Abstract—This paper proposes an improved version of the Self-constructing Neural Fuzzy Inference Network (SONFIN) [1], called Soft-Boosted SONFIN (SB-SONFIN). The design softly boosts the learning process of the SONFIN in order to decrease the error rate and enhance the learning speed. The SB-SONFIN boosts the learning power of the SONFIN by taking into account the numbers of fuzzy rules and initial weights which are two important parameters of the SONFIN, SB-SONFIN advances the learning process by, (1) initializing the weights with the width of the fuzzy sets rather than just with random values; (2) improving the parameter learning rates with the number of learned fuzzy rules. The effectiveness of the proposed soft boosting scheme is validated on several real world and benchmark datasets. The experimental results show that the SB-SONFIN possesses the capability to outperform other known methods on various datasets.


I. INTRODUCTION

The online self-constructing neural fuzzy inference network (SONFIN) [1] features a high learning speed and small network size, and it has been applied to various problems producing promising results [2-11]. The SONFIN has a 6-layer network structure which cognizes a fuzzy model with structure and parameter learning process. The details of each layer’s function and learning procedure are described in [1]. The width of each fuzzy set, the number of fuzzy rules generated on-line, and the initial weights of the nodes are the most sensitive parameters in the SONFIN and play important role in the network structure. Taking into consideration these characteristics of the SONFIN, this paper proposes its enhanced variant, called SB-SONFIN, which boosts its online learning process and provide a compact and efficient model. The SB-SONFIN possesses the ability of neural networks through the Takagi–Sugeno–Kang (TSK)-type fuzzy rule-based model. Fuzzy neural network based adaptive control systems [22-30] generates numbers of rules online and adapt them by the system procedure based on given input parameters and datasets. The generated numbers of rules and given parameters play an important role in the performance of system modeling and compactly work together via synchronous structure and parameter learning. The network structure of SB-SONFIN is based on the structure of SONFIN as shown in Fig. 1.

The explanatory sections of this paper are organized as follows: Section II presents the proposed method called SB-SONFIN; Section III outlines the experimental results; and finally, Section IV provides the conclusions and future work.

II. SB-SONFIN

The SB-SONFIN boosts the learning process of the SONFIN by incorporating novel weight initialization procedure adapted along with the number of fuzzy rules generated...
generated during the structure learning. The SB-SONFIN focuses primarily on adjustment of the number of its rules and parameters, which not only boosts the learning process but also gives an effective and stable system. The SB-SONFIN produces an effective network structure and speeds up the learning process compared to existing methods as shown in the experimental results. Fig. 2 shows the structure learning flow diagram of the SB-SONFIN, where the red rectangles indicate the updated parts of the original learning algorithm employed in the SONFIN. The SB-SONFIN enhances the learning power of the SONFIN in two ways: (1) it initializes the weights with width of the fuzzy sets rather than just with random values; and (2) it softly boosts the parameter learning rates with the number of learned fuzzy rules. These two ways can be formally described as follows:

In layer 2 each node defines a Gaussian membership function as follows:

\[ \mu_j(x_j) = \exp \left\{ -\frac{1}{2} \left( \frac{x_j - m_j}{\sigma_j} \right)^2 \right\} \]  

(1)

where \( m_j \) and \( \sigma_j \) denote the center and width of the \( j \)-th fuzzy set \( A_j \) in the \( j \)-th input variable \( x_j \) (\( j=1,2,\ldots,d \)), respectively.

