A Nonlinear Model Predictive Control Strategy Using Multiple Neural Network Models

Zainal Ahmad[1] and Jie Zhang[2]

[1] School of Chemical Engineering, University Sains Malaysia, Engineering Campus, Seri Ampangan, 14300, Nibong Tebal, Penang, Malaysia
chzahmad@eng.usm.my
[2] School of Chemical Engineering and Advanced Materials, University of Newcastle, Newcastle upon Tyne NE1 7RU, UK
jie.zhang@newcastle.ac.uk

Abstract. Combining multiple neural networks appears to be a very promising approach for improving neural network generalization since it is very difficult, if not impossible, to develop a perfect single neural network. Therefore in this paper, a nonlinear model predictive control (NMPC) strategy using multiple neural networks is proposed. Instead of using a single neural network as a model, multiple neural networks are developed and combined to model the nonlinear process and then used in NMPC. The proposed technique is applied to water level control in a conic water tank. Application results demonstrate that the proposed technique can significantly improve both setpoint tracking and disturbance rejection performance.

1 Introduction

Neural networks have been widely used not only in the engineering field but also in other applications like remote sensing, transportation, power systems, medicine, telecommunication, and banking. The growing interests in applying neural networks are due to the rapid development in computing power which enables neural networks being trained in short time durations when modeling the behavior of complex systems. Furthermore the characteristic of neural network models themselves that learn from examples rather than having to be programmed also contributed their increased applications. The architecture of single neural networks varies from multilayer perceptron to radial basis function and also recurrent neural networks. Currently, applications of single neural networks in process modeling and control are quite significant in industry especially in model based predictive control (MBPC) [1]. This is mainly due to the capability of neural networks in modeling nonlinear processes from process operation data. However, single neural networks usually lack generalization capability due to over-fitting, limitation of training data, and network training trapped in undesirable local minima. Recent studies have shown that this limitation can be overcome by combining multiple neural networks. Fig. 1 shows how multiple neural networks are combined. The individual networks in Fig. 1 model the same relationship and are developed from different data sets and/or different training algorithms. They can also have different structures. Instead of choosing the single "best"

J. Wang et al. (Eds.): ISNN 2006, LNCS 3972, pp. 943 – 948, 2006.

neural network model, all the individual neural networks are combined. Note here that if a single network is selected, then it is a common practice to select the best network on the training and/or testing data. However, this "best" network may not be the best when applied to unseen data. There are a number of methods in combining the networks like stacked neural network and bootstrap aggregated network where multiple networks are created on bootstrap re-samples of the original training data [2],[3].

A nonlinear model predictive control (NMPC) strategy using multiple neural networks is proposed in this paper. NMPC basically requires an accurate model of the nonlinear process to be controlled and predict the controlled variable over a future prediction horizon under a sequence of future control actions. An optimization procedure is then carried out to minimize the control errors over the prediction horizon by finding the appropriate sequence of control actions. Since a multiple neural network can offer enhanced prediction accuracy, it is expected that multiple neural network based NMPC can outperform single neural network based NMPC.

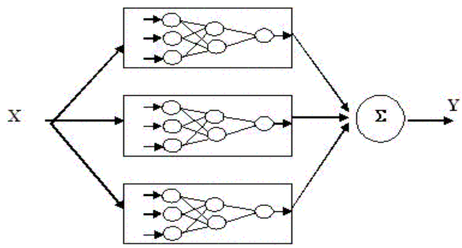

Fig. 1. Combining multiple neural networks

2 NMPC Using Multiple Neural Networks

Model Predictive Control (MPC) basically is a methodology that refers to a class of control algorithms in which a dynamic model of the plant is used to predict and optimize the future behavior of the process. At each control interval, the MPC algorithm computes a sequence of the manipulated variables in such a way to optimize the future behavior of the plant. MPC has been used in industry for nearly 30 years, and has become an industry standard (mainly in the petrochemical industry) due to its intrinsic capability for dealing with constraints and with multivariable systems. Most commercially available MPC technologies are based on a linear model of the controlled process. For processes that are highly nonlinear, the performance of linear model based MPC can be poor. This has motivated the development of Nonlinear Model Predictive Control (NMPC), where a more accurate nonlinear model of the plant is used for prediction and optimization. Many of the current NMPC schemes are based on physical models of the controlled processes. However, in many cases such models are difficult to obtain, and often not available at all. Neural network model based NMPC has been reported by Zhan and Ishida [4]. The basic structure of this predictive control scheme can be found in reference [5] which has been simplified by Hussain. In this paper we propose to use robust multiple neural network models in NMPC. Here instead of using single neural network model, multiple networks are used. The objective function of NMPC is given as

$$\min \sum_{k=1}^{N} \left(y_{sp}(t+k) - y_{predic}(t+k)\right)^2 + \sum_{k=1}^{M} \lambda\left(u(t+k-1) - u(t+k-2)\right)^2 \quad (1)$$

