The influence of algorithms for tuning the parameters of neuromorphic systems on their fault tolerance

To cite this article: S N Danilin et al 2019 J. Phys.: Conf. Ser. 1333 032077

View the article online for updates and enhancements.
The influence of algorithms for tuning the parameters of neuromorphic systems on their fault tolerance

S N Danilin, S A Shchanikov, I A Bordanov and A D Zuev

Vladimir State University named after Alexander and Nickolay Stoletovs, 87, Gorkogo Street, Vladimir, 600000, Russia

E-mail: seach@inbox.ru

Abstract. This article describes the influence of algorithms for tuning the parameters of neuromorphic systems on their fault tolerance. This is relevant to the hardware implementation of neuromorphic systems using memristors (NSM). The study is conducted using the authors' developed a variant of the system approach and methods of simulation of artificial neural networks (ANN). By the example of a multilayer perceptron, it is shown that different ANN learning algorithms in the nominal mode of operation make it possible to achieve similar values of the operation accuracy. But due to the influence of production and operational factors in real conditions of operation, the ANN may fail. The range of allowable values of the destabilizing factors on ANN operation depends on the learning algorithm and may differ several times.

1. Introduction
Currently, the topical scientific direction is the development and study of neuromorphic systems, whose components are based on new electronic elements with the effect of resistive switching - memristors [1]. These elements are made from organic [2] and from inorganic [3] materials. The effect of resistive switching is achieved through various physical processes. The advantage of memristors for use in neuromorphic systems is that they can be used as a synapse for various types of artificial neural networks (ANN) [4, 5].

One of the most important stages in the creation of neuromorphic systems based on memristors (NSM) is the tuning of their parameters or training. In the process of learning, the values of the weights of neuron synapses and (or) biases change in accordance with a given algorithm until the necessary operation accuracy is achieved calculated using some metric (SSE, MSE, MAE, RMSE, etc.). The learning algorithms differ for the ANN models, depending on how the process of learning is implemented - external and internal, software and hardware [6].

The relevant task, in this case, is to determine how the learning algorithm influences the fault tolerance of the ANN and NSM. These studies will help to choose algorithms that allow training ANN with maximum fault tolerance. However, it is more important to identify the features of the learning process of NMS, which provide this fault tolerance in order to develop algorithms for fault-tolerant learning. Let us consider the influence of the learning algorithm on the ANN fault tolerance using the example of a multilayer perceptron.
2. Method
Let us conduct research with the use of the general approach developed by the authors and described in [7]. In accordance with this approach, the failure of the ANN or NSM is an event in time at which the accuracy (performance) of its operation will be out of some predetermined value calculated by one of the metrics - SSE, MSE, MAE, etc. This value is called tolerance. The tolerances for the ANN operation accuracy must be specified in the design specification.

The value of the accuracy of ANN functioning achieved after training is nominal. Under actual operating conditions, this value will change due to the influence of many types of errors and faults. With reference to NSM, these may be errors in the weights of neuron synapses, caused by the instability of the electrophysical characteristics of memristors. At some values of the errors of the weights of neuron synapses, the accuracy of the ANN functioning will become less than allowable, which means the onset of ANN failure.

This study of fault tolerance was conducted using the ANN and NSM simulation models developed by the authors [8]. First of all several ANN models with a multilayer perceptron architecture were synthesized, then they were trained to solve simple mathematical tasks using different learning algorithms and then their fault tolerance was researched.

The method of research of the fault tolerance of ANN:
- Fix the value of the MSE (mean squared error) achieved during the learning process and calculated on the training set.
- In the loop, generate the errors of the weights of the neuron’s synapses of the ANN \( \varepsilon = m_0 \pm \Delta m \% \), \( m_0 = 0 \), \( \Delta m = 0.1 \).
- Simulate the work of the ANN.
- Fix the received value of the MSE calculated on the testing set.
- Brake the loop when the received value of accuracy becomes less than allowable.
- Carry out visualization and analysis of research results.

3. Experiment
For experiment 2 models of ANN were created. The first model is with one hidden layer of 10 neurons. The second model with two hidden layers of 10 neurons. The activation function of the hidden layer neurons is the hyperbolic tangent, the output layer is linear. Several learning algorithms were used:
- Broyden Fletcher Goldfarb Shanno (train_bfgs).
- Newton-CG method (train_cg).
- Gradient descent with momentum backpropagation (train_gdm).
- Resilient Backpropagation (train_rprop).

The first model is trained to approximate the function \( y = \sin(x) \). The training set consists of 25 samples of \( y \), and the testing set consists of 100 samples of \( y \) for an interval of input values \( x \in (0; 2\pi) \) (Figure 1 (a)).

The second model is trained to approximate the zero order Bessel differential equation (1). The training set consists of 50 samples of \( y \), and the testing set consists of 1000 samples of \( y \) for an interval of input values \( x \in (0; 7\pi) \) (Figure 1 (b)):

\[
x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - a^2) y = 0. \quad (1)
\]

ANN input values are normalized to the interval \([0; 1]\). The values of the weights of neuron synapses and biases generated during initialization are the same for all ANNs. Training was carried out before reaching the value of the MSE < 0.001. In the process of research, the number of epochs of training until the required accuracy was recorded.
4. Results and Discussion

The learning results for both models are presented as diagrams (Figures 2-3). Figures (a) show the number of epochs during which the models reached a given accuracy according to the MSE metric. Figures (b) show the values of the operation accuracy of the ANNs calculated using the training and testing sets.