For each incoming data pattern \( \tilde{x} \), the firing strength \( F^i(\tilde{x}) \), computed as in Eq. (2), is used as the criterion to decide whether a new fuzzy rule should be generated or not. Initially, there are no rules and new rule is generated for the first incoming data \( \tilde{x}(1) \) with the center and widths of the fuzzy set determined using Eq. (3).

\[ F^i = \prod_{j=1}^{d} \mu_j^i, i = 1,\ldots,r \]  

(2)

\( m_j^i = x_j(1), \sigma_j^i = \sigma_m \) and \( w_j^i = \sigma_j^i; j = 1,2,\ldots,d \) where \( \sigma_m \) is a pre-specified value that determines the width of each fuzzy set and \( w_j^i \) is the initial weight initialized with the width of the fuzzy set. For subsequent incoming pattern \( \tilde{x}(t) \), the set

\[ I = \arg \max_{1 \leq r \leq F^i(\tilde{x}(t))} F^i(\tilde{x}(t)) \]  

(4)

where \( r(t) \) is the number of rules existing at time \( t \). If \( F^i(\tilde{x}(t)) \leq F_m \), then a new fuzzy rule is generated and \( r(t+1)=r(t)+1 \). After generating new fuzzy rule, the next step is to generate a new fuzzy set on each input variable. A new fuzzy set is generated in the \( j \)-th input variable, and \( h_j(t+1) \) is set equal to \( h_j(t)+1 \). The center and width of the fuzzy sets are set as shown in eq. (5), whereas, the weights are updated with the width of fuzzy sets as shown in Eq. (6).

\[ m_j^{h_j(t+1)} = x_j(t), \sigma_j^{h_j(t+1)} = \beta | x_j(t) - m_j^i | \]  

(5)

\[ w_j^{h_j(t+1)} = \sigma_j^{h_j(t+1)}; j = 1,2,\ldots,d; h_j(t+1) = h_j(t) + 1 \]  

(6)

where \( \beta \) is overlap coefficient and \( h_j(t) \) is number of fuzzy sets in the \( j \)-th input variable, \( x_j \).

Once the structure learning phase based on current training patterns is adjusted, the next parameter learning phase adjusts the parameters of network optimally according to the same training patterns. All the free parameters of SB-SONFIN are tuned, no matter whether the rules are newly generated or already existing. The main goal of proposed approach is to minimize the error function that can be described as follows:

\[ E = \frac{1}{2} \| y(t) - y^d(t) \|^2 \]  

(7)

Fig. 2. Learning flow diagram of SB-SONFIN
where \( y(t) \) is the actual output and \( y'(t) \) is the desired output. For each training pattern starting at input node, a forward pass is used to compute the activity levels of all the nodes in the network in order to obtain the current output \( y(t) \). When starting at output node, a backward pass is used to compute \( \frac{\partial E}{\partial \omega} \) for all the hidden nodes, where \( \omega \) represents an adjustable parameter in the node.

In general, the learning rate is constant with increasing number of fuzzy rules for each new coming pattern. Taking into consideration the dynamic nature of the system, the learning rate should also change with fuzzy rules in order to allow the system to learn faster and take bigger steps to provide accurate results in specific time period. This is the reason why the SB-SONFIN chooses to multiply the fuzzy rule \( r \) with the learning rate \( \eta \) in order to soft-boost the learning process as described in Eq. (10). In order to provide the initial boosting power to the learning process, the adopted general update rule used can be described as follows:

\[
\Delta \omega \propto -\frac{\partial E}{\partial \omega}
\]

\[
\eta(t+1) = \eta(t) + \eta \left( -\frac{\partial E}{\partial \omega} \right)
\]

where \( E \) is the error function, \( \eta \) is the learning rate defined by user, and \( r \) is the number of fuzzy rules generated during the structure learning phase. In order to obtain more details, an interested reader is referred to the parameter learning section of the SONFIN [1].

The effect of soft boosting can be seen during the learning process of the system. Fig. 3-4 show the learning rate of the SB-SONFIN compared with the learning rate of the SONFIN. The soft boosting power present during the learning phase is shown in Fig. 3-4 visualized using Mackey glass time series prediction problem and nonlinear dynamic system identification problem. The experimental results indicate that SB-SONFIN outperforms SONFIN on given problems generating less or equal number of fuzzy rules.

