UDC 681.5:519.24

https://doi.org/10.29202/nvngu/2019-5/21

V. Yesilevskyi[1], Cand. Sc. (Tech.), Assoc. Prof., orcid.org/0000-0002-5935-1505,
S. Dyadun[2,3], Cand. Sc. (Tech.), Assoc. Prof., orcid.org/0000-0002-1910-8594,
V. Kuznetsov[4], Cand. Sc. (Tech.), Prof., orcid.org/0000-0003-4638-5260

1 – National University of Radioelectronics, Kharkiv, Ukraine, e-mail: yes_v_s@ukr.net
2 – Beketov National University of Urban Economy, Kharkiv, Ukraine
3 – Karazin National University, Kharkiv, Ukraine, e-mail: daulding@ukr.net
4 – Starooskolsky Technological Institute A. A. Ugarova (branch) of "National Research Technological University "MISiS", Stary Oskol, Belgorod Region, Russian Federation, e-mail: 1937vnk@gmail.com

LSTM NETWORKS FOR ANAEROBIC DIGESTER CONTROL

Purpose. To study the possibility of use of artificial intelligence systems based on neural networks for the improvement of the efficiency of the biogas production in the anaerobic digester by optimization of the nonlinear control system.

Methodology. Mathematical modeling and computer simulation for building an adaptive control system based on neural LSTM network and reinforcement learning. The study of the learning convergence of the neural network based on a non-linear mathematical model of the technological process.

Findings. The performed computational experiment showed that the learning convergence could be achieved after 600 episodes in a model with two control channels and four measurement channels. It also proved the efficiency of the proposed adaptive control system for the program model of the technological process. The results confirmed the ultimate possibility to use this approach for biogas production control in the anaerobic digester in the real world settings.

Originality. It is shown that the task of controlling the biogas production in the anaerobic digester can be solved based on the reinforcement learning approach implemented for a continuous nonlinear dynamic system. The system can be implemented as two neural networks in the actor-critic architecture. To build the module of critic and the module of the controller of the adaptive control system, it was proposed to use the neural network architecture based on a special type of recursive neural network called LSTM neural networks, which were not previously used to control the production of biogas in the anaerobic digester.

Practical value. The proposed method for constructing a control system for the anaerobic digesters will lead to an increase in the efficiency of biogas production as an alternative renewable energy source. Expansion of the scope of use of anaerobic digesters will also have a significant impact on solving the problem of utilization of industrial and household human waste.

Keywords: *process control, anaerobic digester control, reinforcement learning, neural networks, LSTM networks, adaptive critic systems*

Introduction. The search for alternative energy sources makes the production of biogas in the anaerobic digester as a renewable fuel increasingly important. The biochemical process enables conversion of complex organic materials in digesters into a clean renewable energy source – biogas. Recent research has demonstrated the great potential of biogas not only as a fuel but also as a product for processing various types of waste [1].

However, the creation and maintenance of the necessary technological process in the anaerobic digester is not an easy task and requires new solutions. Due to the complexity of the biogas production process and the simplified control system, the anaerobic digester usually works below their optimal performance. In addition, the prevailing use of PID regulators for production parameter control makes the process of biogas production inefficient – practically, on the edge of economic feasibility.

Today, there is a need for an adapting monitoring and optimization system for fitting the parameters of the biogas production process and the composition of the fermented substrate. Existing systems use predictive control schemes, which are ineffective in a dynamic production environment.

In this paper, we propose a solution, which helps to increase the efficiency of sludge processing in the anaerobic digester. The idea is to create an adaptive control system aimed at increase in both the biogas yield and the methane concentration in it. The successful operation of such a system would create the conditions for the economically feasible introduction of biogas production technology to alternative energy and waste disposal systems.

Related work. The anaerobic digester (Fig. 1) is a continuous chemical anaerobic reactor designed to produce biogas by microbiological desorption of organic substrates. A mixture of raw substrates and biologically active sludge is fermented [2].

The conversion of the substrate loaded to an anaerobic digester into biogas can be performed under two different temperature conditions: mesophilic fermentation is carried out at the temperature of 30–35 °C, and thermophilic fermentation at the temperature of 50–55 °C.

Each type of fermentation corresponds to a specific type of bacteria. Bacteria are very sensitive to temperature. A change in temperature by several degrees can lead to a change in the dominant type of reaction from the methane reaction to the acid fermentation.

