Neural network modelling methods for creating digital twins of real objects

D A Tarkhov¹ and G F Malykhina¹,²
¹ Peter the Great St.Petersburg Polytechnic University, 29 Polytechnicheskaya, St. Petersburg, 195251, Russia
² Russian State Scientific Center for Robotics and Technical Cybernetics
E-mail: malykhina@ftk.spbstu.ru

Abstract. The new approach to the development of digital twins model of physical objects is presented. In the previous papers, we proposed two approaches. The first is based on the unified process of constructing approximate neural network solutions of boundary value problems for equations of mathematical physics. The second approach is based on the construction of multilayer models using differential equations, which makes it possible to dispense with the time-consuming procedure for training neural networks. This paper proposes a third approach based on the construction of multi-row differential equations. We are confident that the approaches we have proposed allow us to significantly simplify and unify the creation and adaptation (keeping up to date) of the digital twins of real objects of various kinds, e.g. technical, biological, socio-economic ones.

1. Introduction
A computer model that copies the behavior of a physical object is called a digital twin [1-3]. The model simulates all possible modes of operation of an object during its life, takes into account the influence of external factors and allows predicting the future behavior of a physical object. In our sight, this digital twin should be based on artificial intelligence technology.

When modeling complex technical objects, the computer analysis based on the engineering software, using the finite element method (FEM) is usually applied. However, modeling a real object with FEM method oneinvolves a number of fundamental difficulties. First, accurate information about the differential equations describing the behavior of an object is usually absent due to the complexity of describing the processes occurring in it. Secondly, the necessity to know the initial and boundary conditions, which is usually less available and less accurate. Thirdly, the properties, the characteristics and the parameters of real object are continuously changed during the operation time. This requires appropriate adaptation of the model, which is difficult to carry out with a model built on the basis of the FEM.

In this case the adaptive model for each element of the physical object seems to be more promising. The model can be refined and rebuilt in accordance with the continuously enlarged observations of the object. This model is most convenient to build in the form of a artificial neural network (ANN). Therefore, the urgent task is the development of modeling technology, a more complete account of historical and newly received data, the improvement of methods for automatic adjustment of the architecture and model parameters, methods of classification and prediction. The set of tasks facing
the digital twin model can be solved using a team of neural networks, each of which displays a certain fragment (element, process) of the object

2. Methods

The first approach, based on the unified process of constructing approximate neural network solutions of boundary value problems for equations of mathematical physics, can be found in [1–7].

Let the corresponding element (process) be described as an initial or a boundary value problem for a differential equation (ordinary or in partial derivatives). We are looking for an approximate solution of the problem in the form of the output of an ANN of a given architecture \( u(x, w) = \sum_{i=1}^{m} c_i \nu(x, a_i) \).

Here \( w = [w_1, w_2, ..., w_m] \) is the weight vector \( w_i = \{c_i, a_i\} \), which aggregates linearly parameters \( c_i \) and nonlinear parameters \( a_i \). The basic neuroelement is represented by the function \( \nu \) depending on the type of neural network and the neuron activation function.

The ANN weights \( w \) vector is in the process of epoch-by-epoch network training, purposely to minimize error functional. For the problem in question the error functional:

\[
J(w) = J_1 + \partial J_2,
\]

where \( J_1 \) expresses the satisfaction score for the equation, \( J_2 \) it is a boundary condition. The component \( J_2 \) may include terms that are responsible for other conditions of the statement: symmetry, relations at the interfaces, the equation of state. Since the digital twin must take into account the current observations of the target physical object, a term \( \tilde{\delta}J_3 \) is added to the functional \( J(w) \). It takes into consideration the accuracy of meeting these conditions.

The first approach has significant drawbacks, such as a long training procedure and the single-layer nature of the neural network. To avoid these disadvantages we have developed a second approach [8–14], which allows us to develop a fairly accurate multilayer ANN solution of a differential equation. Unlike the previous approaches, the method is not based on FEM, but on the finite difference method applied to a variable length interval. The benefit of the suggested method is the automatic inclusion in the final formula of the parameters of the problem. It allows us to avoid multiple re-decisions to investigate the effect of parameters on the obtained solution. This is especially important for building a digital twin of a specific object, taking into account its unique features.

When constructing a differential equation from experimental data, not only the coefficients are unknown beforehand, but also the order and even the structure of the equation are indeterminate. It cause difficulties in creating an accurately model of the physical object. An approach similar to the multi-row algorithm [14, 16] can be used to select a differential model. This approach consists in the selection of complex models. It is essential that the coefficients of the models and their structure are determined by the optimization of various functional.

