Preprints of the 22nd IFAC World Congress
Yokohama, Japan, July 9-14, 2023

Complex nonlinear system modeling using type-2 fuzzy deep learning ⋆

Francisco Vega * **Wen Yu** * **Xiaoou Li** **

* *Departamento de Control Automatico, CINVESTAV-IPN (National Polytechnic Institute), Mexico City, Mexico (e-mail: yuw@ctrl.cinvestav.mx)*
** *Departamento de Computacion, CINVESTAV-IPN (National Polytechnic Institute), Mexico City, Mexico (e-mail: lixo@cs.cinvestav.mx)*

Abstract: Type-2 fuzzy systems have a great adoption in different branches of engineering, due to the fact that this type of fuzzy systems are very well suited to tasks related to nonlinear systems. Data driven models like neural networks and fuzzy systems have some disadvantages, such as the high and uncertain dimensions and complex learning process. In this paper, we show the advantages of type-2 fuzzy systems over type-1 fuzzy systems in modeling nonlinear systems. We combine Type-2 Takagi-Sugeno fuzzy model with the popular deep learning model, LSTM (long-short term memory), to overcome the disadvantages fuzzy model and neural network model. We propose a fast and stable learning algorithm for this model. Comparisons with others similar black-box and grey-box models are made, in order to show the advantages of the type-2 fuzzy LSTM neural networks.

Keywords: deep learning, LSTM, type-2 fuzzy systems, modeling

1. INTRODUCTION

Fuzzy system (FS) and fuzzy neural networks (FNN) have had a really extensive development, since their structure is interpreted as a set of "IF-THEN" rules that are easy to understand. Their theory has arrived at the distinction of two main types of FSs, type-1 FSs (T1FSs) and type-2 FSs (T2FSs). T2FSs are considered extensions of the T1FSs, because the membership value of a type-2 fuzzy set is a type-1 fuzzy number. The membership functions (MFs) of T1FSs are well determined, while the MFs of T2FSs are fuzzy; there are infinitely type-1 MFs (T1MFs) contained in the uncertainty footprint characteristic of type-2 MFs (T2MFs). T2FSs have proven to be better at handling data with uncertainties and noise, as evidenced by the work discussed below. The structure of a FNN, which is known as a self-evolving interval type-2 fuzzy neural network (ST2FNN), is shown in Juang (2008). The ST2FNN is constructed as a Takagi-Sugeno-Kang (TSK) FS with an adaptive structure, where the part of the antecedents in the fuzzy rules is defined with T2MFs. In addition, the ST2FNN was used in the modeling of nonlinear systems, adaptive noise cancelation and prediction of chaotic signals, obtaining good results. A study presented in Aliev (2011) discusses the concept of type-2 fuzzy inference system (FIS) using type-2 fuzzy sets and FNN based refinement. A category of type-2 FNNs was developed based on type-2 fuzzy set constructions with these fuzzy sets forming a collection of IF-THEN rules.

The application of type-2 fuzzy neural networks (T2FNNs) is reviewed in Tavoosi (2021). Different T2FNNs that have been used for system identification are discussed, their disadvantages and advantages are described, as well as their effectiveness in different applications. The previous works show the advantages of using T2FSs over T1FSs, as well as the value of the former in the estimation of nonlinear models. Motivated by the above, in this paper we modified the FNN shown in Yu (2020) and Vega (2020) to boost its performance. The FNN is based on a T1FS for the estimation of nonlinear systems, obtaining favorable results with it, so adapting its structure to a T2FS will improve its performance. In the Fig. 1, we show a Gaussian T2MF with uncertain standard deviation, which is upper bounded by a T1MF (UMF) and lower bounded by another T1MF (LMF) and the gray area is the footprint of uncertainty (FOU). The T2FSs cope with the uncertainties of a system from the fuzzy rules that define it. Unlike the T1MFs present in the T1FSs, the T2MFs of the T2FSs are themselves fuzzy because they are defined in their respective FOU. There are an infinite T1MFs in a FOU, which is why T2FSs have the ability to work with data that have uncertainties and noise in a more efficient way.

