

ЗАСТОСУВАННЯ НЕЙРОННИХ МЕРЕЖ МГУА ДЛЯ ВИЯВЛЕННЯ ЗМІНИ УМОВ РОБОТИ НАСОСНОЇ СТАЦІЇ ВОДОПОСТАЧАННЯ, ЗУМОВЛЕНИХ ВПЛИВОМ СЕЗОННИХ ЧИННИКІВ**Л. В. Давиденко**

Луцький національний технічний університет

ORCID: 0000-0002-0176-2045**В. А. Давиденко, Н. В. Давиденко, С. О. Куницький**

Національний університет водного господарства та природокористування

ORCID: 0000-0001-6183-9879; 0000-0002-9722-745X; 0000-0003-0318-6149

Розглянуто питання ідентифікації циклічних змін фактичних умов роботи об'єктів водопостачання. Обґрунтовано доцільність застосування методів інтелектуального аналізу даних для виявлення закономірностей у формуванні режиму водоподачі. Як режимний показник процесу водопостачання використано добовий графік витрати води з мережі водопостачання. Для ідентифікації змін режиму водоподачі виконано формування класів подібних добових графіків витрати води. Як алгоритм класифікації використано метод «один проти всіх». Для структурно-параметричної ідентифікації моделі класифікатора добових графіків витрати води застосовано нейронні мережі МГУА. Пошук оптимальної моделі виконано в трьох класах нейронних мереж. Для перевірки моделей застосовано стратегію перехресної перевірки по К-блоках. Кращу структуру нейронної мережі вибрано за критерієм регулярності. Результати верифікації моделі класифікатора засвідчили достатньо високу якість класифікації. Побудований класифікатор забезпечує визначення належності добового графіка витрати води до одного з типових класів, що дає змогу ідентифікувати зміну умов роботи об'єкту водопостачання.

Ключові слова: добовий графік витрати води, розпізнавання образів, алгоритм кластеризації, метод групового урахування аргументів, нейронні мережі, модель класифікатора.

АКТУАЛЬНІСТЬ РОБОТИ. Зростання енергетичних витрат та їх впливу на навколишнє середовище висуває проблему енергоефективності в число актуальних у будь-якій сфері. Ефективність режиму роботи будь-якого технологічного об'єкту залежить від низки внутрішніх та зовнішніх чинників. Зміна зовнішніх чинників зумовлює зміну фактичних умов роботи об'єкту. Виявлення та формалізований опис цих змін є першим етапом складної процедури організації ефективного планування режимів роботи об'єкту та контролю їх енергоефективності.

В системі централізованого водопостачання (СЦВ) міст найбільшим споживачем електроенергії є підсистема насосних станцій (НС). Тому забезпечення ефективного режиму їх роботи відповідно до фактичних умов є одним із кроків вирішення проблеми енергоефективності СЦВ. Енергоефективність режиму роботи НС залежить від експлуатаційних характеристик насосних агрегатів (НА), їх заводських напірно-витратних і енергетичних характеристик, числа і схеми включення одночасно працюючих НА, способу управління електромеханічними і технологічними параметрами [1-3]. Сумарні витрати електроенергії на водоподачу залежать від обсягу води, що перекачується, і напору, що створюється при її перекачуванні.

Для систем міського водопостачання домінуючу роль відіграє господарсько-питне водоспоживання, яке характеризується нерівномірністю і формується під впливом багатьох, часто некерованих, чинників. Графік водоспоживання, що змінюється в часі, визначає режим водоподачі НС. Як відомо, напірна характеристика опору водопровідної мережі залежить від гідравлічних параметрів мережі (довжина та діаметри трубопроводів, геодезичні відмітки НС і споживачів), які майже не змінюються в часі, та господарської діяльності споживача, яка має випад-

ковий характер [3]. Водоспоживання є одним з зовнішніх чинників, що визначає поточний водорозбір (витрату води (ВВ) з мережі водопостачання), а отже ефективність режиму водоподачі та разом з іншими технічними та технологічними чинниками - ефективність електроспоживання [4]. Водоспоживання, а отже й водорозбір мають яскраво виражені коливання не лише впродовж доби, що зумовлено щоденним ритмом життя населення, а й впродовж місяця та року [1, 4, 5], що зумовлені зміною укладу життя населення у вихідні, святкові та робочі дні; відключенням опалення, гарячого водопостачання; впливом сезонних та кліматичних (температура повітря, кількість опадів) чинників [5, 6]. Основним критерієм, що визначає ефективність роботи СЦВ та її структурних елементів є забезпечення споживача водою в кількості, рівній його потребі. Отже, для забезпечення ефективно організації технологічного процесу водопостачання, ефективного режиму електроспоживання, формування управлінських впливів необхідне урахування зміни фактичних умов роботи об'єктів водопостачання, що визначаються зміною характеру водоспоживання, зумовленого впливом сезонних та соціальних чинників.