The anaerobic digesters in the thermophilic process use the flow scheme. Raw materials are heated with the superheat-

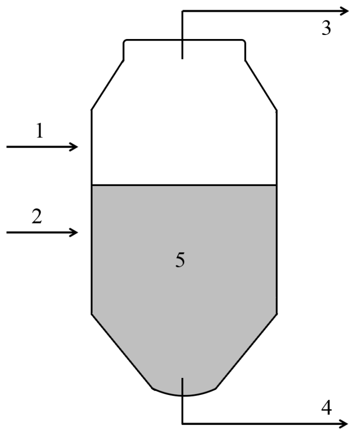

Fig. 1. Functioning of the anaerobic digester:

1 – heating with the steam; 2 – the substrate load; 3 – biogas output; 4 – draining sludge; 5 – mixture for fermentation

ed steam by a steam jet injector. Temperature is the most important parameter that can be controlled in the fermentation process. Thus, most of the other parameters are usually not analyzed. This leads to ineffective work of the digester. Among the other important parameters are:

- the amount of fats, proteins, carbohydrates, as well as the humidity and the temperature of the loaded substrate;
- the acidity;
- biochemical oxygen consumption of the incoming and purified fluid;
- loading speed of the anaerobic digester;
- mixing speed;
- steam temperature and pressure;
- gas pressure in the anaerobic digester.

Knowledge of these parameters allows taking into account the features of the biochemical reactions and adjusting the mode of the digester's operation.

Mathematical modeling of biochemical and physical processes occurring in the digester is quite difficult to implement [3, 4]. The process is characterized by significantly different dynamics for the control channels of temperature and gas flow, concentrations of liquid and solid phases. In addition, bacteria involved in the synthesis of methane use mechanisms which are still not sufficiently understood.

There have been attempts to build digester control systems aimed at optimization of the biogas production [5, 6]. The control schemes proposed in the correspondent papers use control with prediction. Such an approach requires a preliminary identification of the control object, which can lead to inefficient control if the conditions have changed.

Control modules of the dynamic systems for which a formalized model cannot be created, machine-learning methods, in particular, neural networks are successfully used.

Neurocontrol is a method for adaptive control. Initially, the classical multilayer network and correspondent supervised learning methods were proposed for the direct control of the object. The best control results were obtained when using multilayer networks with delay lines.

Despite their popularity, these classical methods do not work optimally with modern biogas production systems. The reason is the highly non-linear nature of the control object, which is inflicted by the uncertainty of external factors. This also increases the complexity of the control problem.

An interesting approach is to use reinforcement learning of neural networks for both discrete [7] and continuous [8] systems. The effectiveness of this approach is proved by modeling. This approach can also be used for our task of the digester control.

Recently, new solutions based on recurrent neural networks that model the human brain functioning more accurately, have appeared in the field of artificial intelligence [9].

LSTM networks (Long short-term memory units) are a special kind of recurrent neural networks. They are relevant due to the high speed of learning and the ability to memorize long-term dependencies. This allows using them effectively for dynamic systems control.

However, the possibilities of using LSTM networks for adaptive control of the anaerobic digester with the help of the reinforcement learning approach have not yet been studied. This publication is aimed at the study of this approach and is organized as follows. First, the goals and objectives of the research are given. Furthermore, the main theoretical prerequisites solving the problem of adaptive control based on LSTM-networks and the structure of the proposed system are considered; a simplified model of the control object is shown, the figures describe the characteristics of the learning process. finally, the results are discussed.

Purpose. The aim of the work is to study the possibility of using LSTM networks for adaptive control systems as an enabler of biogas production increase in an anaerobic digester.

The following tasks were formulated to achieve this aim:

- to determine the appropriate methods to solve the control problem in continuous technological processes that cannot be precisely mathematically modeled and are characterized by uncertainties of the initial conditions and external factors;
- to explore possible options and justify the choice of the type of neural network control system;
- to suggest a structure and a procedure for setting up (learning) of the control system and a scheme for its implementation;
- to model the control system for biogas production in the anaerobic digester based on the actor-critic model using LSTM neural networks and to confirm the fundamental possibility of using control systems of a new type for biogas production in the anaerobic digester.

Theoretical background. In this paper, we describe an adaptive system controlled by intelligent agents [10]. They learn to make decisions in the changing conditions of the biogas production process. Intelligent agents utilize the unsupervised learning strategy in order to choose random actions (or sequences of random actions) and the assessment of the particular action quality.

Using randomly selected actions and accumulated "experience", an agent eventually elaborates a predictive model of the system. To assess the quality of the agent's actions, some feedback that separates the "bad" actions from the "good" ones is needed. Systems of this type, implemented as programmable units are already well known [11, 12].

