Impact of small-world topology on the performance of a feed-forward artificial neural network based on 2 different real-life problems

Okan ERKAYMAZ1,*, Mahmut ÖZER2, Nejat YUMUŞAK3
1Department of Electronics and Computer Science, Technical Education Faculty, Karabük University, Karabük, Turkey
2Department of Electrical & Electronics Engineering, Zonguldak Karaelmas University, Zonguldak, Turkey
3Department of Computer Engineering, Sakarya University, Sakarya, Turkey

Abstract: Since feed-forward artificial neural networks (FFANNs) are the most widely used models to solve real-life problems, many studies have focused on improving their learning performances by changing the network architecture and learning algorithms. On the other hand, recently, small-world network topology has been shown to meet the characteristics of real-life problems. Therefore, in this study, instead of focusing on the performance of the conventional FFANNs, we investigated how real-life problems can be solved by a FFANN with small-world topology. Therefore, we considered 2 real-life problems: estimating the thermal performance of solar air collectors and predicting the modulus of rupture values of oriented strand boards. We used the FFANN with small-world topology to solve both problems and compared the results with those of a conventional FFANN with zero rewiring. In addition, we investigated whether there was statistically significant difference between the regular FFANN and small-world FFANN model. Our results show that there exists an optimal rewiring number within the small-world topology that warrants the best performance for both problems.

Key words: Small-world network, feed-forward artificial neural network, rewiring, network topology

1. Introduction

Artificial neural networks (ANNs) are computational tools inspired by brain networks, and have been applied to many diverse fields of life such as health, finance, statistics, industry, and cognitive sciences. An ANN is typically composed of interconnected layers of simple processing units operating in parallel within the layers. Each unit in the network represents a real neuron [1], and produces an output signal to the postsynaptic neurons when it receives enough input from the presynaptic neurons. ANNs are characterized by their internal nonlinearity, learning capability, prediction performance, and modular structure. Many ANN models have been proposed with various connection architectures, namely feed-forward, feedback, single-layer, multilayer, and so on. Feed-forward ANNs (FFANNs) are the most popular architecture due to their analytical tractability and effectiveness.

In earlier works, the performance of the FFANN was usually investigated both depending on the hidden structure of the networks, which includes the number of hidden layers, and the number of neurons per hidden layer [2,3], and on the learning algorithms [4]. In the FFANN, a neuron in the next layer receives inputs from all neurons in the preceding layer. This connection topology corresponds to a regular network. Since a
regular network may be biologically questionable due to the brain’s sophisticated anatomical structure, random network topologies have been proposed to bridge the gap between artificial and brain networks. In this context, the small-world network topology, proposed by Watts and Strogatz [5], is one of the best models to reflect the functional connectivity and anatomic structure of the brain [6–13]. Moreover, many recent investigations showed that real-life networks such as the World Wide Web, protein interaction networks, email networks, social networks, and metabolic networks, exhibit the small-world property [14–16]. A small-world network is characterized by 2 parameters: the characteristic path length, which is the average node-to-node distance, and the clustering coefficient, which is the tendency for clustering between the neighbors of a neuron. The small-world network is obtained if the value of the clustering coefficient is high, like in regular networks, and the value of the characteristic path length is small, like in random networks. Therefore, small-world networks have properties given neither in regular nor in random networks [17].

Recently, several, though not many, works have explored the impact of small-world topology on the FFANN’s performance. Simard et al. [17] studied supervised learning in a multilayered FFANN and found that small-world connectivity reduces the learning error and learning time when compared to regular and random networks. Shuzhong et al. [18] compared the performances of small-world neural networks and regular networks constructed by various statistical methods, and reported that the network with small-world topology had higher performances in several aspects compared to the regular networks. It has also been found that Hopfield networks with small-world properties produce far better results in terms of memory storage and generalization abilities than ones with random and regular connectivity [19,20].