Figure 1. The results of approximation performed by ANNs: a - sine; b - Bessel.

Figure 2. The results of training of the first ANN model: a – the number of epoch; b – MSE.

Figure 3. The results of training of the second ANN model: a – the number of epoch; b – MSE.
As can be seen from Figures 2 and 3, the ANNs were most quickly and accurately trained using the `train_bfgs` and `train_cg` algorithms. In general, all algorithms were able to achieve during training the required values of the parameters of the ANNs, which are necessary to ensure the specified operation accuracy. Let us study the effect of the ANNs learning algorithm on their fault tolerance. The results of the study of both models are presented in the diagrams (Figure 4).

**Figure 4.** The results of the study of the influence of the learning algorithm on fault tolerance for: a – the first ANN model; b – the second ANN model.

Figure 4 shows that for the first ANN model the fault tolerance is higher when it is training with `train_cg` algorithm, and for the second ANN model - `train_gd`. The allowable errors values of the neurons parameters differ from the maximum value for the first ANN model up to 3 times and for the second ANN model up to 1.3 times depending on the learning algorithm.

On average, the fault tolerance of the first ANN model is 2 times lower than the second. Add another hidden layer with 10 neurons to the first model and repeat the experiment (Figure 5).

**Figure 5.** The results of training of the first ANN model after adding another hidden layer with 10 neurons: a – the number of epoch; b – MSE; c – fault tolerance.

As can be seen from Figure 5 (c), after adding a hidden layer, the allowable value of the synapses errors of the first ANN model increased to 20 times and the fault tolerance is higher when it is training with `train_cg` or `train_gd` algorithms.

Let us conduct the described experiment several times with different initial values of weights and biases. The parameters of neurons are initialized with pseudo-random numbers from the range [-1; 1] according to the uniform distribution law. The results of the study of both models are presented in the diagrams (Figures 6, 7).
As the research results show, fault tolerance strongly depends on the algorithm for tuning the parameters of the NSM. The allowable error values of the neuron parameters differ from each other for the first ANN model up to 11 times and for the second ANN model up to 80 times, depending on the learning algorithm and initialization.

Figure 6 shows that for the first ANN model for different variants of the initial values of the neurons parameters, the fault tolerance is higher when it is trained with the train_gd algorithm. The maximum tolerance obtained experimentally for the first ANN model is $[-7.5\%; 9.0\%]$.

Figure 7 shows that for the second ANN model for different variants of the initial values of the neurons parameters, the fault tolerance is higher when it is trained with the train_gd algorithm. The maximum tolerance obtained experimentally for the second ANN model is $[-10.4\%; 2.6\%]$.

5. Conclusion

- Different algorithms for tuning parameters of the ANN and NSM provide similar values of the operation accuracy in the nominal mode of functioning. However, after hardware implementation using memristors, these values will change due to the inevitable influence of destabilizing factors. The authors have proposed one of the ways to ensure the required fault tolerance of the ANN and the NSM associated with the organization of their learning process.

- It is shown that the learning algorithms of ANN affect their fault tolerance. As a result of the study, the ranges of allowable values of the factors destabilizing the operation of ANN are determined and they differ several times for different learning algorithms. The informed choice of the learning algorithm influence the fault tolerance of the ANN.

- It is necessary to study more deeply the tuning algorithms of the ANN and NSM (including the algorithms for the initialization of the values of their parameters, the learning algorithms, using
metrics, etc.) which are effective tools for creating high-performance computers with the required level of fault tolerance.

Acknowledgments
This work was supported by the grant of the President of the Russian Federation, MK-3927.2019.9.

References
[1] Chua L O 2018 Five non-volatile memristor enigmas solved Applied Physics A 124 563-606
[2] Lapkin D A, Emelyanov A V, Demin V A, Erokhin V V, Feigin L A, Kashkarov P K and Kovalchuk M V 2018 Polyaniiline-based memristive microdevice with high switching rate and endurance Applied Physics Letters 112(4)
[3] Tikhov S V et al 2018 Behavioral Features of MIS Memristors with a Si3N4 Nanolayer Fabricated on a Conductive Si Substrate Semiconductors 52(12) 1540-6
[4] Antonov I N, Belov A I, Mikhailov A N, Morozov O A and Ovchinnikov P E 2018 Formation of Weighting Coefficients in an Artificial Neural Network Based on the Memristive Effect in Metal–Oxide–Metal Nanostructures Journal of Communications Technology and Electronics 63(8) 950-57
[5] Bayat F, Prezioso M, Chakrabarti B, Kataeva I and Strukov D 2018 Implementation of Multilayer Perceptron Network with Highly Uniform Passive Memristive Crossbar Circuits Nature communications 9(1)
[6] Yang C, Kim H, Adhikari S P and Chua L O 2017 A Circuit-Based Neural Network with Hybrid Learning of Backpropagation and Random Weight Change Algorithms Sensors 17(1)
[7] Danilin S N, Shchanikov S A and Sakulin A E 2017 Algorithm for Determining Optimum Operation Tolerances of Memristor-Based Artificial Neural Networks 2017 IVth Int. Conf. on Engineering and Telecommunication (Moscow: Institute of Electrical and Electronics Engineers) pp 140-44
[8] Danilin S N, Shchanikov S A, Sakulin A E, Bordanov I A 2018 Determining the Fault Tolerance of Memristors-Based Neural Network Using Simulation and Design of Experiments Vth Int. Conf. on Engineering and Telecommunication (Moscow: Institute of Electrical and Electronics Engineers)