This approach to the strategy elaboration based on reward is described for the method of optimal strategies definition in a discrete Markov decision process (MDP) [10]. The optimal strategy is an agent's behavior strategy that maximizes the expected total reward. However, unlike the discrete Markov model, the problem of control of the anaerobic digester is continuous; moreover, there is no complete model of the dynamic system. In addition, the reward function is unknown.

Similarly to optimal control, reinforced learning algorithms, minimize the cumulative sum of costs for a given time horizon. In the system, described in this paper, optimal control actions are generated at the training stage during the process of trial and error in direct interaction with the anaerobic digester. Artificial neural networks are chosen as a learning model adapting to the control of a nonlinear multidimensional dynamic system.

Reinforcement Learning. Reinforcement Learning (RL) is a computational approach for sequential decision making under uncertainty.

Fig. 2 shows the interaction between the agent and the environment: the agent takes an action that changes the state of the system and receives a scalar reward signal from the environment. The RL algorithm attempts to maximize the cumu-

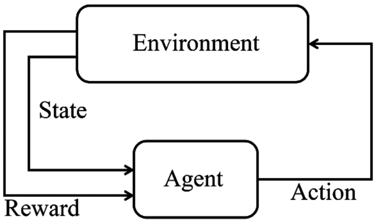

Fig. 2. The basic idea of the reinforcement learning

lative reward by studying the unknown environment during the process of trial and error.

The Markov decision process for a discrete system formalizes the interaction between the agent and the environment. It possesses the following basic components:

$s \in S$: state space

$a \in A$: action space

$r \in R$: reward obtained after the transition from the current state to the next one.

$\pi(a|s)$: a strategy mapping the state s to action a. Usually, a stochastic strategy is used, but deterministic strategies can also be used.

$P(s, r|s, a)$: transition probability distribution, which defines one dynamic step. It is the probability that the environment will return the reward r after the transition to the next state $s\_$ from state s after action a.

The process can be considered an interaction of an agent with the environment at discrete time stamps, $t = 0, 1, 2, 3, \ldots$. At time stamp t, the agent evaluates the state of the environment s_t, which is the argument for the choice function of the action at, available for this state. Evaluation of the reward rt allows finding the accuracy of the strategy at the next step.

For each value of step t, the agent refines the mapping $\pi(a|s)$ by changing the probability of each action selection. Thus, the agent creates a strategy by an iterative procedure, which is called reinforced learning. The strategy generation allows maximizing the cumulative amount of reward. This process is formalized as the G_t rewards function.

The goal of the RL agent is to learn the optimal policy that maximizes the expected amount of G_t rewards

$$G_t = r_{t+1} + r_{t+2} + r_{t+3} + \ldots + r_{t+N}. \tag{1}$$

The goal is to use reward information to determine the expected utility. The utility is defined as the expected amount (possibly decreasing over time) of the obtained rewards if the agent follows the strategy π.

In order to simulate the amount decrease (discounting), the amount is usually weighted. A discount factor $\gamma(0 \le \gamma \le 1)$ is introduced to determine the current value of the reward. In case of the discount, the total reward is determined by the equation (2).

$$G_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \ldots. \tag{2}$$

With $\gamma = 1$, the total reward is the sum of the rewards at each step, and with $\gamma = 0$, if only the last reward is estimated.

The choice of the intermediate value γ makes it possible to adjust the confidence of the estimation. The estimated function (3) can be introduced as the mathematical expectation of the total reward. This function estimates the expected total return of Gt, starting from a certain state s after applying the π strategy.

$$v_\pi(s) = E_\pi[G_t|s]. \tag{3}$$

The basic idea (3) is that the function takes the optimal value of v_π^*, when the optimal strategy is π^*.

The solution of this problem is based on the classical Q-learning algorithm. The algorithm is based on a virtual table (Q-tables) in which the reward for each possible action is recorded for each possible state.

Learning, in this case, means choosing an action according to this table (with some randomness introduced) and adjusting it after reward obtaining. This formulation and the technique go back to the Bellman dynamic programming method. Neural networks can solve this problem. Great opportunities of such algorithms are proved by the Deep Q-Network (DQN) system [13].

This approach is used for discrete systems with a finite number of states and actions, which is the main disadvantage of it. Attempts to break continuous intervals into a finite number of subintervals cause numerous problems. The problem of dimensionality of the obtained table is one of them, though not even the most challenging one.