In these studies, the performance of small-world–type ANNs was used with a synthetic dataset instead of using a real-life problem dataset. Therefore, our aim in this study is to show how real-life problems can be solved by a FFANN with small-world topology. In this context, we consider 2 real-life problems: estimating the thermal performance of solar air collectors and predicting the modulus of rupture (MOR) values of oriented strand boards (OSBs). We use a FFANN with small-world topology to solve both problems and compare the results with those of a conventional FFANN.

2. The FFANN with small-world topology

Small-world network topology is constructed using regular lattice topology. We adapt the standard small-world algorithm of Watts and Strogatz [21,22] to a regular conventional FFANN topology. The construction process starts with disconnecting a randomly selected link from its end point and rewiring it to a randomly selected neuron in the network. Notably, if the new connection already exists between 2 nodes, we cancel this rewiring and select a new node randomly. This process is continued for up to the number of maximum possible rewirings. A schematic illustration for the rewiring process is shown in Figure 1.

![Figure 1. Small-world rewiring process.](image)

Since a small-world network is characterized by 2 parameters, namely the characteristic path length (L) and the clustering coefficient (C), they should be calculated to determine which topology the network has, i.e.
regular, small-world, or random network. However, the L and C in a FFANN could not be calculated because it has unconnected neurons within the same layer. Hence, we used the local efficiency \( D_{\text{Local}} \) and the global efficiency \( D_{\text{Global}} \) parameters, introduced by Latora et al. [23], which correspond to \( 1/C \) and \( L \), respectively. Therefore, the networks exhibit a small-world property when both the \( D_{\text{Local}} \) and \( D_{\text{Global}} \) parameters are small [23–25]. For a network, the global efficiency is defined as:

\[
D_{\text{Global}} = \frac{1}{E_{\text{Global}}},
\]

\[
E_{\text{Global}} = \frac{1}{N(N-1)} \sum_{i \neq j \in N} \frac{1}{d_{ij}},
\]

where \( N \) is the number of nodes in the network, and \( d_{ij} \) denotes the shortest path length between 2 nodes. The local efficiency of a network is calculated by averaging each individual node’s efficiency, as follows [23–25]:

\[
D_{\text{Local}} = \frac{1}{E_{\text{Local}}},
\]

\[
E_{\text{Local}} = \frac{1}{N} \sum_{i \in N} E(G_i),
\]

\[
E(G_i) = \frac{1}{N_i(N_i-1)} \sum_{k \neq l \in N} \frac{1}{d_{kl}},
\]

where \( N_i \) is the number of neighbor nodes that are connected directly to node \( i \), and \( d_{kl} \) is the shortest path length between the neighboring nodes when node \( i \) is disconnected from them.

We used a bipolar-sigmoidal function to activate each neuron and performed the backpropagation learning algorithm for teaching small-world FFANNs. The backpropagation algorithm aims to determine the minimum output error in weight space using the gradient descent method, where the weights for the output and hidden layer neurons are defined as follows:

For neurons in the output layer:

\[
\delta_k(i) = y_k(1 - y_k)(b - y_k),
\]

\[
\Delta W_{mk}(i) = \alpha y_m \delta_k(i),
\]

\[
W_{mk}(i + 1) = W_{mk}(i) + \Delta W_{mk}(i) + m \Delta W_{mk}(i).
\]

For the neurons in the hidden layers:

\[
\delta_m(i) = y_m(i)(1 - y_m(i)) \sum_{k=1}^{l} \delta_k(i) w_{mk}(i),
\]

\[
\Delta w_{im}(i) = \alpha x_i(i) \delta_m(i),
\]

\[
W_{im}(i + 1) = W_{im}(i) + \Delta W_{im}(i) + m \Delta W_{im}(i),
\]
where $\alpha$ is the learning coefficient, which has a range of $[0, 1]$; $W$ is the synaptic weight; $\Delta W$ is the change in weight; $\delta$ is the derivation of the error; and $m$ is the momentum coefficient. The mean square error value is calculated while the network is being trained, and is defined as:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - yb_i)^2,$$

