models are based on the fuzzy set theory, which represents the basis of many models where the handling of the imprecision is of key importance [10]. Various publications are

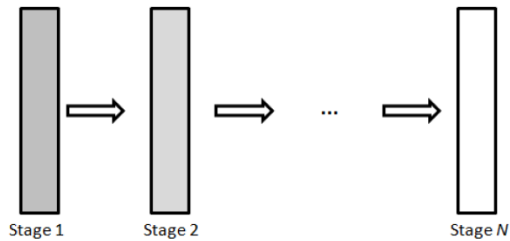


Figure 1. Illustration of state transitions between certain stages of loading system [2]

dealing with the performance of supply chains in connection with the forecasting models [11][12][7]. Most of these techniques belong to the category of heuristic or analytical approaches. Loading systems as subsystems of supply chains may efficiently be modelled by heuristic approaches.

The main contribution of the proposed approach is to identify local heuristic models over a hyper-rectangular grid in the parameter space, reduce their number and blend them together according to the actual value of parameter vector. Since the identification of local models is usually based on noisy measurements and depending on the complexity of the system the number of models to be identified might be significant it is adequate to represent the local models in such a framework in which the number of these models as well as the noise can be reduced efficiently. In other words, we are searching for such a reduced set of MLP models which convex combination approximates the behavior of the system effectively. Similar approaches are used to control non-linear systems on the basis of linear control theories [4], however in our case the models are represented by local MLPs.

The paper is organized as follows: Section 2 gives a brief description of loading systems and their models including the loading processes, in Section 3 a possible approach is described where the loading system is considered as a parameter varying system modeled by locally identified neural networks. Finally Section 4 reports conclusions.

2. Models of Loading Systems

In a loading system numerous state transitions may occur. Since the behavior of such a system is nonlinear and depends on many parameters it can be embedded into an LPV framework and modeled by the above mentioned technique, as well. Numerous measurements are needed to identify the local models in the parameter space. Fig. 3 illustrates the schematic model of the technical and technological process by a complex loading system. There are n types of goods loaded, unloaded and temporary stored in the

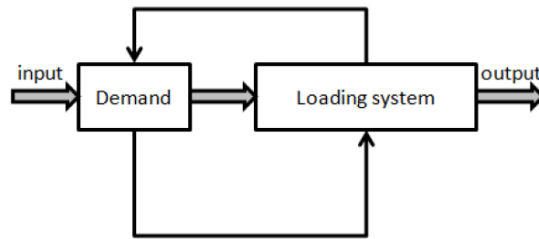


Figure 2. Illustration of the general stochastic demand- adaptive handling system [2]

analysed system. The states of processes running in the system are shown as circles and rectangles, the transitions are represented by directional arrows [2].

In case of stochastic loading systems linking two stochastic processes is causeless to suppose the independency of onset and service, especially the system contains intermittent-duty, mobil, material handling machines. Models developed on that basis are suitable for approximate disquisitions only. In case of a loading system the demand process influences the service process, controls it, accelerates or slows down the executive process. Those loading systems where the input influences the system's processes are referred as adaptive loading systems. Depending on the type of performed changes in the system three cases can be distinguished, namely the spontaneous homotroph and the heterotroph adaptivity. The spontaneous adaptivity is related to intensity changes of services, the homotroph adaptivity considers the number of elements in the system while the heterotroph adaptivity is related to functional change of elements [1],[2].

type of change this case the service intensity, the number of elements in the system

2.0.1. The Input of the Loading System

- LI - Loaded transportation vehicle arriving
- EI - Empty transportation vehicle.