The optimization is subject to the following constraints

$$y_{\min} \le y_{\text{predict}}(t+k) \le y_{\max} \quad (k=1,..,N) \quad (2)$$

$$u_{\min} \le u(t+k) \le u_{\max} \quad (k=0,1,..,M-1) \quad (3)$$

$$\left|u(t+k) - u(t+k-1)\right| \le \Delta u_{\max} \quad (k=1,..,N) \quad (4)$$

where N is the predictive horizon, M is the control horizon, and λ is the control weight or a penalty term for large control action variations. The decision variable u(t+k) (k=0,..,M-1) are the control moves over a manipulated input horizon M (M<N) and are assumed to be kept constant for the remaining of the sampling intervals.

$$u(t+k) = u(t+M-1) \text{ (for } k=M,..,N-1) \quad (5)$$

Although the optimal control actions are obtained for a future control horizon, only the first control action is implemented. Then process output measurements are obtained and the process/model mismatch is compensated. The optimization is performed again at the next sampling interval, applying the following compensation for process (y(t))/model ($y^m(t)$) mismatch:

$$d(t) = c[y(t) - y^m(t)] \quad (0<c<1.0) \quad (6)$$

$$y_{predict}(t+k) = y^m(t+k) + d(t) \quad (\text{for all } k = 1,..,N) \quad (7)$$

where *c* is a tunable parameter.

3 Result and Discussion

Conic water tanks were used in this case study. There is an inlet stream to the tank and an outlet stream from the tank. Manipulating the inlet water flow rate will regulate the water tank level. A detailed schematics diagram can be found in reference [6] and mechanistic model is developed based on material balance and is used to simulate the process. The sampling time used is 10 seconds. The detailed mechanistic model indicates that the relationship between the inlet water flow rate and the water level in the tank is quite nonlinear. The outlet valve characteristic determines that the static gain increases with tank level. Because the tank is of a conical shape, the time constant of the processes increases with the tank level. Thus, both the static and dynamic characteristics of the process vary with the operating condition. All the network building data were generated from the simulation program and normally distributed noise with zero mean and a standard deviation of 0.7cm were added to the simulated tank level. A multiple neural network containing 20 individual networks with fixed structure (single hidden layer with 4 hidden neurons) and a multiple neural network containing 20 individual networks with various structures (single hidden layer with the number of hidden neurons ranging from 1 to 10) were developed to

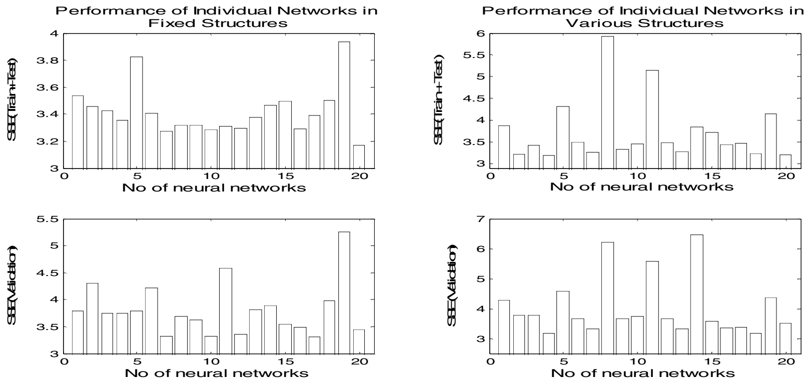

Fig. 2. SSE of one-step-ahead predictions from individual neural networks

model the nonlinear dynamic relationship between the inlet water flow rate and the water level in the tank from the process operating data.

In this study, multiple neural networks were combined using the simple averaging approach. The appropriate model structure was determined through cross validation. Fig. 2 shows the sum of squared errors (SSE) of one-step-ahead predictions from the individual networks on the training and testing data and on the unseen validation data. It can be seen that the individual networks give inconsistent performance on the training and testing data and on the unseen validation data. This indicated that non-robust nature of single neural networks. Neural network generalization capability can be combined by combining multiple neural networks. As shown in Fig. 3, the long range prediction performance of the combined neural network models with fixed and various structures is quite good, even though there are some errors in the predictions but the predictions still follow the trend in the actual data. Then NMPC is developed using the multiple neural networks to control the tank level.