(4)

where $N$ is the number of data samples, and $y_i$ and $yb_i$ denote the desired output and the network output, respectively. In order to evaluate the performance of the network, we used 3 statistical parameters: the root mean square error (RMSE), mean absolute error (MAE), and the coefficient of determination ($R^2$). These parameters are determined as:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - yb_i|,$$

(5)

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - yb_i)^2},$$

(6)

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - yb_i)^2}{\sum_{i=1}^{N} yb_i^2}.$$  

(7)

In order to obtain reliable results, the value of $R^2$ must be closer to 1 and the mean squared error (MSE), MAE, and RMSE values must be closer to 0. In addition to these statistical parameters, we used the k-fold cross-validation method. This method is commonly used to avoid the over-fitting problem in ANN applications [26]. In a k-fold cross-validation, the data set is split into k approximately equal-size partitions. Each time, one of the k partitions is used as the test set and the remaining k – 1 partitions are used as the training set. The test error and statistical correction parameters ($R^2$, RMSE) of the partition are computed. This process is repeated k times. The overall performance of the model is then found with the average of the partition test errors and the correction parameters ($R^2$, RMSE). The model performance is defined as:

$$[E_T, R^2_T, RMSE_T] = \frac{1}{k} \sum_{i=1}^{k} E_i, R^2_i, RMSE_i.$$  

(8)

3. Results

In this study, we show how real-life problems can be solved by the FFANN with small-world topology. Therefore, we consider 2 real-life problems: estimating the thermal performance of solar air collectors and the prediction MOR values of the OSB. We construct a FFANN with small-world topology, use it to solve these problems, and compare the results with those of a conventional FFANN. The statistical parameters are calculated for both networks. The predicted outputs for each model are compared with the experimental findings.

We carry out the simulations on a 2.4 GHz Quad Core Intel processor with a 4 GB memory. The model is created using Microsoft Visual C# and is compiled with the Microsoft Visual Studio 2010 Editor.
3.1. Estimating the thermal performances of solar air collectors

In this case, we consider the problem of estimating the thermal performance of a solar air collector, as reported by Caner et al. [27]. They constructed 2 types of solar air collectors and examined them experimentally, and they then used a 3-layer ANN (8, 20, and 1 neurons within each layer) to estimate the thermal performances of the solar air collectors. We use their experimental findings and compare our results with those of their regular FFAN network. As a first step, we design a 3-layer FFANN (8, 12, and 1 neurons within each layer), as shown in Figure 2.

We define the same variables as Caner et al. [27], as the inputs and the output. Eight variables are used as inputs: the input and output of the temperatures, $T_i, T_o$; stored water, $T_{sw}$; an ambient temperature of the collector, $T_{amb}$; surface temperature of the collector, $T_s$; solar radiation intensity, $I$; measuring time data, $Time$; and model type number, $Model~type$. One output variable is the performance of the solar air collector, $\eta$.

As a second step, we apply the rewiring process for the regular network in Figure 2 to determine the rewiring range of a small-world network. More than 8 rewirings cannot be applied to the network due to finished unconnected neurons in the network. After each rewiring process, the $D_{Local}$ and $D_{Global}$ parameters are calculated and shown in Figure 3. The increasing of the rewiring number (RN) reduces the value of $D_{Local}$, while it increases the value of $D_{Global}$. A small-world network is obtained when both of the parameter values are small. In this context, Table 1 shows that a small-world network is obtainable for the rewiring range of between 2 and 8.

As a third step, we use the 10-fold cross-validation method. The used dataset, which is scaled by the min-max method in the range of [0–1], consists of 80 samples. It is then split into 10 approximately equal-size partitions, where one of them is used for testing and the remainder are used for training; this process is repeated 10 times. The overall test errors (MAE, MSE) and statistical correction parameters ($R^2$, RMSE) of the model are calculated by averaging the test error and statistical correction parameters of each test partition. We perform 60 experimental trials with a minimum training error (MSE) criterion. The ending error criterion (MSE) of the training process is 0.000001. The results are given in Table 1. As indicated in Table 1, the best performance is obtained for an optimal RN of RN = 6, within the small-world network range.