2.0.2. Loading processes

- WU - Transportation vehicle waiting for unloading
- WL - Transportation vehicle waiting for loading
- TS - Temporary storing or warehouse
- SIU - Interrupted state of unloading

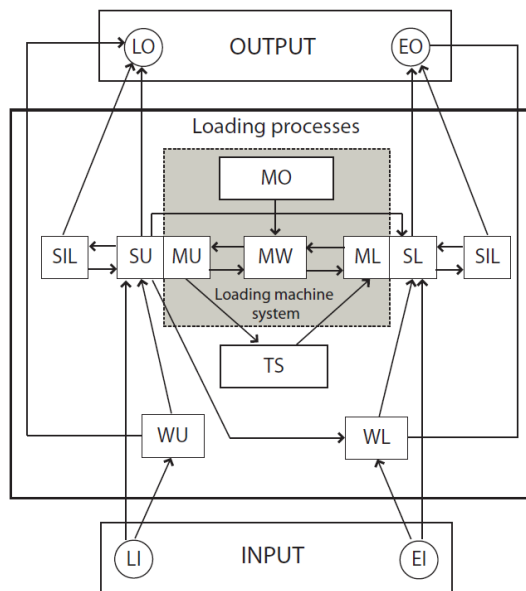


Figure 3. Illustration of loading processes[2]

- SU - State of unloading
- MU - Unloading machine
- MW - Working loading machine waiting for loading
- ML - Loading machine
- SL - State of loading
- SIL - Interrupted state of loading
- MO - Loading machine out of work

2.0.3. The Output of the Loading System

- LO - Transportation vehicle leaving loaded
- EO - Transportation vehicle leaving empty

3. MLPs in HOSVD Based Framework

By monitoring the parameters of the loading system its behavior can numerically be reconstructed and modeled. Since the number of monitored parameters might be large it is reasonable to perform the identification locally and model the whole system by an identified set of simpler models. Although the number of models necessary to accurately describe the dynamics of the whole system might be large (depending on the number of parameters and discretization points), it is reasonable to transform the local models into such a representation in which the models are ordered according to their significance in order to reduce their number. In such a framework the less significant models are responsible for modeling the details and noise, thus by dismissing few from them the noise may also be reduced.

This section deals with the problem of embedding MLPs into HOSVD based framework in order to reduce the number of vertex systems (in our case MLPs). The main advantage of this approach lies in modeling a complex parameter varying system by blending simpler MLP models together. The HOSVD-based framework offers an efficient tool for their reduction.

Let us consider a parameter varying loading system modeled by blending local multilayer perceptron (MLP) models which correspond to different parameter vectors. Let us further suppose that these local MLP models are identical in structure. Their architecture is depicted by Fig. 4. The output of such a local model can be expressed as follows [15]:

$$\mathbf{a}_3 = \varphi_3 \left(\mathbf{W}^{(3)} \varphi_2 \left(\mathbf{W}^{(2)} \varphi_1 \left(\mathbf{W}^{(1)} \mathbf{h} \right) \right) \right) \quad (1)$$

where

$$\mathbf{W}^{(j)} = \begin{pmatrix} w_{11}^{(j)} & w_{12}^{(j)} & \cdots & w_{1S_j-1}^{(j)} & b_{j1} \\ w_{21}^{(j)} & w_{22}^{(j)} & & w_{2S_j-1}^{(j)} & b_{j2} \\ \vdots & & & & \\ w_{S_j1}^{(j)} & w_{S_j2}^{(j)} & & w_{S_jS_j-1}^{(j)} & b_{jS_j} \end{pmatrix},$$

$j = 1..N_L$ (Number of layers (N_L) in our example represented by Fig. 4 is 3),

$$\mathbf{h} = (h_1 \quad h_2 \quad \cdots \quad h_R \quad 1)^T$$

stand for the input vector, while vector

$$\mathbf{a}_3 = (a_{31} \quad a_{31} \quad \cdots \quad a_{3S_3})^T$$

represents the output of the NN in Fig. 4.