Reinforcement learning can be applied to continuous models with the parametrized strategy space. Thus, the RL problem is converted into an optimization problem for the objective function $J(\theta)$ according to the strategy π (θ)

In this case, it is possible to use gradient methods, which gives certain advantages compared with the stochastic choice of strategy for the evaluation function. A known method of this type is the REINFORCE method [14]. The REINFORCE method requires the gradient strategy to the objective function $J(\theta)$. For example, a strategy can be represented by a neural network, its weights are the parameters of the strategy, the input signal is the current state, and the output is equal to the probability of an action choice. If θ is the vector of strategy parameters and J is the performance of the corresponding strategy (for example, the average reward per step), then the parameters are adjusted proportionally to the gradient (4) in case of the gradient strategy.

$$\Delta\theta = \alpha \frac{\partial J}{\partial \theta}. \tag{4}$$

A neural network model with θ parameters can be used as a parameterized model for determining a stochastic or deterministic strategy. In [15], a neural network system based on a deterministic gradient policy (DDPG) for continuous spaces was proposed. The principle of dynamic programming in this case requires using two separate cycles: the agent training cycle and the critic training cycle. In the agent cycle, the neural network learns to approximate the optimal control signal. The critic learns to optimize the function $J(t)$. Methods of adaptive critic are based mainly on the gradient strategy. They allow approximating the model of the control system implicitly through the generation of reward values. The adaptive critic system consists of two neural modules: the control module (actor) and the critic module. In the control module, the learning process is reduced to minimization of the functional cost $J(t)$. The critic module calculates the approximation of the cost function. Thus, in actor-critical methods, an actor is built as a neural network with a parametrized strategy, and the critic is a neural network with a parameterized estimate of rewards.

LSTM neural network. Classical multilayer neural networks are very slow learners, which is critical for control tasks There are two main reasons for this:

1. The computing resources of modern computers are still insufficient for learning by back-propagation, especially when networks have several layers and a large number of hidden nodes.

2. The problem of the vanishing gradient: the error gradient decreases from layer to layer during backward propagation, which leads to a very long learning process.

Moreover, classical multilayer neural networks are not adapted to remember information, which is one of their weaknesses. Recurrent neural networks (RNN) are aimed at overcoming this weakness: they contain cycles that are allowed to store information.

RNN is similar to the classical neural network if one would trace its work in time. The neural network element receives some value x as its input and returns the value h after conversion.

The cycle uses the output h along with the consequent value of x in the next conversion step (Fig. 3).

In fact, the recurrent network can be transformed into a sequence of multilayer neural networks that transmit information to subsequent layers, as shown in Fig. 4.

This chain shows that recurrent neural networks are designed to work with sequences and lists. RNNs are relevant architectures for use in control systems where the sequences of signals are precisely the most important issue for generating control actions.

Recently, such networks have been considerably successful in many applications of artificial intelligence, such as NLP, image and speech recognition, machine translation and others.

Such success was essentially facilitated by the emergence of LSTM (Long short-term memory) networks – a specific type of recurrent neural networks that work better than classical multilayer networks in a large number of tasks. Long-term memorization of information is an important feature of LSTM networks. Almost all known breakthroughs in the field of artificial intelligence have been achieved using LSTM networks.

LSTM networks solve the vanishing gradient problem by adding three gates into the neural cell architecture (input gate, output gate, forget gate). Their task is to use memory of the past states effectively.

LSTM networks have a four-layer, specifically organized structure (Fig. 5).

The following chain can formally describe the LSTM computational scheme:

$$
\begin{aligned}
&f(k) = \sigma(W_f[h(k-1), x(k)] + b_f) \text{ forget gate}\\
&i(k) = \sigma(W_i[h(k-1), x(k)] + b_i) \text{ input gate}\\
&o(k) = \sigma(W_o[h(k-1), x(k)] + b_o) \text{ output gate} \qquad (5)\\
&C_i(k) = \tanh(W_c[h(k-1), x(k)] + b_c) \text{ input state}\\
&C(k) = f(k)C(k) + i(k)C_i(k) \text{ cell state}\\
&h(k) = o(k)\sigma(C(k)) \text{ hidden state,}
\end{aligned}
$$

where k is the step number, $x(k)$ is the input vector, $h(k)$ is the output vector, $C(k)$ is the state vector, $f(k)$, $i(k)$, $o(k)$ are the gate vectors. W_f, W_i, W_o, W_c are matrices of weights and b_f, b_i, b_o, b_c are biases for forget, input, output, and cell elements respectively, $\sigma(\cdot)$ is a sigmoid activation function for input, output and forget gate, $\tanh(\cdot)$ is the activation function for the state of the memory cell based on the hyperbolic tangent.

Information memorization in LSTM networks is implemented by the variable of the cell $C(k)$ state. The state of the cell is transferred through the entire chain of calculations. In some cases, the information may pass with little or no change. The gates values regulate the ability of the network to add or remove information in the cell.

