Newly developed double neural network concept for reliable fast plasma position control

Young-Mu Jeon, Yong-Su Na, Myung-Rak Kim, and Y. S. Hwang
Department of Nuclear Engineering, Seoul National University, Seoul 151-742, Korea

(Presented on 19 June 2000)

Neural network is considered as a parameter estimation tool in plasma controls for next generation tokamak such as ITER. The neural network has been reported to be so accurate and fast for plasma equilibrium identification that it may be applied to the control of complex tokamak plasmas. For this application, the reliability of the conventional neural network needs to be improved. In this study, a new idea of double neural network is developed to achieve this. The new idea has been applied to simple plasma position identification of KSTAR tokamak for feasibility test. Characteristics of the concept show higher reliability and fault tolerance even in severe faulty conditions, which may make neural network applicable to plasma control reliably and widely in future tokamaks. © 2001 American Institute of Physics. [DOI: 10.1063/1.1323251]

I. INTRODUCTION

Neural network (NN), which consists of several layers and nodes on each one, is the structure or the algorithm that recognizes the system and acts onto that through the statistical learning. It is particularly a good tool for parameter estimation of the nonlinear and complicated problem since it can represent system properties very fast through the nonphysical statistical learning. Because of the good properties, NN was introduced and has been studied at DIII-D,1 COMPASS-D,2 ASDEX-U,3 and ITER,4 etc., for some years as a tool for the plasma position and shape identification in tokamak.

Conventional NN has good accuracy of very small root-mean-squared (rms) error. However, the possibility of a large peaked-error deviation renders some tokamak operators reluctant to use it as a routine control method even after its operational demonstration. To promote NN as a reliable tool for plasma control, this possibility should be eliminated while keeping rms error small. In this study, a concept of double neural network (DNN) is introduced as a new tool that increases the reliability of the conventional NN for the control system. DNN will be shown to eliminate the large peaked error and to make the average error smaller than the conventional NN through its modified structure.

Section II describes characteristics of the conventional NN, and then a new concept of DNN is introduced in Sec. III. This method has been tested by applying to the position identification of KSTAR tokamak plasmas.5 These results are discussed in Sec. IV. Conclusions are given in Sec. V.

II. CHARACTERISTICS OF CONVENTIONAL NEURAL NETWORK

Generally speaking, NN indicates a multilayer perceptron (MLP) or a radial basis function (RBF) network.6 Here, MLP is considered as NN. NN consists of an input layer into which the sources are entered, an output layer from which the network response (or action) returns, and some hidden layers which interconnect the input and the output layer. Each layer has several nodes and all nodes are fully connected into nodes in the next layer, and the connection strength is represented by weight value.

In NN, the learning (training) is updating the weight value (or connection length) to minimize the cost function, i.e., the average squared error energy of network output errors. The training data are fed forward through the network, and the updating procedure works by the feedback direction, which is called the back-propagation algorithm.6

Although NN is trained through the nonphysical and statistical procedure, it has good accuracy and fast computational speed. Also, it obtains high fault and noise tolerance from plenty of interconnections (distributed structure) because of the distributed nature of information. However, NN has some drawbacks. Since the relation between inputs and outputs has no physical meaning, sufficient training data are required to learn system properties. Long optimizing (training) processes are required since these processes are performed by empirical approach instead of systematic approach. Another possible disadvantage comes from the possibility of large maximum (MAX) errors even with very small rms errors. The distributed structure of neural network makes the rms error very low, but sometimes makes locally unexpected large MAX errors incurred by itself.

III. DOUBLE NEURAL NETWORK AS A MODIFIED NEURAL NETWORK

To overcome the drawbacks of the conventional NN as described in the preceding section, modified NN with double network structure is introduced. The structure of DNN is shown in Fig. 1. It is composed of a primary network and a secondary network which are connected in series. The primary network is a set of several NNs which are connected in parallel and have the same output set. In Fig. 1, the primary network is represented as {MLP(1), MLP(2),..., MLP(s)} in mathematical notation. Each element network in the primary network, MLP(i), has several layers and nodes as a complete single NN. A set of corresponding outputs is fed into the...
secondary network as its input set. For example, the first input set of the secondary network is \(\{y_{11}, y_{21}, ..., y_{m1}\}\) from the primary network output as shown in Fig. 1. A single element network in the secondary network works similar to the averaging algorithm and produces one corresponding output. Of course, the number of element networks in the secondary network is the same as that of final output parameters to be obtained or of outputs from each element network in the primary network.

DNN with a special structure modified from single NN is expected to reduce both rms and MAX errors simultaneously. These improvements may come from the modified structure of the conventional NN. The structure of NN has many degrees of freedom to be optimized for better fitting. Especially, numbers of layers and nodes should be changed to get optimum structure. Though improved fitting is expected by adding an additional layer in the conventional NN, sometimes it may end up worse fitting due to overrestriction. On the other hand, DNN is an alternative of conventional NN to obtain some specific properties by adding partial connections and two split training processes via double structures without imposing such an overrestriction. The secondary network may play an important role to eliminate the peaked-error distribution of conventional NN as well as to reduce rms errors further down than those of a single NN. Also, optimization processes for primary networks, which are the most time-consuming part, may be performed in a simple manner by compromising with an accurate training process for the relatively simple secondary network in this structure.

Moreover, an additional module can be inserted between the primary and the secondary network because of its divided structure. Since DNN can accommodate an additional module between two networks, the error-filter module can be added to estimate and recover faulted values from the primary network before the secondary network is applied. With some signal faults, an error-filter module is expected to mitigate their effects significantly.