In order to identify the vertex MLP-based models - corresponding to the nodes of a hyper-rectangular grid in the parameter space - measurements are needed. The nodes of

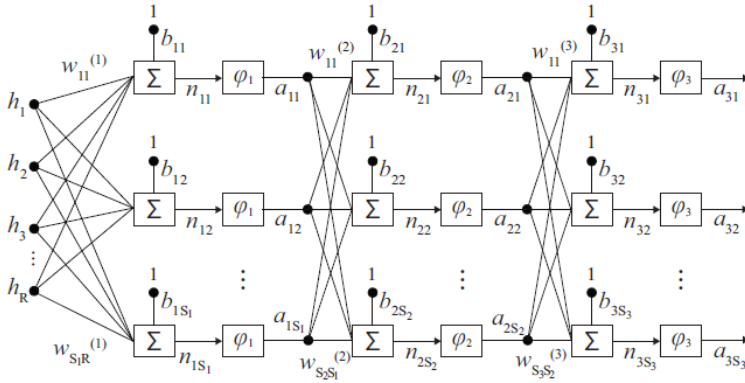


Figure 4. Architecture of vertex neural models [15]. ($R = S_0$)

the hyper-rectangular grid stand for those parameter vector values for which the local behavior of the system has to be identified. Let us denote the number of discretization points for the j th dimension as I_j . Based on such measurements (input-output pairs) the proper weights of the vertex MLP-based models are determined. From the weight matrices of corresponding layers of vertex MLP models an $N+2$ dimensional tensor $\mathcal{B}_i \in \mathfrak{R}^{I_1 \times I_2 \times \dots \times I_N \times S_j \times (1+S_{j-1})}$ can be constructed [15]. Index i indicates the layer index of the MLP. Since in case of a loading system the number of parameters N and the number of identified vertex models corresponding to the certain dimensions may be large it is reasonable to reduce the complexity, i.e. the number of considered vertex models. The reduction can be performed efficiently by decomposing \mathcal{B}_i (containing the weight matrices of corresponding layers) to a set of orthonormal components by HOSVD. Since the approximation is performed in the parameter space, the HOSVD is performed only for the 1st N dimensions of \mathcal{B}_i . The decomposition will result a core tensor \mathcal{D}_i having the same size as \mathcal{B}_i and N pieces of orthonormal matrices in which the columns stand for the components ordered according to their significance. From such a form the less significant components can be dismissed, thus the number of vertex models can be reduced [5],[15]:

$$\mathbf{a}_3(\mathbf{p}) = \varphi_3 \left(\mathbf{W}^{(3)}(\mathbf{p}) \varphi_2 \left(\mathbf{W}^{(2)}(\mathbf{p}) \varphi_1 \left(\mathbf{W}^{(1)}(\mathbf{p}) \mathbf{h} \right) \right) \right), \quad (2)$$

where

$$\mathbf{W}^{(1)}(\mathbf{p}) = \mathcal{D}_1 \boxtimes_{n=1}^N \xi_n^{(1)}(p_n), \quad (3)$$

$$\mathbf{W}^{(2)}(\mathbf{p}) = \mathcal{D}_2 \boxtimes_{n=1}^N \xi_n^{(2)}(p_n), \quad (4)$$

$$\mathbf{W}^{(3)}(\mathbf{p}) = \mathcal{D}_3 \boxtimes_{n=1}^N \xi_n^{(3)}(p_n), \quad (5)$$

where \mathbf{p} stands for the vector of parameters, \mathcal{D}_i the core tensor corresponding to the i th layer and the elements of vector valued functions

$$\xi_n^{(i)}(p_n) = \left(\xi_{n1}^{(i)}(p_n) \quad \xi_{n2}^{(i)}(p_n) \quad \cdots \quad \xi_{nI_n}^{(i)}(p_n) \right)$$

are the function values of components (corresponding to the n th dimension of \mathcal{D}_i) at parameter value p_n .