![Figure 2. The constructed FFANN with zero rewiring for estimating the thermal performances of solar air collectors.](image)

![Figure 3. The change in $D_{Global}$ and $D_{Local}$ with the RN for estimating the thermal performances of solar air collectors.](image)
Table 1. Rewiring-based error analysis results for estimating the thermal performances of solar air collectors.

<table>
<thead>
<tr>
<th>RN</th>
<th>$R^2$</th>
<th>MAE</th>
<th>MSE</th>
<th>RMSE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.959341</td>
<td>0.049025</td>
<td>0.021994</td>
<td>5.922908</td>
</tr>
<tr>
<td>2</td>
<td>0.989063</td>
<td>0.02177</td>
<td>0.004267</td>
<td>2.911202</td>
</tr>
<tr>
<td>4</td>
<td>0.98624</td>
<td>0.025331</td>
<td>0.005196</td>
<td>3.183115</td>
</tr>
<tr>
<td>6</td>
<td>0.994048</td>
<td>0.017879</td>
<td>0.003085</td>
<td>2.297599</td>
</tr>
<tr>
<td>8</td>
<td>0.989019</td>
<td>0.022255</td>
<td>0.00381</td>
<td>2.86277</td>
</tr>
</tbody>
</table>

Table 1 indicates that the best performance is obtained for an optimal RN of RN = 6, within the small-world network range.

Caner et al. [27] reported that their error analysis of their total data revealed the values of 0.9967, 0.0173, and 0.9879 for the statistical parameters of $R^2$, RMSE, and MAE, respectively. Figure 4 shows how the thermal performance of the small-world FFANN for an optimal RN matches the experimental data. Therefore, we conclude that the small-world FFANN improves the performance of a solar air collector for an optimal RN compared to the regular FFANN.

![Figure 4](image)

**Figure 4.** The thermal performance of the small-world FFANN at an optimal RN (RN = 6).

We investigated whether there is significant difference between a regular FFANN and our small-world FFANN model. Thus, we used the independent 2-sample t-test [28]. The t-test was performed with SPSS and the t value and P value were computed using a significance level of 0.01. The obtained results are shown in Table 2.

Table 2. Result of the t-test for the absolute errors.

<table>
<thead>
<tr>
<th></th>
<th>Regular FFANN model</th>
<th>Small-world model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>23.1387</td>
<td>8.5608</td>
</tr>
<tr>
<td>Std. deviation</td>
<td>37.6303</td>
<td>16.4435</td>
</tr>
<tr>
<td>Std. error mean</td>
<td>4.2072</td>
<td>1.8384</td>
</tr>
<tr>
<td>N</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>df</td>
<td>108.11</td>
<td></td>
</tr>
<tr>
<td>T value</td>
<td>3.1751</td>
<td></td>
</tr>
<tr>
<td>P value</td>
<td>0.0020</td>
<td></td>
</tr>
</tbody>
</table>
As seen in Table 2, because $P < 0.01$ (0.0020), the $H_0$ hypothesis is rejected with a significant level of $\alpha = 0.01$, and also it indicates that there is a statistically significant difference between the models. In this context, the small-world network model is statistically better than the regular FFANN model.

3.2. Predicting the MOR values of the OSBs

In this case, we consider the problem of predicting the MOR values of the OSBs that was reported by Yapıcı [29]. He conducted an experimental study as follows: Scots Pine wood ($P. sylvestris$ L.) was used in the production of the OSBs. The strand’s dimension was approximately 80 mm long, 20 mm wide, and 0.7 mm thick. First, the wood strands were dried to a 3% moisture content before an adhesive was sprayed on them for 3 min. Next, an adhesive material without wax, a solid content of 47% liquid phenol-formaldehyde resin, was applied at rates of 3%, 4.5%, and 6%, based on the weight of the oven-dried wood strands.