We begin with the simplest problem of creating a model in the form of a linear homogeneous ordinary differential equation with constant coefficients. If it is required to select the differential equation

\[
a_0 y^{(n)} + \ldots + a_{n-1} y' + a_n y = 0 \]

using data acquisition \( y(x_1) = f_1, y(x_2) = f_2, \ldots, y(x_p) = f_p \) in the best possible way, then a multi-row algorithm can be applied. Algorithm allows us to consider the selected model on each row

\[
y_{l+1} = w_{i,l} y_{i,l} + w_{j,l} y_{j,l} + w_{k,l} \frac{d}{dx}(y_{k,l}).
\]

In this case, a first fraction of the data is used to construct a functional and adjust the weights, a second fraction is available to another functional
creation and the third fraction - for the selection of the best trained models. The accuracy of satisfying the equation can be checked in the required number of points.

As an illustration, we will describe a multi-row algorithm for creation a model based on linear differential equations with constant coefficients:

1. The whole set of experimental data is divided into two fractions.

2. Consider the set of the models of the first row selection $y(1) = w(1)y + w_d(1) \frac{d}{dx}(y)$. The coefficients of these models are adjusted by minimizing the discrepancy between the solutions of the equations and the first part of the experimental data.

3. From these models, we select the best ones according to the accuracy of the second part of the experimental data.

4. From $m$ the resulting models we build models of the second row of the species selection adjusting coefficients for the first fraction of the experimental data.

$$y(2) = w_i(2)y_i(1) + w_j(2)y_j(1) + w_{d,k}(2) \frac{d}{dx}(y_k(1)) \quad i, j, k = 1, \ldots, m,$$  \hspace{1cm} (2)

5. Repeat the selection process as many times as necessary. In the last step, the unique best model is selected.

We can build a more complex model for which the coefficients of the equation are some neural network functions. At the same time, the second and fourth steps of the above algorithm vary significantly. The main problem that manifests itself in this case is that the resulting non-linear differential equation cannot be solved analytically.

The way out of the situation is to use an iterative algorithm that combines the steps of selecting weights of the networks corresponding to the coefficients of the models and the stage of constructing an approximate solution of the corresponding equations, based on minimizing the obviously modified functional (1). To determine the structure of these functions, we can use any genetic algorithm. Similarly, we consider a heterogeneous equation with an unknown right side, which is sought in the class of some neural network functions.

A fundamentally new class of models arises if we consider a multi-row algorithm, in which a new selection series is determined by the equality

$$y_{i+1} = w_i y_{i,l} + w_j(2)y_{j,l} + w_{q,k}\phi(y_{k,l}) + w_{d,k} \frac{d}{dx}(y_{l,k}) \quad i, j, k, l = 1, \ldots, m,$$  \hspace{1cm} (3)

where $\phi$ is some non-linear activation function (we choose the sigmoid type). To construct each such model, it is necessary to determine only four coefficients, which are linearly input, although the result is a very complex non-linear model. It is fundamentally important that the construction of such a model is performed using the same algorithm as for a linear model with variable coefficients.

Even more complex models arise, if we consider $[w_i, w_j, w_{q,k}, w_{d,k}]$ as not a constants, but functions that can be represented by neural networks. To determine the structure of such models, we can propose a double multi-row algorithm, which consists in the following. Each stage of determining the basic model is that a set of linear models of the structure defined above is considered, the coefficients $[w_i, w_j, w_{q,k}, w_{d,k}]$ each of which are neural network functions and the structure of these functions can also be determined by a multi-row algorithm (or another, for example, genetic algorithm or error clustering). Thus, at each stage of the main algorithm, several auxiliary stages are implemented and the number of stages itself can be selected using genetic algorithm.
You can immediately consider such a multi-row model, when the result of the action of a linear differential operator is fed to the input of a layer of neurons at the outputs of the previous layer and the activation function can be just a sigmoid function or some non-linear operator (linear "in small" and having a limited image "in large"). An example of such an operator is the mapping,
\[
K : y(x) \mapsto \left( K(y) \right)(x) = \int \kappa \left( x, y(x'), x \right) dx',
\]
where \( \kappa \) is a bounded function differentiated at zero by the second argument and having the necessary smoothness properties by the first.