We combine the FNN structure known as a type-2 Takagi-Sugeno fuzzy neural network with long-short memory term cells, and propose a new model Type-2 fuzzy LSTM (T2LSTM) and a training algorithm for it. Then, we compare it with the type-1 Takagi-Sugeno fuzzy neural network with long-short memory term cells (T1LSTM). T1LSTM and T2LSTM were tested in the identification

⋆ This work was supported in part by CONACYT under the grant CONACyT-A1-S-8216

and control of nonlinear systems, in order to observe the advantages of the T2FS implementation.

2. T2LSTM STRUCTURE

The T2LSTM is obtained from the NARMA model of a nonlinear system:

$$\begin{aligned} y(k) &= \varphi\left[U_r(k)\right] \\ U_r(k) &= \left[(k-1), \cdots, y(k-n_y), u(k), \cdots \right. \\ &\left. \cdots, u(k-n_u)\right]^T = \left[u_{r_1} \cdots u_{r_m}\right]^T \end{aligned} \tag{1}$$

where $y(k)$ is the variable of interest of the system under study (output signal), $\varphi(\cdot)$ is an unknown nonlinear difference equation, $U_r(k)$ is the state vector with $u(k)$ and $y(k)$ with the former as the input signal for the system; n_y indicates the number of the delayed output signal, n_u indicates the number of the delayed input signal, and m indicates the number of elements u_{r_m} in $U_r(k)$.

To model (1), we used fuzzy IF-THEN rules. For the p-th rule it has:

$$\begin{aligned} R_p : &\text{IF } u_{r_1}(k) \text{ IS } A_{1p} \;\&\; u_{r_2}(k) \text{ IS } A_{2p} \;\&\; \cdots \\ &\cdots \;\&\; u_{r_m}(k) \text{ IS } A_{jp}, \text{ THEN } h_p(k) = \varrho_p(k) \end{aligned} \tag{2}$$

where $h_p(k)$ is an estimation to the function $\varrho_p(k)$ that represents the consequent part of each fuzzy rule. The sets A_{jp}, with $j = 1 \ldots \kappa$, are the fuzzy sets for the fuzzification (using κ fuzzy sets) of each u_{r_m} in (2).

The T2MFs associated with each A_{jp} are of the form shown in Fig. 1 and are described as follows, first the UMF associate to each A_{jp}:

$$\overline{\mu}_{A_{jp},u_{r_m}}(k) = \exp\left(-\frac{\left(u_{r_m}(k) - \varsigma_{jp}\right)^2}{2\overline{\nu}_{jp}}\right) \tag{3}$$

and the LMF:

$$\underline{\mu}_{A_{jp},u_{r_m}}(k) = \exp\left(-\frac{\left(u_{r_m}(k) - \varsigma_{jp}\right)^2}{2\underline{\nu}_{jp}}\right) \tag{4}$$

in these Gaussian functions, the center is $\varsigma_{jp} \in \mathbb{R}$ and the widths are $\overline{\nu}_{jp}, \underline{\nu}_{jp} \in \mathbb{R}^+$. For the estimation of (1), the contribution of each input element to the premise part of a fuzzy rule in (2) is obtained by the T-norm:

$$\overline{z}_p(k) = \prod_{j=1}^{\kappa} \overline{\mu}_{A_{jp},u_{r_m}}(k)$$

$$\underline{z}_p(k) = \prod_{j=1}^{\kappa} \underline{\mu}_{A_{jp},u_{r_m}}(k)$$

assuming $j = m$.

The vectorial representation of the value of each element of (3) and (4) in each fuzzy set is:

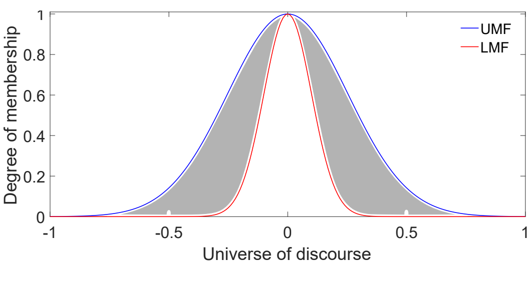


Fig. 1. Type-2 Gaussian MF, bounded by UMF and LMF, with uncertain standard deviation.