Метою статті є формування принципів виявлення та ідентифікації зміни фактичних умов роботи НС водопостачання на основі інтелектуального аналізу даних, отриманих з системи моніторингу режимів водоподачі в СЦВ.

МАТЕРІАЛИ І РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ. Сучасні технологічні об'єкти це складні системи. Їх стан та режим роботи описують великою кількістю вхідних і вихідних параметрів. При цьому функціональні залежності між цими параметрами можуть бути невідомі, або ж математичний опис фізичного зв'язку між ними є складним. Складовою вирішення проблем, пов'язаних із забезпеченням ефективних

режимів роботи складних технологічних об'єктів, організацією їх контролю, своєчасним реагуванням на нештатні ситуації тощо, є завдання створення систем моніторингу режимів роботи технологічних систем та їх об'єктів. Сучасний розвиток інформаційних технологій, технічних засобів автоматизації та комп'ютеризації забезпечує можливість накопичення інформації про фактичні параметри та характеристики режимів роботи всієї системи та кожного її об'єкту. Вивчення цієї інформації на основі використання методів інтелектуального аналізу даних забезпечує виявлення прихованих закономірностей формування технологічних режимів.

Одним із підходів ідентифікації зміни фактичних умов роботи обладнання є аналіз профілів характеристик режиму (режимних показників) на рівні системи чи її об'єкту, спрямований на вилучення з даних максимального ступеня інформації з метою планування та поліпшення режимів роботи об'єктів, планування енергетичних витрат тощо [7]. Профіль режимного показника це графік його зміни в часі.

Режимним показником процесу водопостачання є добовий графік витрати води (ГВВ). Нехай кожен добовий ГВВ є реалізацією $q_i(t)$ випадкового процесу водоспоживання $Q(t)$, $t \in T$ на річному інтервалі спостереження T , де, T – період спостереження випадкового процесу $T \in [1; 8760]$, i – номер доби, $i \in [1; 365]$, t – номер години доби $t \in [0; 24]$.

Для ідентифікації циклічних змін процесу водоподачі в [4, 8] запропоновано виконати пошук класів подібних добових ГВВ з використанням теорії розпізнавання образів. Об'єктами кластеризації є добові ГВВ; ознаками – характеристики профіля добового ГВВ та його нерівномірності; кластером – група добових ГВВ з подібним профілем.

Нехай кожен профіль добового ГВВ $d_i = q_i(t)$ з множини $D = \{d_1, d_2, \dots, d_n\}$ описано множиною ознак $X = \{x_1, x_2, \dots, x_m\}$, тобто, кожен добовий ГВВ $d_j \in W$, $j = \overline{1, n}$ представлений описом виду $d_j = (x_{ij} \in X)$, $i = \overline{1, m}$. Для опису добового ГВВ виконано його характеристику за об'ємами водоподачі та показники нерівномірності: x_1 – об'єм добової ВВ; x_2 – ВВ у нічний період доби; x_3 – ВВ у денний період доби; x_4 – ВВ у проміжний період ранішнього та вечірнього водоспоживання; x_5 – максимальне значення добової ВВ; x_6 – мінімальне значення добової ВВ; x_7 – мінімальне значення ВВ протягом денних годин доби; x_8 – середньодобове значення ВВ з мережі; x_9 – середнє значення ВВ протягом денних годин; x_{10} – дисперсія; x_{11} – дисперсія за день. Необхідно: 1) синтезувати підмножину $\{x_n^*\} = X^* \subset X^m$, $\eta = 1, m^*$, $m^* < m$ із наведених ознак опису ГВВ (ансамбль інформативних ознак [9]), найкращу за заданим критерієм оптимальності, яка б дозволила: 2) класифікувати всі ГВВ D на $k < n$, $k = \overline{1, K}$ однорідних груп; 3) побудувати класифікатор для визначення належності нових добових ГВВ до одного з типових класів.