Similarly, one can consider the multi-layer model with partial derivatives. The multi-row algorithm for constructing such a model does not practically differ from the one discussed above, only on each selection row should the models of the type be considered
\[
y_{l+1} = w_i y_{i,l} + w_j y_{j,l} + w_{q,l} \phi \left( y_{k,l} \right) + w_{p,k} \frac{\partial}{\partial x_p} \left( y_{k,l} \right),
\]
(4)

3. Conclusion
Methods for solving differential equations of mathematical physics, using RBF ANN, are the basis for creating dynamic twin models of physical and technical objects. The paper proposes a new class of mathematical models, combining the properties of neural networks and differential (integro-differential) equations. Algorithms for the formation of such models based on experimental data are considered. Such models may be interesting in a situation where standard mathematical models are inadequate to the studied physical processes.

Acknowledgments
This work was supported by the Russian Science Foundation (project code 18-19-00474).

References
[1] Vasilyev A and Tarkhov D 2014 Models of Complex Systems on the Basis of Artificial Neural Networks Nonlinear Phenomena in Complex Systems, Mathematical "Education and Upbringing" Publishing, Minsk, Belarus 17(2) 327-335
[2] Gorhochenko V I., Lazovskaya T V., Tarkhov D A., Vasiliev A N and Zhukov M V 2016 Neural Network Technique in Some Inverse Problems of Mathematical Physics, Advances in Neural Networks - 13th International Symposium on Neural Networks, Lecture Notes in Computer Science 310–316
[3] Shemyakina T A., Tarkhov D A and Vasiliev A N 2016 Neural Network Technique for Processes Modeling in Porous Catalyst and Chemical Reactor Advances in Neural Networks — ISNN 2016: 13th International Symposium on Neural Networks Lecture Notes in Computer Science 547–554
[4] Kaverzneva T, Lazovskaya T, Tarkhov D and Vasilyev A 2016 Neural network modeling of air pollution in tunnels according to indirect measurements Journal of Physics: Conference Series 772 1-9
[5] Lazovskaya T N, Tarkhov D A and Vasilyev A N 2017 Parametric Neural Network Modeling in Engineering Recent Patents on Engineering 11 1 10–15
[6] Budkina E M, Kuznetsov E B, Lazovskaya T V, Tarkhov D A, Shemyakina T A and Vasilyev A N 2017 Neural network approach to intricate problems solving for ordinary differential equations Optical Memory and Neural Networks 26 96–109
[7] Antonov V, Tarkhov D and Vasilyev A 2018 Unified approach to constructing the neural network models of real objects Mathematical Models and Methods in Applied Sciences 41 9244-9251
[8] Lazovskaya T and Tarkhov D 2016 Multilayer neural network models based on grid methods IOP Conf. Series: Materials Science and Engineering 1-10
[9] Lazovskaya T, Tarkhov D and Vasilyev A 2016 Multi-Layer Solution of Heat Equation. Advances in Neural Computation Machine Learning and Cognitive Research, Studies in Computational Intelligence 736 17–22

[10] Vasilyev A N, Tarkhov D A, Tereshin V A, Berminova M S and Galyautdinova A R 2017 Semi-empirical Neural Network Model of Real Thread Sagging Studies in Computational Intelligence 736 138–146

[11] Zulkarnay I U, Kaverzneva T T, Tarkhov D A, Tereshin V A, Vinokhodov T V and Kapitsin D R 2016 A Two-layer Semi Empirical Model of Nonlinear Bending of the Cantilevered Beam Journal of Physics: Conference Series 1044 1-9

[12] Bortkovskaya M R, Vasilyev P I, Zulkarnay I U, Semenova D A, Tarkhov D A, Udulov P P and Shishkina I A 2017 Modeling of the membrane bending with multilayer semi-empirical models based on experimental Proceedings of the II International scientific conference “Convergent cognitive information technologies” (Convergent’2017) (RWTH Aachen University) 24–26

[13] Vasilyev A, Tarkhov D and Shemyakina T 2016 Approximate analytical solutions of ordinary differential equations Selected Papers of the XI International Scientific-Practical Conference Modern Information Technologies and IT-Education (RWTH Aachen University) 393-400

[14] Tarkhov D and Shershneva E 2016 Approximate analytical solutions of mathieu’s equations based on classical numerical methods. Selected Papers of the XI International Scientific-Practical Conference Modern Information Technologies and IT-Education 356-362

[15] Malada H R and Ivakhnenko A G 1994 Inductive Learning Algorithms for Complex Systems Modeling 384

[16] Malykhina G F and Tarkhov D A 2018 Digital Twin Technology as a Basis of the Industry in Future 18-th PCSF 2018 Professional Culture of the Specialist of the Future. The European Proceedings on Social &Behavioral Sciences LI 45 416-428