### III. EXPERIMENTAL RESULTS

The performance of SB-SONFIN is demonstrated on Mackey glass time series prediction problem, nonlinear dynamic system identification problems and ten other benchmark datasets from the UCI repository [21]. The experiments, whose results are analysed in this section, were performed in MATLAB 7.9 environment and Intel i5 3.1 GHZ CPU with 4GB RAM running on Windows XP 7(32-bit). The results of the SB-SONFIN are compared with those of other on-line and dynamic models, such as SONFIN [1], GEFB-OFSNN [12], RBF-AFS [13], OLS [14], DFNN [15], FAOS-PFNN [16], Mean-Shift method [17], KNN method [17], Space partitioning method [17], Khayat’s model [18], SOFNNGA [19] and SOFNN [20]. The performance evaluation of the SB-SONFIN is done in terms of number of rules and root mean square error (RMSE) during the training and testing phase, whereas, Smaller RMSE and less numbers of rules indicate better performance. The parameter settings adopted in all the reported experiments are mentioned in Table I.
A. Mackey Glass Time Series Prediction (MGTP)

In order to verify the SB-SONFIN capabilities, we chose to employ one of the classical benchmarks, namely, chaotic Mackey glass time series prediction problems [1, 12-16] having the discrete model of time series as follow:

\[ x(t + 1) = (1 - a) x(t) + \frac{b x(t - \tau)}{1 + x(t - \tau)^m} \]

(11)

where \( a = 0.1, b = 0.2, \tau = 17 \) and \( x(0) = 1.2 \). The problem is to predict the value \( x(t+p) \) from the prediction model when \( p = 6 \).

\[ x(t + p) = f \{ x(t), x(t - 6), x(t - 12), x(t - 18) \} \]

(12)

We took a set of 1000 data for each training and testing purpose. Table II shows the performance comparison of the SB-SONFIN with SONFIN [1], GEBF-OSFNN [12], RBF-AFS [13], OLS [14], DFNN [15] and FAOS-PFNN [16].

B. Nonlinear Dynamic System Identification Problem (NDSIP)

NDSIP is another benchmark nonlinear dynamic system identification problem [17-20] that can be described as follows:

\[ y(t + 1) = \frac{y(t) y(t - 1)}{1 + y^2(t) + y^2(t - 1)} + \mu(t) \]

(13)

If a series-parallel identification model is used to identify the NDSIP, the model can be described as follows:

\[ \hat{y}(t + 1) = f \{ y(t), y(t - 1), \mu(t) \} \]

(14)

where \( \mu(t) = \sin(2\pi t/25) \) is input, \( y(t+1) \) is output and the network contains three inputs and one output. The initial input values \( y(0) = 0 \) and \( y(1) = 0 \) is used.

In our experiments, we generated a set of 200 data for each training and testing dataset. Table III shows the performance comparison of the SB-SONFIN with the SONFIN [1], Mean-Shift method [17], KNN method [17], Space partitioning method [17], Khayat’s model [18], SOFNGA [19], and SOFNN [20]. SOFNGA KNN method, SONFIN and Khayat’s model have the same number of rules but their training and testing RMSE is higher than the RMSE of SB-SONFIN. The SONFIN, Mean-shift method and SOFNN has the same number of rules as the proposed system but in terms of training and testing RMSE, the performance of the SB-SONFIN is superior.

C. Benchmark Data from UCI

The SB-SONFIN was also tested on 10 different benchmark datasets collected from the UCI repository [21] and normalized to fit the [0 1] range. For all benchmark datasets, 80% of the data are taken for training and 20% of the data are chosen for testing purpose. Table IV shows the performance of SB-SONFIN in comparison with the performance of SONFIN. The SB-SONFIN achieves better performance than the SONFIN while generating similar or smaller number of rules on all given 10 dataset.


