

## МЕТОД СИНТЕЗУ ЕМПІРИЧНИХ МОДЕЛЕЙ НА ЗАСАДАХ ГЕНЕТИЧНИХ АЛГОРИТМІВ

<sup>1</sup>М.І.Горбійчук, <sup>1</sup>М.І.Козуляк, <sup>2</sup>О.Б.Василенко, <sup>3</sup>І.В.Щупак

<sup>1</sup>ІФНТУНГ, 76019, м. Івано-Франківськ, вул. Карпатська, 15, тел. (0342) 504521,  
e-mail: gorb@nung.edu.ua

<sup>2</sup>Науково-дослідний і проектний інститут ВАТ «Укрнафта»,  
76019, м. Івано-Франківськ, вул. Північний бульвар, 2, тел. (0342) 776140,  
e-mail: felix122@rambler.ru

<sup>3</sup>ДП «Укрметртестстандарт», 03680, м. Київ, вул. Метрологічна, 4, тел. (044) 1265389,  
e-mail: shchupak@gmail.com

Розроблено метод синтезу емпіричних моделей, які подані у вигляді поліноміальної регресії, на засадах генетичних алгоритмів. Даний метод, на відміну від комбінаторного методу, значно скорочує кількість обчислювальних операцій, що відкриває можливості для синтезу складних моделей технологічних процесів. Такі моделі можуть бути використані у задачах оптимізації та діагностування технічних систем. Ефективність розробленого методу перевірена на прикладі синтезу емпіричних моделей процесу компримування природного газу.

Ключові слова: емпірична модель, генетичний алгоритм, критерій селекції, індуктивний метод самоорганізації, функція пристосування, модель оптимальної складності

Разработан метод синтеза эмпирических моделей, которые поданы в виде полиномиальной регрессии, на принципах генетических алгоритмов. Данный метод, в отличие от комбинаторного метода, значительно сокращает количество вычислительных операций, что открывает возможности для синтеза сложных моделей технологических процессов. Такие модели могут быть использованы в задачах оптимизации и диагностирования технических систем. Эффективность разработанного метода проверена на примере синтеза эмпирических моделей процесса сжатия природного газа.

Ключевые слова: эмпирическая модель, генетический алгоритм, критерий селекции, индуктивный метод самоорганизации, функция приспособления, модель оптимальной сложности

The method of synthesis of empiric models which are given as polynomial regression is developed, on principles of genetic algorithms. This method, unlike a combinatorics method, considerably abbreviates the amount of calculable operations that reveals possibilities for the synthesis of difficult models of technological processes. Such models can be used in the tasks of optimization and diagnosing of the technical systems. Efficiency of the developed method is tested on the example of synthesis of empiric models of natural gas compression process.

Keywords: empiric model, genetic algorithm, criterion of selection, objective method of self-organization, function of adaptation, model of optimum complication

В основі емпіричного моделювання процесів і явищ лежить фундаментальний метод найменших квадратів (МНК), суть якого полягає у наступному. Припустимо, що деякий процес або явище розглядаються як певна система, функціонування якої характеризується вектором вхідних величин  $\bar{x} = (x_1, x_2, \dots, x_n)^T$  і вихідною величиною  $y$ . Серед вхідних величин однією із змінних може бути і час  $t$ . За вхідними і вихідними величинами ведеться спостереження, у результаті якого отримано  $\bar{x}^{(1)}, \bar{x}^{(2)}, \dots, \bar{x}^{(N)}$  і  $Y^{(1)}, Y^{(2)}, \dots, Y^{(N)}$ , де  $N$  – кількість спостережень. Сукупність векторів  $\bar{x}^{(1)}, \bar{x}^{(2)}, \dots, \bar{x}^{(N)}$  утворюють матрицю, рядки якої  $\bar{x}^{(1)}, \bar{x}^{(2)}, \dots, \bar{x}^{(N)}$ , тобто

$$X = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_n^{(2)} \\ \dots & \dots & \dots & \dots \\ x_1^{(N)} & x_2^{(N)} & \dots & x_n^{(N)} \end{bmatrix}.$$

Матрицю  $X$  називають матрицею спостережень. Задача полягає у тому, щоб на основі спостережень за вхідними і вихідними величинами системи побудувати модель, яка якомога точніше наближує (апроксимує) вихід системи до виходу моделі. За критерій такого наближення у МНК вибирають суму квадратів відхилень виходу моделі від виходу системи.