Завдання кластеризації полягає у визначенні груп добових ГВВ, які є найбільш близькими за деяким

критерієм. Зважаючи на відсутність інформації щодо можливих класів доцільним є використання процедур розпізнавання образів «без учителя», зокрема, кластерного аналізу (КА) [4]. На початковому етапі слід застосувати алгоритм ієрархічного КА як попередній для визначення числа класів, а потім – метод *k-means* [8]. Це дозволить виявити приховані закономірності та сформулювати групи подібних добових ГВВ, тобто навчальну вибірку, в якій для кожного добового ГВВ d_j відома його класифікація $y \in [1, k]$, де k – число класів. Однак КА не дає ні правил, ні критеріїв оцінки якості класифікації.

Наступний етап – застосування методів розпізнавання образів «із учителем», який дає можливість класифікувати добові ГВВ на основі «навчання» за навчальними вибірками і побудувати класифікаційні функції для визначення належності добових ГВВ до одного з класів. Застосування класифікаційних функцій за своєю суттю відображає прогнозування імовірності належності об'єкту до деякого класу. Отже, правильність їх побудови разом з багатьма іншими факторами визначається об'ємом вибірки, утвореної об'єктами, що належать до одного кластеру.

Зважаючи на можливість утворення малих вибірок, для побудови класифікаційних функцій доцільно є застосування підходу самоорганізації моделей, який може бути використаний в умовах мінімального обсягу апіорної інформації. Також можливим є прогнозування без урахування деяких визначальних чинників. Це пояснюється тим, що в складних системах фактори корельовані між собою, отже, вимір одного фактора містить інформацію про інші фактори, пов'язані з вимірюваним. Таким чином, завершальним кроком процедури класифікації добових ГВВ є побудова правил класифікації, здатних до самоорганізації. Згідно [9] погляд на кластеризацію як на модель дозволяє перенести в теорію КА основні поняття й принципи теорії самоорганізації моделей на основі МГУА. В основі даного методу лежать не тільки принципи навчання з учителем, а й самоорганізація, характерна для систем, здатних до самонавчання, що дозволяє виконувати спрямований адаптивний пошук оптимальних рішень [10] (автоматично знаходити взаємозалежності в даних, вибирати оптимальну структуру кластерів, налаштувати (оцінювати) параметри). За допомогою МГУА, на підставі значень категоріальних і метричних змінних добовий ГВВ може бути зарахований до одного з заданих класів. З позицій загальної постановки задачі КА [9], має бути вирішене завдання, яке полягає у 1) конструюванні оптимального за заданою системою критеріїв ансамблю інформативних ознак та 2) одночасній побудові оптимальної за заданими критеріями кластеризації [10].

Таким чином, алгоритм кластеризації розглядається як функція $f: X^m$, яка будь-якому добовому ГВВ $d \in X^m$ ставить у відповідність номер кластера $y \in \mathfrak{R} = \{r_1, r_2, \dots, r_k\}$.

Необхідно знайти модель правила класифікації (модель класифікатора):

$$\bar{y}(d) = f(X_M, \Theta_M), \quad (1)$$

Де Θ_M - вектор невідомих параметрів;
 X_M - підматриця, $\dim X_M = n \times M$, $M \leq m$.

Завдання структурно-параметричної ідентифікації моделі (1) потребує вирішення задачі оптимізації множини аргументів матриці X_M , перетворених в заданому класі функцій, задачі оптимізації значень параметрів Θ_M функції f , що залежить від множини аргументів X_M . Отримана модель класифікатора добових ГВВ повинна задовольняти умову оптимального значення вибраного критерію відбору.

Вважається, що наявна множина ознак є достатньою для поділу класів [11], тобто виконується гіпотеза про компактність класів $R_i \cap R_j = \emptyset$, де R_i, R_j , $i \neq j$ - області розміщення більшості об'єктів в просторі $\mathbf{X} \in \mathfrak{R}^m$ вихідних ознак x_1, x_2, \dots, x_m . В якості вихідних задано множину спостережень d_1, d_2, \dots, d_n , $d \in D$, які є вектор-рядками матриці даних $\mathbf{X} \in \mathfrak{R}^m$, $\dim \mathbf{X} = n \times m$, вектор-стовпці якої є змінні x_i , $i = 1, \dots, m$ (m - кількість вхідних змінних; n - кількість спостережень). Задані також вектори фіктивної змінної - виходу \bar{y} , $\dim \bar{y} = n \times 1$, елементи \bar{y}_i , $i = \overline{1, n}$ якого задають для різних спостережень приналежність об'єктів d_1, d_2, \dots, d_n до класів $\{\bar{r}_1, \bar{r}_2, \dots, \bar{r}_k\}$, $k = \overline{1, K}$ [11]. Передбачається, що $y_j \in \mathfrak{R}$ є сумою значень відомої функції $\bar{y}(d_j)$, $j = \overline{1, n}$ та невідомої випадкової величини ξ_j з нульовим математичним сподіванням і обмеженою дисперсією $\sigma^2 \leq \infty$, $y_i = \bar{y}(d_i) + \xi_i$ [11]. Шум ξ і функція $\bar{y}(d)$ не залежать один від одного, $d \in \mathbf{X}^m$; де вектор $\xi = (\xi_1, \dots, \xi_n)^T$, $\dim \xi = n \times 1$.