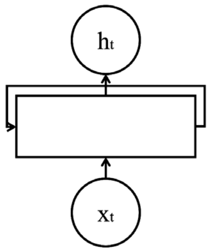

Fig. 3. The recurrent network cycle:
X_t – a sequence of inputs; h_t – a sequence of states

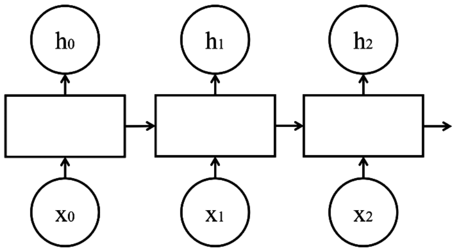

Fig. 4. Expanded recurrent network:
x_i – is a sequence of inputs; h_i – is a sequence of states

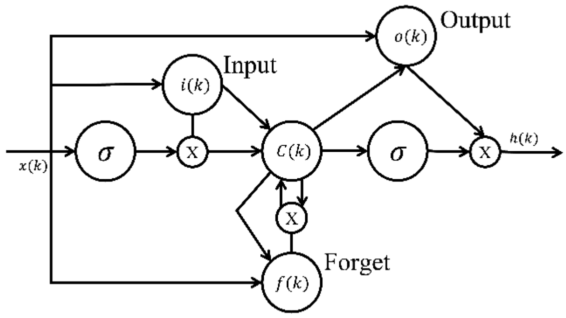

Fig. 5. LSTM scheme used in a recurrent neural network

These advantages of the LSTM network can be used to control the operation of the digester. Networks with such an architecture are trained faster than classical back-propagation networks since they have no vanishing gradient problem.

Additionally, it is very important that they are adapted for storing sequences, especially for dynamic systems which operate with the output signal as a result of the sequence of input control signals.

Structure of the anaerobic digester control system based on reinforcement learning. The proposed reinforcement learning driven control system for the anaerobic digester control consists of two LSTM networks, one for the actor and one for the critic. In Fig. 6 the RL architecture of interaction between actor and critic controllers is shown. Several steps are performed to monitor the state.

First, the controller gives out a set of control actions. The actor calculates the output of the control action after observing the state. The critic calculates the value of the reward function. This value is used as a baseline for updating the critic strategy gradient.

Fig. 7 shows a diagram for the reinforcement learning system in the mode of the object control. The controller receives the vector of the current reference signal $r(k+1)$ and the state vector $S(k)$, which is a sequence of the last outputs of the system $\{y(k), y(k-1), \ldots, y(k-N)\}$. The neurocontroller generates a control signal $u(k)$, consequently the control object generates the output $y(k+1)$.

The neural control system learning is performed on the model of the object (or on the real object) and includes simul-

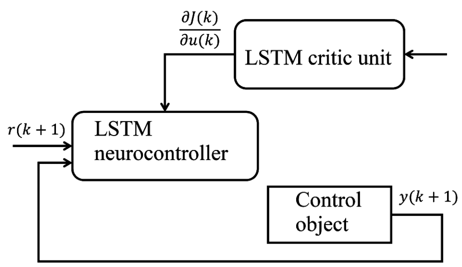

Fig. 6. The implementation of reinforcement learning

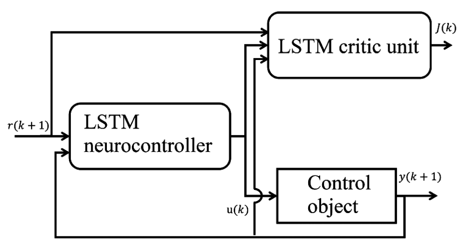

Fig. 7. Implementing control in reinforcement learning

taneous learning of both the critic module and the control controller. The critic module learns by changing its weights according to the gradient descent method. The weights of the controller are adjusted by the same scheme.

The training procedure continues until either the steps limit is reached, or the sufficient control quality is achieved.

Simulation of anaerobic digester control based on actor-critic model. To test the possibility of building a biogas control system in a digester, computer simulation of the technological process was carried out. An adequate biogas production model had to be chosen.

Analysis of classical mathematical models of the kinetics of biochemical processes has shown that the model of enzymatic reactions described by the Mono equations can be used. According to it, the simplest closed system of kinetic equations for a flow system can be represented as follows [6]

$$\frac{dX}{dt} = \mu X - \frac{Q}{V}X;$$

$$\frac{dS}{dt} = -\mu X/Y_s + \frac{Q}{V}(S_0 - S);$$

$$\frac{dT}{dt} = \frac{Q}{V}(T_0 - T) + Gu; \qquad (6)$$

$$m = \mu_{max}/(1 + K/S);$$

$$\mu_{max} = 0.013T - 0.129,$$

where S is the concentration of the substrate, X is the concentration of sludge in the digester, Q is the flow rate, V is the volume of the anaerobic digester, S_0 is the concentration of the incoming substrate; Y_s is a coefficient showing the amount of the absorbed substrate used for biomass creation, m is the specific growth rate of the substrate, μ_{max} is the maximum specific growth rate of the substrate corresponding to the maximum saturation level, K is the constant of half saturation, Gu is the heating intensity.