У МНК припускають, що структура моделі відома, яку у більшості випадків, вибирають лінійною відносно її параметрів

$$y = a_0 f_0(\bar{x}) + a_1 f_1(\bar{x}) + \dots + a_k f_k(\bar{x}), \quad (1)$$

де  $a_0, a_1, \dots, a_k$  – параметри моделі. припустимо, що певним чином вибрано параметри моделі (1) і відома матриця спостережень  $X$ . Тоді у кожній точці експерименту можна обчислити вихід моделі

$$\bar{y} = F\bar{a}, \quad (2)$$

$$\text{де: } F = \begin{bmatrix} f_0(\bar{x}^{(1)}) & f_1(\bar{x}^{(1)}) & \dots & f_k(\bar{x}^{(1)}) \\ f_0(\bar{x}^{(2)}) & f_1(\bar{x}^{(2)}) & \dots & f_k(\bar{x}^{(2)}) \\ \dots & \dots & \dots & \dots \\ f_0(\bar{x}^{(N)}) & f_1(\bar{x}^{(N)}) & \dots & f_k(\bar{x}^{(N)}) \end{bmatrix};$$

$\bar{a} = (a_0, a_1, \dots, a_k)^T$  – вектор параметрів моделі (1);  $\bar{y} = (y^{(1)}, y^{(2)}, \dots, y^{(N)})$  – обчислене значення виходу моделі (1) у кожній точці спостережень. За відомими  $Y^{(i)}$  і  $y^{(i)}$ ,  $i = \overline{1, N}$ , можна обчислити критерій апроксимації

$$J(\bar{a}) = \sum_{i=1}^N (Y^{(i)} - y^{(i)})^2, \quad (3)$$

який подамо у вигляді векторного виразу

$$J(\bar{a}) = (\bar{Y} - \bar{y})^T (\bar{Y} - \bar{y}).$$

Враховуючи значення  $\bar{y}$ , що визначається формулою (2), отримаємо

$$J(\bar{a}) = (\bar{Y} - F\bar{a})^T (\bar{Y} - F\bar{a}). \quad (4)$$

Значення параметрів моделі (1) обчислюють за умови, що критерій апроксимації набуде мінімального значення, яке знаходять з умови

$$\frac{\partial J(\bar{a})}{\partial \bar{a}} = \bar{0}.$$

Обчисливши градієнт функції від правої частини рівняння (4) і враховуючи його нульове значення отримаємо рівняння

$$F^T F\bar{a} = F^T \bar{Y}, \quad (5)$$

яке називають нормальним рівнянням МНК.

Безпосередньо із рівняння (5) можна знайти

$$\bar{a} = (F^T F)^{-1} F^T \bar{Y}. \quad (6)$$

Використовувати формулу (6) можна лише тоді, коли розмірність вектора параметрів  $\bar{a}$  невелика і матриця  $F^T F$  є добре обумовленою. Якщо така умова не виконується, то для розв'язку рівняння (6) слід використовувати один із числових методів, наприклад, метод Гаусса зі зворотнім ходом [1].

У більшості випадків на вихід системи  $Y$  накладається перешкода. Якщо допустити, що вона є адитивною і має нормальний закон розподілу, то оцінки параметрів моделі є незміщеними і ефективними [2].

На практиці, як правило, структура моделі (1) невідома, що призводить до необхідності довільного вибору як числа функцій, так і вигляду самих функцій у моделі (1). Критерій (3), який використовується для визначення параме-

трів моделі (1) за формулою (6), є внутрішнім критерієм [3], і його використання призводить до помилкового правила: чим складніша модель, тим вона точніша. Складність моделі, наприклад поліноміальної, визначається числом членів і найвищим степенем полінома: чим більше членів полінома, тим меншим є значення критерію апроксимації (3).