The press periods and press pressure were 3, 5, and 7 min under 35, 40, and 45 kg/cm$^2$, respectively. The shelling rate was 40% for the core layer and 60% for the face layer, and the density of the boards was aimed at 0.70 g/cm$^3$ density. A total of 27 OSB panels, with dimensions of $56 \times 56 \times 1.2$ cm, were made for the experiments. There were 54 in total and 2 for each. Hand-formed mats were pressed in a hydraulic press. These panels were labeled from 1 to 27. All of the mats were pressed under automatically controlled conditions at $182 \pm 3$ °C. After pressing, the boards were conditioned to a constant weight at 65 ± 5% relative humidity and at a temperature of $20 \pm 2$ °C until they reached a stable weight [30]. Ten samples were taken from the boards to perform the MOR values according to the related standard [31].

In the measurement of the MOR values, Yapıcı [29] used a Zwick/Roell Z050 universal test device with a capacity of 5000 kg and measurement capability of 0.01 N in accuracy. In the testing, the loading mechanism was operated at a velocity of 5 mm/min. As a result of the measurements, Yapıcı [29] obtained the dataset, including the experimental data from 27 OSB panels.

Here, we use his experimental dataset for predicting the MOR values of the OSBs. We follow the same steps as we did for the first problem in Section 3.1. As a first step, we construct a 4-layer FFANN (3, 8, 8, 2 neurons within each layer), as shown in Figure 5.

In defining the input and output variables, we followed the variables reported by Yapıcı [29], where 3 variables are used as inputs: adhesive material, time (min), and press pressure (kg/cm$^2$), and 2 variables are used as outputs: MOR (flexure parallel) and MOR (flexure perpendicular).

As a second step, we apply the rewiring process for the regular network in Figure 5 to determine the rewiring range of a small-world network. More than 40 rewirings cannot be applied to the network due to finished unconnected neurons in the network. After each rewiring process, the $D_{Local}$ and $D_{Global}$ parameters...
are calculated, as shown in Figure 6. As in Figure 3, the increasing of the RN reduces the value of \( D_{\text{Local}} \), while it increases the value of \( D_{\text{Global}} \). Since the small-world network is obtained when both parameter values are small, the small-world network is obtainable for the rewiring range of between 8 and 40.

**Figure 6.** The change in \( D_{\text{Global}} \) and \( D_{\text{Local}} \) with the RN for predicting the MOR values of the OSBs.

As a third step, we use the 10-fold cross-validation method. The used dataset, which is scaled by the min-max method in the range of [0–1], consists of 27 samples. It is then split into 10 approximately equal-size partitions, and one of them is used for the testing and the remainder are used for the training; this process is repeated 10 times. The overall test errors (MAE, MSE) and statistical correction parameters (\( R^2 \), RMSE) of the model are calculated by averaging the test error and statistical correction parameters of each test partition. We perform 60 experimental trials with the minimum training error (MSE) criterion. The ending error criterion (MSE) of the training process is 0.000001.

The results are given in Table 3, which indicates that the best performance is obtained for an optimal RN of RN = 16, within the small-world network range. Figure 7 also shows how the MOR values of the small-world FFANN for an optimal RN match the experimental data. Therefore, we conclude that the small-world FANN provides the best prediction of the MOR values of the OSBs.