$$\overline{\zeta}_j = \exp\left[\left(U_r(k) - \chi_j\right)^2 \otimes \left(-\frac{1}{2}\overline{\Upsilon}_j\right)\right] \tag{5}$$

$$\underline{\zeta}_j = \exp\left[\left(U_r(k) - \chi_j\right)^2 \otimes \left(-\frac{1}{2}\underline{\Upsilon}_j\right)\right] \tag{6}$$

with $\chi_j, \overline{\Upsilon}_j \underline{\Upsilon}_j \in \mathbb{R}^m$ as the center and width vectors for $\overline{\zeta}_j, \underline{\zeta}_j \in \mathbb{R}^m$, respectively.

The vectors $\overline{\zeta}_j, \underline{\zeta}_j$ represent the value of each element of $u_{r_m}(k)$ in a fuzzy set A_j, and $\otimes$ is the operator for the element to element product in vectors. This representation will be useful for the adjustment of the parameters of the FNN.

The estimation of (1) is obtained by the defuzzification of the FS (2) with p rules:

$$\begin{aligned} \hat{y}(k) &= \frac{\sum_{n=1}^{p} \overline{z}_n h_n(k)}{\sum_{n=1}^{p} \overline{z}_n} + (1-\beta)\frac{\sum_{n=1}^{p} \underline{z}_n h_n(k)}{\sum_{n=1}^{p} \underline{z}_n} \\ &= \sum_{n=1}^{p} \overline{\hat{z}}_n h_n(k) + (1-\beta)\sum_{n=1}^{p} \underline{\hat{z}}_n h_n(k) \end{aligned} \tag{7}$$

where:

$$\overline{\hat{z}}_p = \overline{z}_p / \left(\overline{z}_1 + \overline{z}_2 + \cdots + \overline{z}_n\right)$$
$$\underline{\hat{z}}_p = \underline{z}_p / \left(\underline{z}_1 + \underline{z}_2 + \cdots + \underline{z}_n\right)$$

In (7), the premise part can be represented in a vectorial way as $\overline{Z}_F, \underline{Z}_F \in \mathbb{R}^p$, where all the elements of this new vector are the organized multiplications as was explained in (3) and (4). Also, each element of $\overline{Z}_F$ and $\underline{Z}_F$ is normalized. Acording with the theory of T2FSs, the design parameter, β, weights the sharing of the lower and the upper firing levels of each fired rule, this parameter can be a constant and in this paper we take $\beta = 0.5$.

For multiple estimations, the elements of $\overline{Z}_F$ and $\underline{Z}_F$ can be organized in such a way that the premise parts repeats for every estimation, hence the consequent parts are the only ones that are different for several estimation in a same system.

The $h_n(k)$ elements in (7) are calculate by LSTM cells. The cells that we used in the T2LSTM have several parts, which are describe as follows:

$$F(k) = \sigma\left(W^f U_r(k) + V^f H(k-1)\right) \tag{8}$$
$$I(k) = \sigma\left(W^i U_r(k) + V^i H(k-1)\right) \tag{9}$$
$$S(k) = \psi\left(W^s U_r(k) + V^s H(k-1)\right) \tag{10}$$
$$C(k) = F(k) \otimes C(k-1) + I(k) \otimes S(k) \tag{11}$$
$$O(k) = \sigma\left(W^o U_r(k) + V^o H(k-1)\right) \tag{12}$$
$$H(k) = O(k) \otimes \psi\left(C(k)\right) \tag{13}$$

where $F(k)$, $I(k)$, $S(k)$, $C(k)$, $O(k)$ and $H(k) \in \mathbb{R}^p$ are the fitness of the internal state, the fitness of the internal input, the internal input, the internal state, the fitness of the output, and he output of the cells, respectively. The synaptic weights are: W^f, W^i, W^s and $W^o \in \mathbb{R}^{p\times m}$; V^f, V^i, V^s and $V^o \in \mathbb{R}^{p\times p}$ as diagonal matrices or V^f, V^i, V^s and $V^o \in \mathbb{R}^p$ as vectors, according to the need. The functions $\sigma(\cdot)$ and $\psi(\cdot)$ are the sigmoid and hyperbolic tangent functions, respectively, $U_r(k) \in \mathbb{R}^m$ is the input in (1).