These advantageous properties of DNN are tested by applying this new concept to a simple system of plasma position identification in KSTAR tokamak. Their results are described in Sec. IV.

**IV. APPLICATION TO KSTAR PLASMA POSITION IDENTIFICATION**

The new algorithm has been tested by applying to the plasma position identification during the ramp-up stage of KSTAR tokamak plasmas.\(^6,7\) Plasma equilibria are assumed to be up–down symmetric, and evolving from the outboard-limited small plasma to the inboard-limited full-sized one. In this test, a partial set of magnetic diagnostic sensors has been selected from full KSTAR magnetic sensors shown in Fig. 2 by the sensitivity test for the plasma position. Thirty-five sensors out of total 78 are selected from those in the upper half plane.

Forty-six equilibrium data for NN training are produced by tokamak simulation code (TSC).\(^8\) The plasma position parameter, \(R_o-a\), changes from 1.30 m to 1.75 m with the increment of 0.01 m, and magnetic measurement data are...
normalized by plasma current. To test fault tolerance, the signal value of a specific faulted sensor is assumed to be zero with the faulty condition. Also, Gaussian distribution errors are imposed to simulate realistic experimental environments.

### A. Description of DNN modeling

DNN configuration for plasma position identification in KSTAR is described in Fig. 3. The primary network consists of six element networks (MLPs) whose multilayer structures are $6-4-2-1$ except the last one with $5-3-2-1$, and each one uses subgroup signals of the previous selected sensors as its input. Each output from all element networks is the plasma radial position parameter, $R_{o-a}$, and this is fed into the secondary network after error filtering. Error-filter module is designed by assuming that the input set of each element network in secondary network has some distributions such as the Gaussian distribution. If there exists any significant deviation in input values of the secondary network, i.e., beyond the 95%-reliability range from their mean value, the significantly deviated input value is substituted by the average value of the other input values.

The results from this double network is compared with that of a conventional single NN which is composed of only one MLP with $35-20-10-1$ structure using all selected magnetic sensor signals as input values. Table I shows input data set of each subgroup in the primary network of DNN.

### B. Effects of double neural network

Network output errors for 46 testing data including 5% Gaussian errors are shown in Fig. 4. A peaked point with very large-valued error appears in the dashed circular line of Fig. 4, an error distribution of single network. On the other hand, a mild error distribution with small average value is observed in DNN. Sometimes, the presence of this kind of peak in the single NN is one major reason why the fast and accurate neural network is not welcome by plasma control engineers. Although the average error of conventional single NN is sufficiently small below the error threshold (control system criterion), the system will suffer crucial damages or be out of control with this kind of unexpected large peaked error. However, the mild error distribution of DNN with small average error and nonpeaked value simultaneously confirms the high reliability of the network.

Test results by increasing Gaussian errors are given in Fig. 5. DNN shows the best result of error-reducing effect for both rms and MAX errors. DNN is a reliable method even without an extra error filter as long as the level of random errors are maintained in a reasonable level, i.e., for the case of no sensor fault event.

### C. Effects of error filtering

Generally NN has a strong fault tolerance due to the distributed structure, i.e., layers, modes, and their interconnections. However, conventional NN is not enough to over-
come severe faulty conditions like sensor faults or large unexpected error peaks. On the other hand, DNN, series-connected double structure, shows high fault tolerance and the structural advantage of accommodating an error filtering module between the primary and secondary networks as shown in Fig. 3. With the error filter, robust fault tolerance and higher accuracy of the network are obtained even in a severe condition such as a complete sensor fault.

In addition, various cases of magnetic diagnostic sensor faults are tested in this study as shown in Table II. Results by combining DNN with error filtering are shown in Fig. 6 for various magnetic diagnostic sensor fault cases. The rms error of network outputs in Fig. 6(a) and the MAX error in Fig. 6(b) are reduced significantly to almost the same level of no fault condition. Error filtering module inserted between the primary and secondary networks in DNN is good enough to eliminate effect of any severe error such as a complete faulty sensor showing small average error and error deviation.

**D. Effects of secondary network structure**

The main role of the secondary network seems like an averaging. Adding a simple average algorithm as a secondary network seems to be much easier than adding another NN as a secondary one only for realizing the double network (DN) structure. However, test results shown in Figs. 4 and 5 represent that NN is better than averaging algorithm as a choice of the secondary network, though the effects are not significant for some cases. For uniform distributed outputs from the primary network, averaging algorithm instead of NN is good enough for the secondary network of DN. However, the difference can be significant when outputs of the primary network have some patterns different from random distribution. For example, all outputs are above the desired value, or below the desired value, or some definite and others dispersed. In these cases, DNN is more accurate and robust than DN with averaging algorithm since the second NN can be trained to reflect these patterns.

**V. CONCLUSION**

Newly developed DNN concept was introduced and has been tested for several conditions. DNN is shown to be more accurate and robust than conventional NN, and it is immune from any unexpected large-error event. By combining with error-filtering step, which can be added between the primary and the secondary networks, DNN gives very strong fault tolerance even in a severe condition such as a complete sensor fault. Therefore, these characteristics of DNN make it such a reliable tool for the plasma position identification that it may expedite the application of NN to the actual plasma control system. This DN concept can be applied to other complicated systems extensively.

**ACKNOWLEDGMENT**

This work is supported by the Korea Ministry of Science and Technology.