Тому для вибору структури моделі (1) був запропонований індуктивний метод самоорганізації моделей [3], ідейну сторону якого визначає теорема Геделя. У відповідності до цієї теореми жодна система аксіом не може бути логічно замкнутою: завжди можна знайти таку теорему, для доведення якої необхідне зовнішнє доповнення – розширення початкової системи аксіом. Стосовно задачі визначення структури моделі (1) геделівський підхід означає застосування зовнішнього критерію, який дає можливість однозначного вибору єдиної моделі із заданого класу моделей. Критерій називають зовнішнім, якщо його визначення засновано на застосуванні нових даних, які не використовувались при синтезі моделі (1). Це означає, що всі дані, отримані у результаті експерименту, розбиваються на дві частини  $N_A$  і  $N_B$ . Перша із них – навчальна, а друга – перевірна.

У більшості випадків для вибору структури моделі використовують критерій регулярності

$$\Delta^2(B) = \frac{\sum_{i=1}^{N_B} (Y_i - y_i)^2}{\sum_{i=1}^{N_B} Y_i^2} \quad (7)$$

і мінімуму зміщення

$$\Delta^2(A, B) = \frac{\sum_{i=1}^N (y_i(A) - y_i(B))^2}{\sum_{i=1}^N Y_i^2}. \quad (8)$$

Якщо вибраний критерій регулярності (7), то вибирають такий розподіл даних експерименту [4]:  $N_A = 0,7N$  і  $N_B = 0,3N$ , а при виборі критерію (8) –  $N_A = 0,5N$  і  $N_B = 0,5N$ .

Реалізація індуктивного методу самоорганізації моделей здійснюється поетапно: перший етап генерація моделей-претендентів (у певному порядку підвищення складності); другий етап – відбір найкращої моделі за критерієм селекції (7 або 8).

Розрізняють три способи генерації моделей-претендентів.

Перший із них комбінаторний метод, за яким вибирають моделі із виразу (1) шляхом прирівнювання до нуля деяких його коефіцієнтів. Таким чином, отримуємо сукупність моделей. Вибір найкращої моделі здійснюється на основі одного із критеріїв селекції.

Другий спосіб відомий як метод групового урахування аргументів (МГУА), у якому генерація моделей здійснюється на основі багаторядної процедури. У першому ряду селекції

утворюють всі можливі пари аргументів, і для кожної із них знаходять часткові моделі, наприклад, у вигляді повного полінома. Із всіх часткових моделей вибирають  $K$  найкращих, за вибраним критерієм селекції. Із виходів цих  $K$  моделей знову утворюють комбінації всіх можливих пар, які є входами моделей другого ряду селекції. Для кожної із цих пар знову формують часткові моделі і т. д. Оцінка коефіцієнтів часткових моделей здійснюється за допомогою МНК. Нарощування рядів селекції відбувається до тих пір поки основний критерій селекції спадає.

Третій метод подібний до другого. Різниця лише у тому, що на кожному ряді селекції часткові моделі утворюють шляхом прирівнювання до нуля певного числа їх коефіцієнтів.

Недоліком комбінаторного методу селекції моделей є необхідність перебору великого числа моделей. Якщо вихідною моделлю вибрано повний поліном степеня  $m$ , то загальне число моделей-претендентів складає  $2^M - 1$ , де  $M$  – загальне число членів повного полінома степеня  $m$ . Навіть сучасні ЕОМ не здатні реалізувати такі алгоритми через значне число змінних і високий степінь полінома. МГУА породжує моделі, у яких фігурують проміжні змінні кожного із рядів селекції, що значно утруднює процес переходу до вхідних змінних системи, що моделюється. Сказане відноситься і до третього методу, оскільки він посуті є модифікацією МГУА.