From (7), the output of the FS is

$$\hat{y}(k) = \left[\overline{Z}_F + (1-\beta)\underline{Z}_F\right] H(k) \tag{14}$$

where $H(k) = [h_1(k) \cdots h_p(k)]^T$ corresponds to the consequent parts and $\overline{Z}_F, \underline{Z}_F \in \mathbb{R}^n$ are the elements of the premise parts of the rules described by (2).

The cells, as well as the number of rules, are defined as $p = \kappa^m$ in the case of one estimation; for several estimations it has $p = l(\kappa^m)$ where l is the number of estimations, thus $\hat{y} \in \mathbb{R}^l$, as was described for (7). The complete structure of the T2LSTM is contained in (5)-(14).

By applying the function approximation theories of FSs, (1) can be represented as:

$$y(k) = \left[\overline{Z}_F(W^*) + \cdots \right.$$
$$\left. \cdots + (1-\beta)\underline{Z}_F(W^*)\right] H(W^*) + \mu(k) \quad (15)$$

where W^* corresponds to the unknown weights which can minimize the unmodeled dynamic $\mu(k)$.

We assume that (1) is bounded-input and bounded-output (BIBO) stable, *i.e.*, $y(k)$ and $U_r(k)$ in (1) are bounded. By the bound of the membership functions (3) and (4), $\mu(k)$ in (15) is bounded.

Remark 1. It can be noted that the main difference between the T2LSTM and the T1LSTM lies in (14), since the latter equation is defined for a T1LSTM as follows:

$$\hat{y}(k) = Z_F H(k)$$

for the T1FS, the premise part of the fuzzy rules, Z_F, is simpler than in the T2FS case. This largely due to the use of T2MFs instead of T1MFs, which as explained above, offers higher robustness. But on the other hand, it increases the computational complexity as shown later in the paper.

3. COMPARISONS

We employed two exercises to compare the new T2LSTM with a similar T1LSTM, taking into account the remark 1. We talk about the performance of the T2FS, the exercises exhibit the advantages of the T2FS over the T1FS. Both exercises were handled as applications in real time under the same conditions for the algorithms.

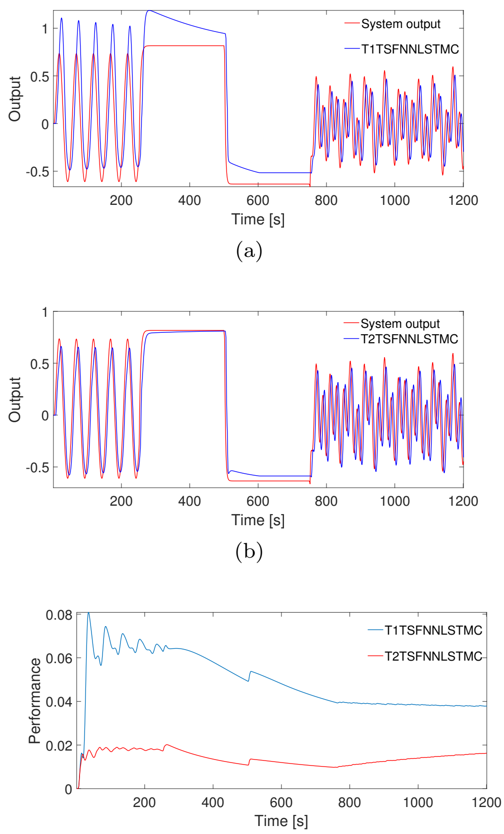
The first exercise consist on a model generation for a nonlinear system, which is defined as:

$$y(k+1) = 0.72y(k) + 0.025y(k-1)u(k-1) +$$
$$+ 0.01u^2(k-2) + 0.2u(k-3) \quad (16)$$

$$u(k) = \begin{cases} \sin\left(\frac{\pi k}{25}\right), & k \le 250 \\ 1, & 250 < k \le 500 \\ -1, & 500 < k \le 750 \\ 0.3\sin\left(\frac{\pi k}{25}\right) + 0.1\sin\left(\frac{\pi k}{32}\right) + & \\ +0.6\sin\left(\frac{\pi k}{10}\right), & 750 < k \end{cases}$$

with $y(0) = 0$, $u(0) = 0$, and a sampling time of $T = 1$s.