The gas yield can be described in terms of growth rate and gas yield factor Y_g

$$G = QY_g\mu X. \qquad (7)$$

We performed computer simulation of the critic learning in the actor-critic system and the implementation of the control action according to the abovementioned scheme. In the simulation, we used the values of variables for the thermophilic process from Table.

For this model we chose the heating intensity Gu and the flow rate of the substrate in the digester Q as the control actions.

For the simulation, we used TensorFlow, An open source machine learning library for research and production developed by Google. This library includes modules for working with various types of neural networks.

The computational experiments showed the convergence of the learning process (Fig. 8). The test results showing the process of changing the biogas yield for the model with the learning control system are shown in Fig. 9. The results prove the fundamental possibility of the proposed system application to control of the process of biogas production in the anaerobic digester.

Discussion of research results for the control system based on LSTM networks. As a result of the learning of the proposed biogas production control system in the digester, it is shown that an increase in the efficiency of the technological process can be achieved through the use of an adaptive system of multidimensional regulation. This system can be built based on neural network control using deep learning networks that have recently demonstrated great results in many areas of artificial intelligence.

It is shown that for dynamic control systems the use of LSTM networks is preferable comparing to classical multilayer networks.

Table

Parameter values for simulation the conditions of operation of the anaerobic digester

Parameter	Value
K	1458
Q	300
S	300
S_0	9000
T	55
T_0	30
V	3000
X	0.5
Y_s	5.75e–5
Y_g	3.045
μ	1
μ_{max}	0.52

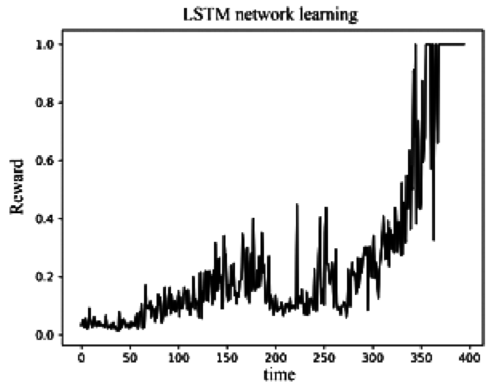

Fig. 8. Curve of the learning process:

on the X axis – the step number (time); on the Y axis – the level of reward

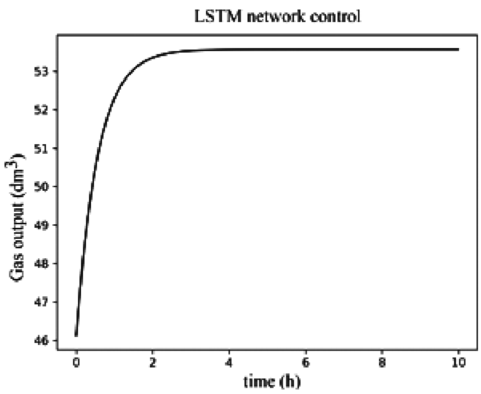

Fig. 9. Curve showing gas output after training:

on the X axis – the step number (time); on the Y axis – the calculated level of the gas output

It has to be mentioned that the mathematical model of the technological process was essentially simplified. The aim was to prove the concept and verify the ultimate possibility of the proposed control system use. The choice of real-world control actions and measurable process parameters is possible only in the course of a full-scale experiment, which is not the goal of this research. However, the obtained result proves the possibility and effectiveness of the proposed approach to building an adaptive control system. The proposed system will provide an increase in the efficiency of the anaerobic digester, which allows solving the problem related to alternative energy production and utilization of household and industrial wastes.

Further improvements of the proposed control system are related to the experiments with the extended mathematical model of the process.

Conclusions.

1. The proposed structure of the biogas production control system in the digester allows using adaptive optimal control methods for complex technological processes that do not have an accurate description of the mathematical model and are characterized by high uncertainty of the initial conditions and external factors. There are more than 10 possible parameters, measured as process states; the number of control channels depends on the complexity of the equipment of a particular digester and can vary from 2 to 10. In this case, the state parameters are in the nonlinear mutual influence.