Із усіх трьох методів найпривабливішим є комбінаторний метод, оскільки він дає можливість отримати модель, де аргументами виступають вхідні величини системи. Для зняття проблеми великої розмірності застосуємо генетичний підхід. Як емпіричну модель будемо розглядати поліном степеня  $m$

$$y = \sum_{i=0}^{M-1} a_i \prod_{j=1}^n x_j^{s_{ji}}, \quad (9)$$

де:  $M$  – кількість членів полінома;

$a_i$  – коефіцієнти полінома;

$s_{ji}$  – степені аргументів, які повинні задовольняти обмеженню

$$\sum_{j=1}^n s_{ji} \leq m.$$

Число членів  $M$  полінома (9) визначають за такою формулою [5]

$$M = \frac{(m+n)!}{m!n!}. \quad (10)$$

При комбінаторному методі синтезу моделі із повного полінома (9) отримують емпіричну модель, де частина параметрів приймає значення нуль. Інші параметри, будуть відмінні від нуля. Утворимо упорядковану структуру довжиною  $M$ , в якій на  $i$ -тому місці буде стояти одиниця або нуль в залежності від того чи параметр  $a_i$ ,  $i=1, M$  моделі (9) відмінний від нуля, чи нульовий. У теорії генетичних алгори-

змів така упорядкована послідовність має назву хромосоми або особи, а атомарний елемент хромосоми (одиниця або нуль) – це ген. Набір хромосом утворює популяцію. Важливим поняттям у теорії генетичних алгоритмів є функція пристосування, яка визначає ступінь пристосування окремих осіб у популяції. Вона дає змогу із всієї популяції вибрати особи, що є найбільш пристосованими, тобто такі, які мають найбільше (найменше) значення функції пристосування. У задачі синтезу емпіричних моделей функцією пристосованості виступає критерій селекції (7 або 8).

Таким чином, задачу синтезу емпіричної моделі сформуємо так: із початкової популяції хромосом шляхом еволюційного відбору вибрати таку, хромосому, яка забезпечує найкраще значення функції пристосування (мінімальне значення критерію селекції (7) або (8)).

Генетичний алгоритм складається із таких кроків [6].

K1. Формування початкової популяції (ініціалізація). На першому кроці роботи алгоритму випадковим чином формується популяція із  $I$  осіб, кожна із яких є хромосомою довжиною  $M$ . Число генів у хромосомі визначається за формулою (10).

K2. Оцінка пристосованості хромосоми у популяції. Для кожної хромосоми обчислюється критерій селекції (7) або (8). Здійснюється така процедура так. Якщо вибраний критерій селекції (7), то формуються матриці  $F_A$  і  $F_B$  розміром  $N_A \times M$  і  $N_B \times M$ . Із матриці  $F_A$  вилучається  $i$ -тий стовпець, якщо на  $i$ -тій позиції у хромосомі знаходиться нуль; у протилежному випадку  $i$ -тий стовпець залишається без змін. У результаті отримаємо матрицю  $F_A$ , із якої вилучено  $c_A$  стовпців, де  $c_A$  – кількість нулів у вибраній хромосомі із початкової популяції. Розмір такої матриці –  $N_A \times (M - c_A)$ . Аналогічним чином формується матриця  $F_B$  розміром  $N_B \times (M - c_A)$ . На множині точок  $N_A$  обчислюються ненульові коефіцієнти  $a_{A,i}$  моделі (9) шляхом розв'язку нормального рівняння Гауса (5), яке видозміниться наступним чином:

$$F_A^T F_A \bar{a}_A = F_A^T \bar{Y}_A, \quad (11)$$

де:  $\bar{a}_A = (a_{A,0}, a_{A,1}, \dots, a_{A, M-c_A-1})^T$  – вектор параметрів моделі, яка асоційована з черговою хромосомою із початкової популяції;

$\bar{Y}_A = (Y^{(1)}, Y^{(2)}, \dots, Y^{(N_A)})^T$  – вектор експериментальних даних на множині точок  $A$ .

За відомим коефіцієнтами  $a_A$  поліноміальної моделі на множині точок  $B$  обчислюють

$$\bar{y}_B = F_B \bar{a}_A. \quad (12)$$

Знаючи  $\bar{y}_B$ , за формулою (7) обчислюють функцію пристосування  $\Delta^2(B)$  для кожної хромосоми із початкової популяції. У результаті на кроці K2 отримують значення  $\Delta_j^2(B)$ ,  $j = \overline{1, I}$ .