So, (16) was solved for 1,200s and the system was sampled, creating the vector $y(k)$ with $k = 1, \ldots, 1,201$. We took the values of $y(k)$ to define $U_r(k) = [y(k-4), y(-8)]^T$, which was used to make the $\hat{y}(k)$ estimate. We employed the first 601 iterations to train the FNNs, meanwhile the rest data were used for testing the algorithms.



(a) (b) (c)

Fig. 2. Nonlinear system identification with tne FNNs: (a) T1TSFNNLSTMC, (b) T2TSFNNLSTMC, (c) Performance comparison.

We set $p = 9$ fuzzy rules for the FNNs ($m = 2$, $\kappa = 3$, $l = 1$). Here, the modeling error $E(k)$ at the end of each phase is defined as in (**??**) and it represents the performance of the algorithms, a low value indicates a better performance. We used (**??**) for the training of the FNNs because is the error that we want to minimize.

The comparison results are shown in the Table 1 and the Fig.2, the latter has three parts: (a) shows the system output and the output of the T1LSTM, (b) shows the system output and the output of the T2LSTM, and (c) shows a performance comparison between the T1LSTM and the T2LSTM.

It can be seen that both FNNs have similar behavior, however, the T2LSTM learned the system dynamics faster compared to the T1LSTM, which is why the former has a smaller modeling error or better performance during the whole process.

Table 1. Performance of the FNNs during the identification ($\times 10^{-2}$)

System	Training	Testing
T1TSFNNLSTMC	4.80	3.78
T2TSFNNLSTMC	1.19	1.62

4. CONCLUSIONS

In this paper, we use the advantages of type-2 fuzzy system and LSTM neural networks to design a novel model for nonlinear system modeling. The main advantage is this model can reduce the modeling error compared with the other fuzzy neural networks. However, the cost of the computational complexity of the algorithm is more than the others. This disadvantage can be reduced by current computational technology, which allows the easy application of this model.

REFERENCES

P. Singh, A brief review of modeling approaches based on fuzzy time series, International Journal of Machine Learning and Cybernetics, vol. 8, no. 2, pp. 397-420.

J. Qiu, H. Gao, and S. X. Ding, Recent Advances on Fuzzy-Model-Based Nonlinear Networked Control Systems: A Survey, IEEE Transactions on Industrial Electronics, vol. 63, no. 2, pp. 1207-1217, Feb. 2016.

Z. Pezeshki and S. M. Mazinani, Comparison of artificial neural networks, fuzzy logic and neuro fuzzy for predicting optimization of building thermal consumption: a survey, Artificial Intelligence Review, vol. 52, no. 1, pp. 495-525

J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling , arXiv preprint arXiv:1412.3555.

H. Salehinejad, S. Sankar, J. Barfett, E. Colak, and S. Valaee, Recent advances in recurrent neural networks , arXiv preprint arXiv:1801.01078, 2017.

S. Hochreiter and J. Schmidhuber, Long short-term memory , Neural computation, vol. 9, no. 8, pp. 1735-1780, 1997.

O. Ogunmolu, X. Gu, S. Jiang, and N. Gans, Nonlinear systems identification using deep dynamic neural networks , arXiv preprint arXiv:1610.01439, 2016.

Y. Wang, A new concept using lstm neural networks for dynamic system identification , in 2017 American Control Conference (ACC). IEEE, 2017, pp. 5324 -5329.

F. Nicola, Y. Fujimoto, and R. Oboe, A lstm neural network applied to mobile robots path planning , in 2018 IEEE 16th International Conference on Industrial Informatics (INDIN). IEEE, 2018, pp.349 -354.

Y. Liu, Y. Zhou, and X. Li, Attitude estimation of unmanned aerial vehicle based on lstm neural network , in 2018 International Joint Conference on Neural Networks (IJCNN). IEEE, 2018, pp. 1-6.