3.1. Complexity Reduction

Let us express equations (3)-(5) in the following form:

$$\mathbf{W}^{(1)}(\mathbf{p}) = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} \prod_{n=1}^N \xi_{ni_n}^{(1)}(p_n) \mathbf{d}_{i_1, i_2, \dots, i_N}^{(1)} \quad (6)$$

$$\mathbf{W}^{(2)}(\mathbf{p}) = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} \prod_{n=1}^N \xi_{ni_n}^{(2)}(p_n) \mathbf{d}_{i_1, i_2, \dots, i_N}^{(2)} \quad (7)$$

$$\mathbf{W}^{(3)}(\mathbf{p}) = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} \prod_{n=1}^N \xi_{ni_n}^{(3)}(p_n) \mathbf{d}_{i_1, i_2, \dots, i_N}^{(3)} \quad (8)$$

It follows from the properties of the HOSVD, that the rightmost columns of singular matrices in each dimension represent those weighting functions which are responsible for describing the details involved in the measurement data. According to the above indexing the ξ_{ni_n} functions with larger i_n index are of less significance, thus by dismissing them the number of local models can be reduced by keeping the approximation error at lower level. It is the best low rank approximation of the original layer tensor, i.e. giving the minimal Frobenius norm $\left\| \mathbf{W}^{(i)} - \widetilde{\mathbf{W}}^{(i)} \right\|$, where i denotes the layer index and $\widetilde{\mathbf{W}}^{(i)}$ stands for the approximation of $\mathbf{W}^{(i)}$ as follows:

$$\widetilde{\mathbf{W}}^{(1)}(\mathbf{p}) = \sum_{i_1=1}^{I_{1r}} \sum_{i_2=1}^{I_{2r}} \cdots \sum_{i_N=1}^{I_{Nr}} \prod_{n=1}^N \xi_{ni_n}^{(1)}(p_n) \mathbf{d}_{i_1, i_2, \dots, i_N}^{(1)} \quad (9)$$

$$\widetilde{\mathbf{W}}^{(2)}(\mathbf{p}) = \sum_{i_1=1}^{I_{1r}} \sum_{i_2=1}^{I_{2r}} \cdots \sum_{i_N=1}^{I_{Nr}} \prod_{n=1}^N \xi_{ni_n}^{(2)}(p_n) \mathbf{d}_{i_1, i_2, \dots, i_N}^{(2)} \quad (10)$$

$$\widetilde{\mathbf{W}}^{(3)}(\mathbf{p}) = \sum_{i_1=1}^{I_{1r}} \sum_{i_2=1}^{I_{2r}} \cdots \sum_{i_N=1}^{I_{Nr}} \prod_{n=1}^N \xi_{ni_n}^{(3)}(p_n) \mathbf{d}_{i_1, i_2, \dots, i_N}^{(3)}, \quad (11)$$

where $0 < I_{j_r} < I_j, j = 1..N$, matrices $\mathbf{d}_{i_1, i_2, \dots, i_N}^{(i)}$ stand for the weight matrices of vertex models for the i th layer. The output of the approximated MLP can then be expressed as:

$$\tilde{\mathbf{a}}_3 = \varphi_3 \left(\tilde{\mathbf{W}}^{(3)} \varphi_2 \left(\tilde{\mathbf{W}}^{(2)} \varphi_1 \left(\tilde{\mathbf{W}}^{(1)} \mathbf{h} \right) \right) \right) \quad (12)$$

4. Conclusions

In the paper a theoretical approach has been introduced for modeling loading systems having various parameters. Instead of looking onto the system as a unit, we proposed to model it by the combination of simpler models identified in the parameter space. As simpler models local MLPs have been considered and transformed into polytopic representation in order to blend them efficiently together according to the actual parameter vector as well as to reduce the number of identified models according to their significance. As future work we are going to test the suitability of the approach on real word data.