Матриця вихідних даних має m вихідних змінних і n спостережень. Коли загальна кількість спостережень щодо числа класів невелика, використовується підхід «штучного» збільшення обсягу навчальної вибірки. Матриця ознак може містити не тільки ознаки, що спостерігаються, але і їх нелінійні перетворення [11]. Передбачається, що функції перетворень $f_j(d) \in \Phi$, забезпечують перехід з простору ознак розмірності m в простір ознак розмірності $f_j(d): \mathfrak{R}^m \rightarrow \mathfrak{R}$, де $\Phi = \{f_j(d)\}_{j=1, m}$ - задана скінченна множина функцій перетворень $d \in \mathbf{X}^m$ [11].

Завдання побудови вирішального правила в МГУА представляється як завдання індуктивної побудови моделі, що ускладнюється в процесі роботи алгоритму. Нехай моделі $y(d_j)$ лінійно залежать від змінних $f_j(x_i)$, $i = \overline{1, m}$, де $d_j \in \mathfrak{R}$ - елементи вектор-рядка $d \in \mathbf{X}^m$. Шуканим правилом класифікації добових ГВВ буде залежність, представлена поліномом Колмогорова-Габбора [11]:

$$\bar{y}(d) = \theta_0 + \sum_{i=1}^M \theta_i x_i + \sum_{i=1}^M \sum_{j=1}^M \theta_{ij} x_i x_j + \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^M \theta_{ijk} x_i x_j x_k, \quad (2)$$

Функція $\bar{y}(d)$ є нелінійною за вихідними змінними і лінійною за параметрами θ [11]. Відшукується краща та загальна для всіх $d=1, \dots, n$ структура моделі класифікатора $\bar{y}(d)$, яка на параметрах Θ_M , отриманих за навчальними вибірками, найбільш точно з точки зору зовнішнього критерію представить складові (1). Як класифікатор використовується краща індуктивна модель, що отримана за багаторядним алгоритмом синтезу моделей МГУА в результаті послідовного випробування моделей та забезпечує мінімальне значення зовнішнього критерію.

Як алгоритм синтезу моделей обрано нейронні мережі МГУА. Пошук моделі класифікатора виконувався в таких класах: клас 1 - лінійна НМ; клас 2 - поліноміальна НМ першого порядку; клас 3 - поліноміальна НМ другого порядку.

Для побудови класифікатора використано програмну оболонку GMDH Shell DS 3.8.8 [12].

Алгоритм класифікації GMDH Shell використовує метод «один проти всіх» (One-vs-All). Це дає змогу звести багатокласову класифікацію до бінарної класифікації. Для заданої навчальної вибірки будується K бінарних класифікаторів (за кількістю класів), кожен з яких відокремлює один з класів від всіх інших [13]. Кожен клас a_j з'являється з іншими ($k-1$) класами, тобто в кожному з j випадків вибір здійснюється з двох варіантів: «клас a_j » і «не клас a_j ». One-vs-All метод полягає в навчанні K класифікаторів за принципом [14]:

$$f_i(x) = \begin{cases} \geq 0, & \text{якщо } y(x) = i \\ < 0, & \text{якщо } y(x) \neq i \end{cases} \quad (3)$$

які відокремлюють кожен клас від інших. Для кожного $x \in X$ обчислюються всі класифікатори. Кожен класифікатор навчається на тренувальній вибірці:

$$\{(x, 1) | y = i, (x, y) \in S\} \cup \{(x, -1) | y \neq i, (x, y) \in S\}, i = 1, k \quad (4)$$

де i - порядковий номер бінарного класифікатора

Кожен з бінарних класифікаторів через точність, вірогідність тощо визначає ступінь довіри до заданого класу (що прецедент належить класу k). Підсумкове рішення по всіх класах приймається за схемою «переможець забирає все» - переможцем вважається клас, який має максимальне значення функції F . Класифікація довільного вектору ознак $X \in \mathfrak{R}^m$ визначається голосуванням [14]:

$$F(x) = \arg \max_{i=1, \dots, N} f_i(x) \quad (5)$$

Тобто, вибирається клас, що забезпечує найбільший ступінь впевненості.