2. It is shown that the problem of control of the biogas production in the anaerobic digester can be solved by reinforcement learning driven approach implemented for a continuous dynamic system (unlike the classical approach for systems with a discrete state space). The system can be implemented as two neural networks in the actor-critic architecture. The neural control system includes two neural network modules: the module of critic learning and the module of controller learning.

3. A special type of recursive neural network called LSTM neural networks is proposed for the first time for control of the biogas production in the anaerobic digester.

4. Computer simulations confirmed the possibility of this approach implementation for an adaptive control system of biogas production in an anaerobic digester. We achieved the convergence of the learning process for 600 episodes on a simplified model with two control channels and four measurement channels. We also proved the possibility of subsequent extension of the model without fundamental reorganization of the control system.

References.

1. Grando, R. L., de Souza Antune, A. M., Da Fonseca, F. V., Sánchez, A., Barrena, R., & Font, X. (2017). Technology overview of biogas production in anaerobic digestion plants: a European evaluation of research and development. *Renewable and Sustainable Energy Reviews*, *80*, 44-53.

2. Malina, J. (2017). *Design of Anaerobic Processes for Treatment of Industrial and Muncipal Waste, Volume VII*. Routledge.

3. Rincón, A., Villa, J., Angulo, F., & Olivar, G. (2014). A Dynamic Analysis for an Anaerobic Digester: Stability and Bifurcation Branches. In *Mathematical Problems in Engineering*, *2014*.

4. Jimenez, J., Latrille, E., Harmand, J., Robles, A., Ferrer, J., Gaida, D., ... & Mendez-Acosta, H. (2015). Instrumentation and control of anaerobic digestion processes: a review and some research challenges. *Reviews in Environmental Science and Bio/Technology*, *14*(4), 615-648.

5. Oppong, G., & Montague, G. (2013). Towards Model Predictive Control on Anaerobic Digestion Process. *IFAC Proceedings Volumes*, *46*(32), 684-689. DOI: 10.3182/20131218-3-IN-2045.00063.

6. Lei Xue, Li, D., & Xi, Y. (2015). Nonlinear model predictive control of anaerobic digestion process based on reduced ADM1. In *10th Asian Control Conference (ASCC), 2015.* 31 May–3 June 2015. DOI: 10.1109/ASCC.2015.7244539.

7. Mnih, V., Kavukcuoglu, K., Silver, D., Andrei A. Rusu, Veness, J., Marc G. Bellemare, ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature, 518*(7540), 529-533.

8. Timothy P. Lillicrap, Jonathan J. Hunt, Pritzel, A., Heess, N., Erez, T., Tassa, Y., ... & Daan Wierstra (2015). *Continuous control with deep reinforcement learning*. Retrieved from https://arxiv.org/abs/1509.02971.

9. Spielberg, S., Gopaluni, R. B., & Loewen, P. D. (2017). Deep reinforcement learning approaches for process control. In *6th International Symposium on Advanced Control of Industrial Processes (AdCONIP)* (pp. 201-206). DOI: 10.1109/ADCONIP.2017.7983780. 2017.

10. Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: a modern approach*. Malaysia: Pearson Education Limited.

11. David, S., Lever, G., Heess, N., Degris, T., Wierstra, D., & Riedmiller, M. (2014). Deterministic policy gradient algorithms. In *Proceedings of the 31st International Conference on Machine Learning*, Beijing, China, 2014. JMLR: W&CP, volume 32.

12. Ghavamzadeh, M., Mannor, S., Pineau, J., & Tamar, A. (2015). Bayesian reinforcement learning: A survey. *Foundations and Trends® in Machine Learning*, *8*(5-6), 359-483.

13. Van Hasselt, H., Guez, A., & Silver, D. (2016). Deep Reinforcement Learning with Double Q-Learning. *AAAI*, *2*, p. 5.

14. Mnih, V., Badia, A.P., Mirza, M., Graves, A., Lillicrap, T., Harley, T., ... & Kavukcuoglu, K. (2016, June). Asynchronous methods for deep reinforcement learning. In *International conference on machine learning* (pp. 1928-1937).

15. Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., ... & Wierstra, D. (2015). Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971.*

LSTM-мережі для управління метантенком

В. С. Єсілевський[1], *С. В. Дядюн*[2,3], *В. М. Кузнєцов*[4]

1 – Національний університет радіоелектроніки, м. Харків, Україна, e-mail: yes_v_s@ukr.net

2 – Національний університет міського господарства імені О.М. Бекетова, м. Харків, Україна

3 – Національний університет імені В.Н. Каразіна, м. Харків, Україна, e-mail: daulding@ukr.net

4 – Старооскольський технологічний інститут імені А.А. Угарова, „Національний дослідницький технологічний університет „МІСіС", м. Старий Оскол, Белгородська обл., Російська Федерація, e-mail: 1937vnk@gmail.com

Мета. Дослідження можливості застосування систем штучного інтелекту на основі нейронних мереж для підвищення ефективності технологічного процесу виробництва біогазу в метантенку шляхом оптимізації нелінійної системи управління.