References

- [1] Prezenszki, J., Várlaki, P.: *Sztochasztikus igényfolyamattal vezérelt adaptív viselkedésű rakodási rendszerek vizsgálati módszereinek elemzése*, Közlekedéstudományi Szemle, No. 2, pp. 78–85, 1976
- [2] Orbán, G., Várlaki, P.: *Fuzzy Modelling for Service Strategy and Operational Control of Loading Systems*, Acta Technica Jaurinensis, Series Logistica, Vol. 2. No. 3., pp. 375–391, 2009
- [3] De Lathauwer, L., De Moor, B., Vandewalle, J.: *A multilinear singular value decomposition* SIAM Journal on Matrix Analysis and Applications, Vol. 21, No. 4, pp. 1253–1278, 2000
DOI:10.1137/S0895479896305696
- [4] Harmati, I., Rövid, A., Várlaki, P.: *Approximation of Force and Energy in Vehicle Crash Using LPV Type Description*, WSEAS TRANSACTIONS on SYSTEMS, Vol. 9, No. 7, pp. 734–743, 2010 URL: <http://dl.acm.org/citation.cfm?id=1865416.1865420>
- [5] Szeidl, L., Várlaki, P.: *HOSVD Based Canonical Form for Polytopic Models of Dynamic Systems*, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol. 13 No. 1, pp. 52–60, 2009, URL: <http://www.fujipress.jp/finder/xslt.php?mode=present&inputfile=JACII001300010007.xml>
- [6] Roemer, F., Becker, H., Haardt, M., Weis, M.: *Analytical Performance Evaluation for HOSVD-based Parameter Estimation Schemes*, 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP),

No. 13-16, pp.77–80, 2009

DOI: 10.1109/CAMSAP.2009.5413232

- [7] Song Q., Chissom B. S.: *Forecasting Enrollments with Fuzzy Time Series-Part II*, Fuzzy Sets and Systems, Vol. 54, pp. 1–9, 1993
DOI: 10.1016/0165-0114(94)90067-1
- [8] Hakan, T., Ozalp, V.: *Fuzzy Forecasting Applications on Supply Chains*, WSEAS TRANSACTIONS on SYSTEMS, Vol. 7, No. 5, pp. 600–609, 2008, URL: <http://dl.acm.org/citation.cfm?id=1456007.1456025>
- [9] Mohammad, H., Zarandi, F., Türkşen, I.B., Saghiri, S.: *Supply Chain: Crisp and Fuzzy Aspects*, International Journal of Applied Mathematics and Computer Science, Vol.12, No.3, pp. 423–435, 2002, URL: <http://pldml.icm.edu.pl/pldml/element/bwmeta1.element.bwnjournal-article-amcv12i3p423bwm>
- [10] Petrovic, D., Roy, R., Petrovic, R.: *Supply chain modelling using fuzzy sets*, Int. J. Production Economics, Vol. 59, pp. 443–453, 1999
DOI: 10.1016/S0925-5273(98)00109-1
- [11] Hakan, T., Ozalp, V.: *The Effects of Fuzzy Forecasting Models on Supply Chain Performance*, 9th WSEAS International Conference on FUZZY SYSTEMS, Sofia, Bulgaria, pp. 107–112, 2008, URL: <http://dl.acm.org/citation.cfm?id=1416056.1416073>
- [12] Tsaur, R.C.: *Fuzzy Grey GM(1, 1) Model Under Fuzzy Systems*, International Journal of Computer Mathematics, Vol. 82, No. 2, pp. 141–149, 2005
DOI:10.1080/0020716042000301770
- [13] Mohammad, H., Zarandi, F., Mohammad, M., Zarani, F., Saghiri, S.: *Five crisp and fuzzy models for supply chain of an automotive manufacturing system*, International Journal of Management Science and Engineering Management, Vol. 2, No. 3, England, UK, pp. 178–196, 2007
DOI:10.1080/17509653.2007.10671020
- [14] Chang, Y., Makatsoris, H.: *Supply Chain Modeling Using Simulation*, I. J. of SIMULATION, Vol. 2 No. 1, pp. 24–30, 2001, URL: <http://ijssst.info/Vol-02/No-1/Chang.pdf>
- [15] Rövid, A., Szeidl, L., Várlaki, P.: *The HOSVD Based Canonical Form of Functions and Its Applications*, In Aspects of Computational Intelligence, Springer Verlag, L. Madarász, J. Zivcák (editors.), pp. 221–231, 2012
DOI: 10.1007/978-3-642-30668-6_15