Щоб уникнути проблеми перенавчання (коли модель добре працює на тестових прикладах, але погано на реальних даних), для перевірки моделі застосовано стратегію перехресної перевірки по

К-блоках [15] (вся навчальна вибірка ділиться на *k*-частин, алгоритм класифікації навчається на всій навчальній вибірці, крім частини *i*, після чого, тестується на *i*-й частині).

Для оцінки якості класифікації використовуються метрики [15]: точність, повнота, метрика Ван Різбергена (F-міра). Точність і повнота не залежать від співвідношення розмірів класів і коректно відображають якість роботи алгоритму. Чим вища точність і повнота, тим краща якість класифікатора. F-міра – гармонічне середнє між точністю та повнотою [16] і є оцінкою якості класифікатора. Оцінка F досягає кращого значення, що означає досконалу точність і чутливість, при значенні 1 і прагне до нуля, якщо точність або повнота прагне до нуля. Оцінка F може

забезпечити більш реалістичну оцінку продуктивності тесту, використовуючи точність та повноту. Для аналізу правильно (чітко) визначених результатів використана метрика акуратність - опис систематичних помилок, показник статистичного упередження. Акуратність використовується як статистичний показник того, наскільки добре тест бінарної класифікації правильно визначає або виключає стан. Разом з часткою правильних відповідей обчислюється базовий рівень - частку правильних відповідей алгоритму, що завжди видає найбільш потужний клас.

Результати верифікації моделей-класифікаторів різного типу за сезоном для навчальної та перевіркової вибірки наведено в табл. 1.

Таблиця 1 – Результати верифікації моделей класифікатора

Клас моделі	Навчальна вибірка			Екзаменаційна вибірка			Критерій регулярності
	Вірно класифікованих	Невірно класифікованих	Зважена F-міра	Вірно класифікованих	Невірно класифікованих	Зважена F-міра	
1	246 (84,5%)	45 (15,5%)	0,839	63 (86,3%)	10 (13,7%)	0,858	0,03
2	266 (91,4%)	25 (8,9%)	0,913	66 (90,4%)	7 (9,6%)	0,906	0,0021
3	277 (95,2%)	14 (4,8%)	0,952	69 (94,5%)	4 (5,5%)	0,945	0,002

За результатами випробувань вибрано модель класифікатора, що забезпечує кращі характеристики результатів класифікації за критерієм регулярності. Аналіз отриманих результатів дозволяє стверджувати: кращі результати забезпечує поліноміальна НМ МГУА другого порядку:

$$Y = -0,025 + x_{11} \cdot 0,051 + N_2 \cdot 1,095 \quad (6)$$

де

$$N_2 = 0,042 - x_5 \cdot 0,094 + N_4 \cdot 0,841$$

$$N_4 = 0,0213 - x_1 \cdot 0,453 - x_1 \cdot x_7 \cdot 0,248 + x_1^2 \cdot 0,117 + x_7 \cdot 0,277 + x_7^2 \cdot 0,081 \quad (7)$$

Ефективність роботи моделі-класифікатора за сезоном характеризує матриця класифікації для навчальної та перевіркової вибірки (табл. 2). Результати верифікації класифікатора свідчать про достатньо високу якість класифікації.

Таблиця 2 – Матриця класифікації класифікатора

Фактична належність до класу	Навчальна вибірка						Екзаменаційна вибірка					
	Класифіковано як					Повнота	Класифіковано як					Повнота
	1	2	3	4	Разом		1	2	3	4	Разом	
1	74	4	0	0	78	0,949	16	0	0	0	16	1
2	3	115	0	0	118	0,975	1	36	0	0	36	0,972
3	0	0	42	0	42	1	1	0	7	0	8	0,875
4	0	4	3	46	52	0,868	0	2	0	11	13	0,846
Разом	77	123	45	46	291		18	37	7	11	73	
Точність	0,961	0,935	0,933	1			0,889	0,946	1	1		
F-міра	0,955	0,954	0,966	0,929	0,952		0,941	0,959	0,933	0,917	0,945	
Базовий рівень	0,732	0,595	0,856	0,818	0,405		0,781	0,507	0,89	0,822	0,493	
Правильність	0,976	0,962	0,99	0,976	0,952		0,973	0,959	0,986	0,973	0,945	