IV. CONCLUSIONS

The main goal of the proposed approach is to construct a model that enhances the learning process of the SONFIN system. Elevating the idea, the SB-SONFIN improved the SONFIN performance through processing a new way to initialize the weight vector and redefining the learning parameters. Since the initial weights and generated number of rules play an important role in all on-line modeling, the SB-SONFIN uses the width of the fuzzy sets to initialize the weights during the on-line structure learning process for each newly generated rule. The performance of the SB-SONFIN is compared with existing methods in terms of the RMSE of the testing phase and the number of generated fuzzy rules. Taking into account the obtained experimental results; it can be easily observed that the SB-SONFIN outperforms the other models along with the SONFIN on all the adopted datasets.

ACKNOWLEDGMENT

The authors would like to thank editors and reviewers for many useful comments and suggestions, which significantly helped to improve the presentation quality of the paper.

REFERENCES


### TABLE IV

<table>
<thead>
<tr>
<th>Method</th>
<th>Training RMSE</th>
<th>Testing RMSE</th>
<th>No. of Rules</th>
<th>Training RMSE</th>
<th>Testing RMSE</th>
<th>No. of Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yatch</td>
<td>0.0103</td>
<td>0.0158</td>
<td>5</td>
<td>0.0079</td>
<td>0.0138</td>
<td>5</td>
</tr>
<tr>
<td>Slump test</td>
<td>0.0667</td>
<td>0.0824</td>
<td>3</td>
<td>0.0463</td>
<td>0.0480</td>
<td>3</td>
</tr>
<tr>
<td>Concrete comp.</td>
<td>0.0857</td>
<td>0.0856</td>
<td>4</td>
<td>0.0617</td>
<td>0.0584</td>
<td>4</td>
</tr>
<tr>
<td>Servo</td>
<td>0.0676</td>
<td>0.0535</td>
<td>4</td>
<td>0.0334</td>
<td>0.0446</td>
<td>4</td>
</tr>
<tr>
<td>Airfoil</td>
<td>0.0799</td>
<td>0.1111</td>
<td>7</td>
<td>0.0643</td>
<td>0.0883</td>
<td>7</td>
</tr>
<tr>
<td>Housing</td>
<td>0.0506</td>
<td>0.0990</td>
<td>5</td>
<td>0.0460</td>
<td>0.0883</td>
<td>5</td>
</tr>
<tr>
<td>Bike sharing day</td>
<td>0.0065</td>
<td>0.0091</td>
<td>5</td>
<td>0.0009</td>
<td>0.0005</td>
<td>3</td>
</tr>
<tr>
<td>Energy efficiency</td>
<td>0.0269</td>
<td>0.0312</td>
<td>12</td>
<td>0.0263</td>
<td>0.0300</td>
<td>7</td>
</tr>
<tr>
<td>Skill Craft</td>
<td>0.0400</td>
<td>0.1068</td>
<td>10</td>
<td>0.0174</td>
<td>0.0879</td>
<td>9</td>
</tr>
<tr>
<td>Parkinson</td>
<td>0.0373</td>
<td>0.0890</td>
<td>11</td>
<td>0.0375</td>
<td>0.0649</td>
<td>6</td>
</tr>
</tbody>
</table>
Mukesh Prasad (M’13) received his Ph.D. degree in computer science from National Chiao Tung University, Hsinchu, Taiwan in 2015 and master degree in computer applications from Jawaharlal Nehru University, New Delhi, India, in 2009. He is currently associated with National Chiao Tung University, Hsinchu, Taiwan as Post-Doctoral Researcher. His current research interests include machine learning, pattern recognition, fuzzy systems, neural networks and wireless sensor networks. Dr. Mukesh has published papers in numerous international journals and conferences including IEEE Transactions, ACM, Elsevier and Springer.