The optimization method used is the sequential quadratic programming implemented as the *constr.m* function in the Matlab™ Optimisation Toolbox. The NMPC

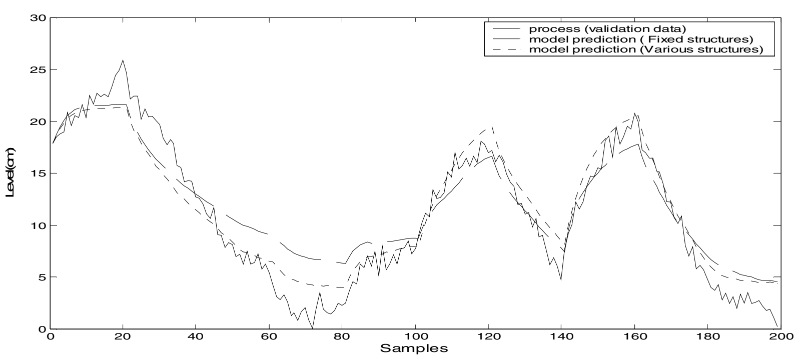

Fig. 3. Long range predictions on validation data from multiple neural network models with fixed and various structures

Table 1. Nominal values for simulation study in NMPC

Variables	Meanings	Nominal values
N	Prediction horizon	7
M	Control horizon	1
y_{min}	Minimum of the controlled variable	5 cm
y_{max}	Maximum of the controlled variable	30 cm
u_{max}	Maximum of the manipulated variable	250 cm^3/s
u_{min}	Minimum of the manipulated variable	70 cm^3/s
Δu_{max}	Maximum change in manipulated variable	180 cm^3/s
λ	Control weight	0.02
c	Integral term	0.1

parameters are listed in Table 1. To eliminate the offset of the prediction output, model plant mismatch is compensated using Eq (4) and Eq (5). An integral action was also added in the control action to eliminate any static offsets. Overall the setpoint tracking performance of NMPC based on multiple neural networks with various structures is better than based on multiple neural networks with fixed structures as shown in Fig. 4 and Table 2 especially in the high region of the operating condition. Even though the NMPC base on the "best" single network model gives quite poor performance in the

Table 2. Sum of squared control errors

	Setpoint tracking		Disturbance rejection	
	Fixed structures	Various structures	Fixed structures	Various structures
NMPC using 'best'	18495	8426	13.81	31.52
NMPC using multiple	5795	5419	14.60	7.89

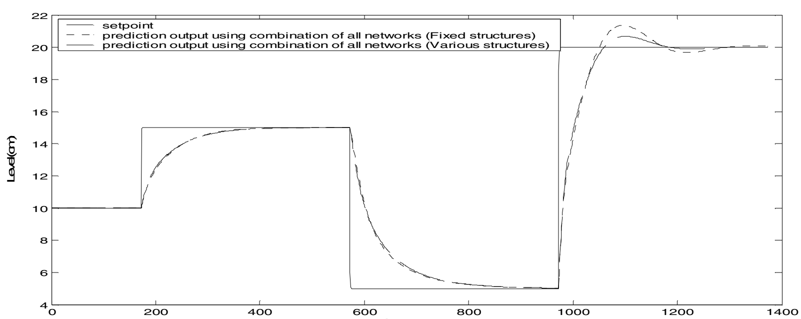

Fig. 4. Setpoint tracking performance for NMPC based on multiple neural networks with fixed and various structures

setpoint tracking as shown in Table 2 but by combining all these networks, better control performance is obtained. In disturbance rejection performance, multiple neural networks are superior to single neural networks as shown in Table 2. The sum of squared control errors from NMPC based on multiple neural networks is much lower than that from NMPC based on single neural networks.

4 Conclusions

Implementing NMPC using multiple neural network models is investigated in this paper. Instead of using the conventional single neural network models, multiple neural network models are used in order to take the advantage of their enhanced long range prediction performance. It is shown in the case study that NMPC based on multiple neural network models gives improved control performance in both setpoint tracking and disturbance rejection.

Acknowledgements

This work was supported by University Science Malaysia (for Z. Ahmad) and UK EPSRC through the Grant GR/R10875 (for J. Zhang).

References

1. Chen, J., Yea, Y.: Neural Network-Based Predictive Control for Multivariable Processes. Chemical Engineering and Communication 189 (2002) 866-894
2. Sridhar, D. V., Bartlett, E. B., Seagrave, R. C.: An Information Theoretic Approach for Combining Neural Network Process Models. Neural Networks 12 (1999) 915-926
3. Zhang, J.: Sequential Training of Bootstrap Aggregated Neural Networks for Nonlinear Systems Modeling. American Control Conference, Vol. 1 (2002) 531-536
4. Zhan, J., Ishida, M.: The Multi-step Predictive Control of Nonlinear SISO Processes with a Neural Model Predictive Control (NMPC) Method. Computers and Chemical Engineering 21(2) (1997) 201-210
5. Hussein, M. A.: Review of the Application of Neural Networks in Chemical Process Control-Simulation and Implementation. Artificial Intelligence in Engineering 13 (1999) 55-68
6. Zhang, J.: Developing Robust Neural Network Models by Using Both Dynamic and Static Process Operating Data. Ind. Eng. Chem. Res. 40 (2001) 234-241