Значимість змінних (табл. 3) визначається за ступенем їх впливу на середньоквадратичне відхилення моделі класифікатора: змінні в моделі одну за одною замінюють їх середнім значенням і вимірюють середньоквадратичну похибку моделі [15]:

$$Impact = (R_{var} - R_{ori}) / (R_{all} - R_{ori}) * 100\% \quad (8)$$

де R_{var} - RMSE змінної, яка розглядається;

R_{all} - RMSE моделі, де всі змінні замінені їх середнім значенням;

R_{ori} - RMSE з нульовим впливом; Оскільки змінна в моделі множиться на іншу, підноситься до квадрату, то вплив перевищує 100%.

За допомогою даного класифікатора на основі значень ознак об'єкт класифікації (добовий ГВВ) може бути зарахований до одного з типових класів.

Таблиця 3 – Значимість змінних моделі

Змінна	RMSE моделі за тестом важливості (R_{var})	Значимість змінних, Impact, %
X_1	0,574	162,5
X_7	0,5	126,5
X_5	0,265	30,21
X_{11}	0,246	11,64

ВИСНОВКИ. Ідентифікація циклічних змін фактичних умов роботи технологічного об'єкту вимагає

моніторингу параметрів режиму роботи та чинників зовнішнього середовища. Аналіз добових ГВВ із застосуванням теорії розпізнавання образів дає змогу виявити приховані закономірності у їх формуванні залежно від сезону, а отже встановити факт зміни фактичних умов роботи об'єкту водопостачання. Застосування методів самоорганізації моделей, зокрема, МГУА нейронних мереж, дає змогу автоматично здійснити структурно-параметричну ідентифікацію моделі класифікатора. Побудований класифікатор забезпечує визначення належності нових добових ГВВ до одного з типових класів. Застосування запропонованих принципів є основою ефективного планування режимів роботи НС водопостачання.

ЛІТЕРАТУРА

1. Євтушенко А. О., Неня В. Г., Сотник М. І., Хованський С. О. Визначення оптимального складу насосної станції системи комунального водопостачання. *Вісник КДПУ ім. М. Остроградського*. 2008. Вип. 4/2008 (51). Ч. 1. С. 158–162.
2. Лезнов Б. С. Энергосбережение и регулируемый привод в насосных и воздуховодных установках. Москва: Энергоатомиздат, 2006. 360 с.
3. Бойко В. С., Неня В. Г., Сотник М. І., Хованський С. О. Аналіз частотного регулювання відцентрових насосів водопостачання з метою енергозбереження. *Вісник КДПУ ім. М. Остроградського*. 2009. Вип. 4/2009 (57). Ч. 1. С. 168–171.
4. Розен В. П., Давиденко Л. В., Давиденко Н. В. Урахування циклічних змін процесу водоподачі під час контролю ефективності електроспоживання насосних станцій. *Енергетика: економіка, технології, екологія*. 2018. Вип. 2. С. 68–74.
5. Алексеева Ю. А., Коренькова Т. В. К вопросу построения модели водопотребления в гидротранспортных комплексах. *Вісник КДПУ ім. М. Остроградського*. 2008. Вип. 4/2008 (51). Ч. 1. С. 136–139.
6. Шушкевич Е. В. Эффективное управление системой подачи и распределения воды Московского мегаполиса. *Водоснабжение и санитарная техника*. 2011. № 1. С. 24–30.
7. Teiwes, H., Blume, S., Herrmann, C., Rцssinger, M., Thiede, S.: Energy Load Profile Analysis on Machine Level. *Procedia CIRP*. 2018. vol. 69. pp. 271–276.
8. Давиденко Н. В. Формування груп однотипних добових графіків витрати води з мережі водопостачання методом К-середніх. *Науково-технічний збірник «Комунальне господарство міст»*. 2018. Вип. 142. С. 8–13.
9. Сарычева Л. В. Объективный кластерный анализ данных на основе МГУА. *Проблемы управления и информатики*. 2008. № 2. С. 86–104.
10. Осипенко В. В. Два підходи до розв'язання задачі кластеризації у широкому сенсі з позицій індуктивного моделювання. *Енергетика і автоматика*. 2014. №1. С. 83–97.
11. Кондрашова Н. В., Павлов А. В., Павлов А. В., Павлов В. А. Адаптивная система классификаторов МГУА. *Scientific Journal «ScienceRise»* 2014. №3/2(3). С. 68–73.
12. GMDH Shell for Data Science. URL: <https://gmdhsoftware.com/> (Дата звернення: 17.03.2021).
13. David M. J. Tax and Robert P. W. Duin Using two-class classifiers for multiclass classification. *16th International Conference on Pattern Recognition (ICPR)*. Vol. 2. 2002. P. 124–127.
14. Карасиков М. Е., Максимов Ю. В. Поиск эффективных методов снижения размерности при решении задач многоклассовой классификации путем её сведения к решению бинарных задач. *Машинное обучение и анализ данных*. 2014. Т. 1, № 9. С. 1273–1290.
15. GMDH Shell for Data Science uses the power of the GMDH (Group Method of Data Handling). URL: <http://www.gmdh.net/> (Дата звернення: 17.03.2021).
16. Осідач А. О. Кластеризація веб-документів на основі алгоритму хешування. *Вісник КрНУ імені Михайла Остроградського*. 2016. Вип. 4/2016 (99). С. 30–35