Professor Chin-Teng Lin (F’95) received a B.S. degree from National Chiao-Tung University (NCTU), Hsinchu, Taiwan, in 1986, and Master’s and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1989 and 1992, respectively. He is currently the Provost and a Chair Professor in Department of Electrical and Computer Engineering and the Director of the Brain Research Center, NCTU. He was selected as an IEEE Fellow for his contributions to biologically inspired information systems in 2005 and was later selected as an International Fuzzy Systems Association Fellow in 2013. He is the co-author of Neural Fuzzy Systems (Prentice-Hall) and the author of Neural Fuzzy Control Systems with Structure and Parameter Learning (World Scientific). He has published more than 200 journal papers in the areas of neural networks, fuzzy systems, multimedia hardware/software, and cognitive neuro-engineering, including over 88 IEEE journal papers. Prof. Lin has been the Editor-in-Chief of the IEEE TRANSACTIONS ON FUZZY SYSTEMS since 2011. He served on the Board of Governors for the IEEE Circuits and Systems Society from 2005 to 2008, the IEEE Systems, Man, Cybernetics Society from 2003 to 2005, the IEEE Computational Intelligence Society from 2008 to 2010, and the Chair of IEEE Taipei Section from 2009 to 2010. He was the Distinguished Lecturer of the IEEE CAS Society from 2003 to 2005. He served as the Deputy Editor-in-Chief of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-II from 2006 to 2008. He was the Program Chair of the IEEE International Conference on Systems, Man, and Cybernetics in 2005 and a General Chair of the 2011 IEEE International Conference on Fuzzy Systems.

Dong-Lin Li received his Ph.D. degree in electrical engineering from National Chiao Tung University, Hsinchu, Taiwan in 2013 and master degree in electrical engineering from National Chung Hsing University, Taichung, Taiwan, in 2006. He is currently associated with National Chiao Tung University, Hsinchu, Taiwan as Post-Doctoral Researcher. His current research interests include machine learning, fuzzy systems, neural networks and image processing and smart living technology. Dr. Li has published papers in international journal and conferences including IEEE Transactions, ACM, Elsevier and Springer.

Chao-Tien Hong received his Ph.D., Master and B.S degree in Electronic and Control Engineering from National Chiao Tung University, Hsinchu, Taiwan in 2015, 2007, and 2005 respectively. He is currently associated with Taiwan Semiconductor Manufacturing Company, Hsinchu, Taiwan as Research and Development Engineer. His current research interests include machine learning, pattern recognition, image processing and fuzzy systems, and neural networks. Dr. Chao-Ting has published papers in international journal and conferences including IEEE Transactions, ACM, Elsevier and Springer.

Wei-Ping Ding received the B.S. degree in Computer Science and Technology, Nantong University, Nantong, China, in 2002, the M.S. degree in Software Engineer from Soochow University, Suzhou, China, in 2005, and the Ph.D. degree in Computer Application, Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2013. His current research interests include co-evolutionary algorithms,
granular computing, data mining, machine learning and their applications in medicine. He is an Associate Professor in School of Computer Science and Technology, Nantong University, Nantong, Jiangsu, China. He is a member of Association of Computing Machinery (ACM), IEEE Computer Society (IEEE-CS), and China Computer Federation (CCF). He has authored or co-authored more than 60 papers in journals and conference proceedings.

Dr. Ding was a recipient of National Natural Science Young Foundation of China in 2013. He was awarded an excellent-young teacher of Jiangsu Province sponsored by Qing Lan Project, Jiangsu Province, China, in 2014.

**Professor Jyh-Yeong Chang** (S’84–M’86) received the B. S. degree in control engineering and the M. S. degree in electronic engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1976 and 1980, respectively, and the Ph.D. degree in electrical engineering from North Carolina State University, Raleigh, NC, USA, in 1987. During 1976–1978 and 1980–1982, he was a Research Fellow at the Chung Shan Institute of Science and Technology, Lung-Tan, Taiwan. Since 1987, he has been an Associate Professor with the Department of Electrical Engineering, National Chiao Tung University, where he is currently a Professor. His current research interests include neural fuzzy systems, video processing and surveillance, and e-health systems.