**Table 3.** Rewiring-based error analysis results for predicting the MOR values of the OSBs.

<table>
<thead>
<tr>
<th>RN</th>
<th>( R^2 )- output 1</th>
<th>( R^2 )- output2</th>
<th>MAE</th>
<th>MSE</th>
<th>RMSE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.981913</td>
<td>0.94994</td>
<td>0.022299</td>
<td>0.004031</td>
<td>5.127057</td>
</tr>
<tr>
<td>4</td>
<td>0.997913</td>
<td>0.982364</td>
<td>0.012101</td>
<td>0.001127</td>
<td>2.801207</td>
</tr>
<tr>
<td>8</td>
<td>0.998752</td>
<td>0.980487</td>
<td>0.009118</td>
<td>0.000658</td>
<td>2.098895</td>
</tr>
<tr>
<td>12</td>
<td>0.995861</td>
<td>0.977478</td>
<td>0.010825</td>
<td>0.001001</td>
<td>2.382575</td>
</tr>
<tr>
<td>16</td>
<td>0.998496</td>
<td>0.995746</td>
<td>0.006923</td>
<td>0.000364</td>
<td>1.583459</td>
</tr>
<tr>
<td>20</td>
<td>0.996022</td>
<td>0.991652</td>
<td>0.010226</td>
<td>0.000781</td>
<td>2.226508</td>
</tr>
<tr>
<td>24</td>
<td>0.996045</td>
<td>0.976741</td>
<td>0.011605</td>
<td>0.001094</td>
<td>2.582601</td>
</tr>
<tr>
<td>28</td>
<td>0.992838</td>
<td>0.984314</td>
<td>0.013337</td>
<td>0.001312</td>
<td>2.979785</td>
</tr>
<tr>
<td>32</td>
<td>0.995462</td>
<td>0.977063</td>
<td>0.014287</td>
<td>0.001307</td>
<td>3.100186</td>
</tr>
<tr>
<td>36</td>
<td>0.737677</td>
<td>0.97984</td>
<td>0.025947</td>
<td>0.012859</td>
<td>5.760594</td>
</tr>
<tr>
<td>40</td>
<td>0.986455</td>
<td>0.974657</td>
<td>0.016255</td>
<td>0.002324</td>
<td>3.730499</td>
</tr>
</tbody>
</table>
Figure 7. The prediction performance of the small-world FFANN at an optimal RN (RN = 16): a) MOR (flexure parallel) and b) MOR (flexure perpendicular).

We investigated whether there is a statistically significant difference between the regular FFANN and our small-world ANN model, as given in section 3.1. The obtained results are shown in Table 4.

Table 4. Results of the t-test for the absolute errors: a) MOR (flexure parallel) and b) MOR (flexure perpendicular).

<table>
<thead>
<tr>
<th></th>
<th>Regular FFANN model</th>
<th>Small-world model</th>
<th>Regular FFANN model</th>
<th>Small-world model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>4.08059795</td>
<td>0.320180225</td>
<td>4.435303658</td>
<td>0.303957035</td>
</tr>
<tr>
<td>Std. deviation</td>
<td>3.295172256</td>
<td>0.229358048</td>
<td>4.782524812</td>
<td>0.447958736</td>
</tr>
<tr>
<td>Std. error mean</td>
<td>0.634156196</td>
<td>0.044139977</td>
<td>0.920397329</td>
<td>0.086209699</td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Df</td>
<td>26.25</td>
<td></td>
<td>26.46</td>
<td></td>
</tr>
<tr>
<td>T value</td>
<td>5.9155</td>
<td></td>
<td>4.4691</td>
<td></td>
</tr>
<tr>
<td>P value</td>
<td>0.000003</td>
<td></td>
<td>0.000132</td>
<td></td>
</tr>
</tbody>
</table>

As seen in Table 4, the P value is smaller than the significance level (0.01) for the 2 outputs of the experimental model. Thus, the H0 hypothesis is rejected. In addition, it indicates that there is a statistically significant difference between the regular FFANN and the small-world model. Thus, we conclude that the small-world network model is better than the regular FFANN model, as given in Section 3.1.

4. Conclusion

In summary, we investigated the impact of the small-world topology on the performance of the FFANN based on 2 real-life problems: estimating the thermal performance of solar air collectors and predicting the MOR values of OSBs. We use a FFANN with small-world topology to solve both problems and compare the results with those of conventional FFANNs. We show that the rewired FFANN performs better than a FFANN with zero rewiring. We also show that the performances of both of the rewired FFANNs are the best if the RN is within the small-world rewiring range, which we call an optimal RN. In addition, we demonstrate that small-world networks perform statistically significantly better than the regular FFANN. Consequently, we propose that there exists an optimal RN within the small-world topology that provides the best performance for both problems.
Acknowledgments
The authors would like to thank Dr Fatih YAPICI and Engin GEDIK of the Technical Education Faculty of Karabük University for providing their experimental data.

References