APPLICATION OF GMDH NEURAL NETWORKS TO DETECT CHANGES IN THE OPERATION CONDITIONS OF THE WATER SUPPLY PUMPING STATION CAUSED BY THE INFLUENCE OF SEASONAL FACTORS

L. Davydenko,

Lutsk National Technical University

ORCID: 0000-0002-0176-2045

V. Davydenko, N. Davydenko, S. Kunytskyi

National University of Water and Environmental Engineering

ORCID: 0000-0001-6183-9879; 0000-0002-9722-745X; 0000-0003-0318-6149

Purpose. Development of a mechanism identification of changes in the actual operation conditions of the water supply facility depending on the season. **Methodology.** Methods of intellectual analysis of profiles of mode indicators were used to identify regularities in the formation of the operation mode of the facility. The mathematical apparatus of pattern recognition algorithm with training was used to classify the profiles of mode indicators. Methods of self-organization of models of complex systems were used for structural and parametric identification of the classifier model. **Results.** The necessity of the analysis of the characteristics of operation mode of the pumping station of water supply, obtained from the monitoring system, to identify regularities in the formation of water supply was substantiated. The daily graph of water consumption from the water supply network was used as a mode indicator of the water supply process. Indicators of water supply volumes and graph unevenness were used to describe the water supply graph. The expediency of application of self-organization methods for solving the problem of classification and construction of the classifier model was substantiated. Structural and parametric identification of the classifier model for daily

water consumption graphs was performed using GMDH Neural Networks. The search for the optimal model was performed in three classes of neural networks. Better neural network structure was chosen on the basis of criterion of regularity. The K-block cross-validation strategy was used to test the models. The results of the verification of the classifier model showed the high quality of the classification. **Originality.** A method for identifying changes in the operation conditions of the water supply facility due to the influence of seasonal factors, based on the usage of the classifier of profiles of daily water consumption graphs from the water supply network, was proposed. **Practical value.** The constructed model of the classifier allows defining of belonging of a profile of the daily water consumption graph, received from monitoring system of the water supply mode, to one of typical classes. The fact of class change indicates a change in the actual operation conditions of the water supply facility. **Conclusions.** Analysis of daily graphs of water consumption using the pattern recognition algorithm makes it possible to establish the change in the actual operation conditions of the water supply facility caused by the influence of seasonal factors. The usage of neural networks of GMDH makes it possible to perform automatically structural and parametric identification of the classifier model. Application of the offered principles is a basis of effective planning of operation modes of pumping station of water supply. References 16, tables 3.

Keywords: daily water consumption graph, pattern recognition, clustering algorithm, group method of data handling, neural networks, classifier model.