Chia-Feng Juang and Yu-Wei Tsao, A Self-Evolving Interval Type-2 Fuzzy Neural Network With Online Structure and Parameter Learning, EEE Transactions on Fuzzy Systems, vol. 16, no. 6, pp. 1411-1424, Dec. 2008.

R. A. Aliev et al., Type-2 fuzzy neural networks with fuzzy clustering and differential evolution optimization, nformation Sciences, vol. 181, no. 9, pp. 1591-1608, May 2011.

J.-Y. Chang, Y.-Y. Lin, M.-F. Han, and C.-T. Lin, A functional-link based interval type-2 compensatory fuzzy neural network for nonlinear system modeling, 2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011), Jun. 2011.

Yang-Yin Lin, Jyh-Yeong Chang, N. R. Pal, and Chin-Teng Lin, A Mutually Recurrent Interval Type-2 Neural Fuzzy System (MRIT2NFS) With Self-Evolving Structure and Parameters, IEEE Transactions on Fuzzy Systems, vol. 21, no. 3, pp. 492-509, Jun. 2013.

E. Kayacan, E. Kayacan, and M. Ahmadieh Khanesar, Identification of Nonlinear Dynamic Systems Using Type-2 Fuzzy Neural Networks: A Novel Learning Algorithm and a Comparative Study, IEEE Transactions on Industrial Electronics, vol. 62, no. 3, pp. 1716-1724, Mar. 2015.

H.-G. Han, Z.-Y. Chen, H.-X. Liu, and J.-F. Qiao, A self-organizing interval Type-2 fuzzy-neural-network for modeling nonlinear systems, Neurocomputing, vol. 290, pp. 196-207, May 2018.

W. Yuan and L. Chao, Online Evolving Interval Type-2 Intuitionistic Fuzzy LSTM-Neural Networks for Regression Problems, IEEE Access, vol. 7, pp. 35544-35555, 2019.

M. A. Cunha Ferreira, R. Tanscheit, and M. Vellasco, Automatic Generation of a Type-2 Fuzzy System for Time Series Forecast based on Genetic Programming, Proceedings of the 2019 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology (EUSFLAT 2019), 2019.

N. F. Rahim, M. Othman, R. Sokkalingam, and E. Abdul Kadir, Type 2 Fuzzy Inference-Based Time Series Model, Symmetry, vol. 11, no. 11, p. 1340, Oct. 2019.

M. Bilgili, A. Yildirim, A. Ozbek, K. Celebi, and F. Ekinci, Long short-term memory (LSTM) neural network and adaptive neuro-fuzzy inference system (ANFIS) approach in modeling renewable electricity generation forecasting, International Journal of Green Energy, vol. 18, no. 6, pp. 578-594, Dec. 2020.

J. Tavoosi, A. A. Suratgar, M. B. Menhaj, A. Mosavi, A. Mohammadzadeh, and E. Ranjbar, Modeling Renewable Energy Systems by a Self-Evolving Nonlinear Consequent Part Recurrent Type-2 Fuzzy System for Power Prediction, Sustainability, vol. 13, no. 6, p. 3301, Mar. 2021.

A. Safari, R. Hosseini, and M. Mazinani, A novel deep interval type-2 fuzzy LSTM (DIT2FLSTM) model applied to COVID-19 pandemic time-series prediction, Journal of Biomedical Informatics, vol. 123, p. 103920, Nov. 2021.

J. Tavoosi, A. Mohammadzadeh, and K. Jermsittiparsert, A review on type-2 fuzzy neural networks for system identification, Soft Computing, vol. 25, no. 10, pp. 7197-7212, Mar. 2021.

W. Yu and F. Vega, Nonlinear system modeling using the takagi-sugeno fuzzy model and long-short term memory cells, Journal of Intelligent and Fuzzy Systems, vol. 39, no. 3, pp. 4547-4556, Oct. 2020.

F.Vega, W.Yu, Fuzzy Modeling Using LSTM Cells for Nonlinear Systems, IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Glasgow, UK, July 19-24, 2020