Методика. Математичне та імітаційне комп'ютерне моделювання для побудови системи адаптивного управління на основі нейронної LSTM-мережі з використанням методів навчання з підкріпленням. Дослідження збіжності процесу навчання нейронної мережі на нелінійній математичній моделі технологічного процесу.

Результати. Проведений обчислювальний експеримент показав збіжність процесу навчання за 600 епізодів на моделі із двома керуючими каналами й чотирма каналами вимірювання та ефективність роботи запропонованої адаптивної системи управління для програмної моделі технологічного процесу. Отримані результати підтверджують принципову можливість використання цього підходу для управління реальним процесом виробництва біогазу в метантенку.

Наукова новизна. Показано, що задача управління отриманням біогазу в метантенку може бути вирішена на основі підходу навчання з підкріпленням, реалізованого для неперервної нелінійної динамічної системи. Система може бути реалізована у вигляді двох нейронних мереж в архітектурі актор-критик. Для побудови модуля критики й модуля нейроконтролера системи адаптивного управління запропоновано використовувати архітектуру нейронної мережі на основі спеціального виду рекурсивної нейронної мережі – LSTM нейронних мереж, що раніше не використовувалася для управління виробництвом біогазу в метантенку.

Практична значимість. Запропонований метод побудови системи управління метантенку призведе до підвищення ефективності виробництва біогазу як альтерна-

тивного поновлюваного джерела енергії. Розширення сфери застосування метантенків також вплине на вирішення проблеми утилізації промислових і побутових відходів життєдіяльності людини.

Ключові слова: *управління процесом, управління метантенком, навчання з підкріпленням, нейронні мережі, мережі LSTM, системи адаптивної критики*

LSTM-сети для управления метантенком

В. С. Есилевский[1], *С. В. Дядюн*[2,3], *В. Н. Кузнецов*[4]

1 – Национальный университет радиоэлектроники, г. Харьков, Украина, e-mail: yes_v_s@ukr.net
2 – Национальный университет городского хозяйства имени А. М. Бекетова, г. Харьков, Украина
3 – Национальный университет имени В. Н. Каразина, г. Харьков, Украина, e-mail: daulding@ukr.net
4 – Старооскольский технологический институт имени А. А. Угарова, „Национальный исследовательский технологический университет „МИСиС“, г. Старый Оскол, Белгородская обл., Российская Федерация, e-mail: 1937vnk@gmail.com

Цель. Исследование возможности применения систем искусственного интеллекта на основе нейронных сетей для повышения эффективности технологического процесса производства биогаза в метантенке путем оптимизации нелинейной системы управления.

Методика. Математическое и имитационное компьютерное моделирование для построения системы адаптивного управления на основе нейронной LSTM-сети с использованием методов обучения с подкреплением. Исследование сходимости процесса обучения нейронной сети на нелинейной математической модели технологического процесса.

Результаты. Проведенный вычислительный эксперимент показал сходимость процесса обучения за 600 эпизодов на модели с двумя управляющими каналами и четырьмя каналами измерения и эффективность работы предложенной адаптивной системы управления для программной модели технологического процесса. Полученные результаты подтверждают принципиальную возможность использования этого подхода для управления реальным процессом производства биогаза в метантенке.

Научная новизна. Показано, что задача управления получением биогаза в метантенке может быть решена на основе подхода обучения с подкреплением, реализованного для непрерывной нелинейной динамической системы. Система может быть реализована в виде двух нейронных сетей в архитектуре актер-критик. Для построения модуля критики и модуля нейроконтроллера системы адаптивного управления предложено использовать архитектуру нейронной сети на основе специального вида рекурсивной нейронной сети – LSTM нейронных сетей, которая ранее не использовалась для управления производством биогаза в метантенке.

Практическая значимость. Предложенный метод построения системы управления метантенком приведет к повышению эффективности производства биогаза как альтернативного возобновляемого источника энергии. Расширение сферы применения метантенков также окажет существенное влияние на решение проблемы утилизации промышленных и бытовых отходов жизнедеятельности человека.

Ключевые слова: *управление процессом, управление метантенком, обучение с подкреплением, нейронные сети, сети LSTM, системы адаптивной критики*

Рекомендовано до публікації докт. техн. наук Л. О. Кириченком. Дата надходження рукопису 10.01.19.