REFERENCES

1. Yevtusheko, A.O., Nenia, V. H., Sotnyk, M. I., Khovanskyi, S. O. (2008), Vyznachennia optymalnogo skladu nasosnoi stantsii systemy komunalnogo vodopostachannia [Determining the optimal composition of the pumping station of the municipal water supply system]. *Transactions of Kremenchuk Mykhailo Ostrohradskyi State University*, 4 (51), 1, 158-162. [in Ukrainian]
2. Leznov, B. S. (2006), Energoberezhnie i reguliruemiy privod v nasosnykh i vozdukhoduvnykh ustanovkakh [Energy saving and variable speed drive in pumping and blowing systems]. Moscow: Energoatomizdat, 360 p. [in Russian]
3. Boiko, V. S., Nenia, V. H., Sotnyk, M. I., Khovanskyi, S. O. (2009), Analiz chastotnoho rehulivannia vidstetrovykh nasosiv vodopostachannia z metoiu energoberezhennia [Analysis of frequency control of centrifugal pumps of water supply for the purpose of energy saving]. *Transactions of Kremenchuk Mykhailo Ostrohradskyi State University*, 4 (57), 1, 168-171. [in Ukrainian]
4. Rozen, V. P., Davydenko, L. V., Davydenko, N. V. (2018), Urakhuvannia tsyklichnykh zmin protsesu vodopodachi pid chas kontroliu efektyvnosti elektrospozhyvannia nasosnykh stantsii [Taking into consideration of cyclic changes in the water supply process during control of the power consumption efficiency of pumping stations]. *Energy: economics, technology, ecology*, 2, 68-74. [in Ukrainian]
5. Alekseeva, Yu. A., Korenkova, T. V. (2008), K voprosu postroeniia modeli vodopotrebleniia v hidrotransportnykh kompleksakh [On the issue of construction a model of water consumption in hydrotransportation complexes]. *Transactions of Kremenchuk Mykhailo Ostrohradskyi State University*, 4 (51), 1, 136-139. [in Russian]
6. Shushkevych, E. V. (2011), Effectivnoe upravlenie sistemoi podachu i raspredileniia vody Moskovskogo megalopolisa [Effective management of the water supply and distribution system of the Moscow megalopolis]. *Vodonasabzhenie i sanitarnaia tekhnika*, 1, 24-30. [in Russian]
7. Teiwes, H., Blume, S., Herrmann, C., Rüssinger, M., Thiede, S. (2018), Energy Load Profile Analysis on Machine Level. *Procedia CIRP*, 69, 271-276.
8. Davydenko, N. V. (2018), Formuvannia hrup odnotypanykh dobovykh hrafikiv vytraty vody z merezhi vodopostachannia metodom K-serednykh [Formation of the groups of typical daily water consumption schedules from the water supply network by K-medium method]. *Naukovo-tekhnichnyi zbirnyk "Komunalne hospodarstvo mist"*, 142, 8-13. [in Ukrainian]
9. Sarycheva, L. V. (2008), Obektivnyi klasternyi analiz dannykh na osnove MHUA [Objective cluster data analysis based on GMDH]. *Problemy upravleniia i informatiki*, 2, 86-104. [in Russian]
10. Osypenko, V. V. (2014), Dva pidkhody do rozv'iazannia zadachi klasteryzatsii u shyrokomu sensii z pozytsii induktyvnoho modeliuвання [Two approaches to solving the clustering problem in a broad sense from the standpoint of inductive modeling]. *Energetyka i avtomatyka*, 1, 83-97. [in Ukrainian]
11. Kondrashova, N. V., Pavlov, A. V., Pavlov, A. V., Pavlov, V.A. (2014), Adaptivnaia sistema klassifikatorov MHUA [Adaptive system of GMDH classifiers]. *Scientific Journal «ScienceRise»*, 3/2(3), 68-73. [in Russian]
12. GMDH Shell for Data Science. URL: <https://gmdhsoftware.com/> (Accessed: Mar. 17, 2021).
13. David M.J. Tax, Robert P.W. Duin (2002), Using two-class classifiers for multiclass classification. *16th International Conference on Pattern Recognition (ICPR)*, 2, 124-127.
14. Karasikov, M. E., Maksimov, Yu. V. (2014), Poisk effektivnykh metodov snizheniia razmernosti pri reshenii zadach mnogoklassovoi klassifikatsii putem ieio svedeniia k resheniiu binarnykh zadach [Search for effective methods for reducing the dimensionality in solving problems of multiclass classification by reducing it to solving binary problems]. *Mashynnoe obuchenie i analiz dannykh*, 1, 9, 1273-1290. [in Russian]
15. GMDH Shell for Data Science uses the power of the GMDH (Group Method of Data Handling). URL: <http://www.gmdh.net/> (Accessed: Mar. 17, 2021).
16. Osidach, A. O. (2016), Klasteryzatsiia veb-dokumentiv na osnovi alorytmu kleshuvannia [Clustering of web documents based on the hashing algorithm]. *Transactions of Kremenchuk Mykhailo Ostrohradskyi National University*, 4 (99), 30-35. [in Ukrainian]