У тому випадку, коли використовують критерій селекції (8) як функцію пристосованості, тоді складають рівняння (11), яке методом Гауса розв'язують відносно вектора параметрів  $\bar{a}_A$ . Після цього обчислюють  $\bar{y}_A = F_A \bar{a}_A$  і  $\bar{y}_B$  за формулою (12). Отримані значення  $\bar{y}_A$  і  $\bar{y}_B$  дають змогу знайти значення  $\Delta_j^2(A, B)$ ,  $j = \overline{1, I}$  для кожної хромосоми із початкової популяції.

K3. Перевірка умови зупинки алгоритму. Визначають

$$\Delta^2(B) = \min_j \Delta_j^2(B) \quad (13)$$

або

$$\Delta^2(A, B) = \min_j \Delta_j^2(A, B). \quad (14)$$

Якщо мінімальне значення (13) або (14) критерію селекції (7) або (8) не перевершує деякого додатного значення  $\varepsilon$ , то відбувається зупинка алгоритму. Зупинка алгоритму також може відбутися у випадку, коли його виконання не призводить до покращення функції пристосування або у тому випадку, коли алгоритмом уже виконано задане число ітерацій.

Після виконання однієї із трьох умов із популяції вибирається хромосома  $ch^*$ , для якої виконується умова (13) або (14). Ця хромосома задає структуру моделі оптимальної складності і формує матрицю  $F^*$  таким чином, що із початкової матриці  $F$  вилучаються стовпці, які асоційовані з нульовими генами хромосоми  $ch^*$ . Перерахунок параметрів моделі (9) здійснюється на множині всіх точок початкового масиву даних.

K4. Селекція хромосом. За розрахованими на другому кроці алгоритму значеннями функції пристосування здійснюється відбір тих хромосом, які будуть брати участь в створенні потомків для наступної популяції. Такий вибір проводиться у відповідності з принципом природного відбору, коли найбільші шанси у створенні нової популяції мають хромосоми з найкращим значенням функції пристосування, тобто такі, що забезпечують мінімальне значення критерію селекції (7) або (8). Найбільш поширеними методами селекції [6] є метод рулетки і метод турнірної селекції.

Суть метода рулетки у тому, що особі для нового покоління вибираються пропорційно до їх значень функції пристосованості. Кожна хромосома отримує у пулі родичів таку кількість своїх копій, яке визначається виразом [6]

$$E(ch_i) = P(ch_i) \cdot I,$$

де:  $P(ch_i)$  – ймовірність селекції хромосоми;

$I$  – кількість хромосом  $ch_i$ ,  $i = \overline{1, I}$  у популяції.

Величина  $P(ch_i)$  підраховується за формулою

$$P(ch_i) = \frac{R(ch_i)}{\sum_{i=1}^I R(ch_i)},$$

де  $R(ch_i)$  – значення функції пристосування  $ch_i$  хромосоми (у контексті задачі, що розглядається, це критерій селекції (7) або (8)). Очевидно, що метод рулетки можна застосовувати тоді, коли функція пристосованості додатна. Цей метод може використовуватись тільки у задачах максимізації.

Турнірний метод можна використовувати як у задачах максимізації, так і у задачах мінімізації функцій. При турнірній селекції всі хромосоми розбиваються на підгрупи з наступним вибором із кожної утвореної підгрупи хромосоми з найкращою пристосованістю. Підгрупи можуть мати довільний розмір, але частіше за все популяції ділять на підгрупи по 2 – 3 особи у кожній. На рис. 1 показана схема, яка ілюструє турнірний метод селекції для підгруп із  $z$  осіб.

K5. Формування нової популяції потомків здійснюється за допомогою двох основних операторів: схрещування і мутації. Слід зауважити, що оператор мутації відіграє другорядну роль у порівнянні з оператором схрещування. Це означає, що у генетичному алгоритмі схрещування проводиться майже завжди, а мутація – досить рідко. Вірогідність схрещування досить висока ( $0,5 \leq P_c \leq 1$ ), тоді як ймовірність мутації вибирається досить малою ( $0 \leq P_m \leq 0,1$ ).

Оператор мутації з ймовірністю  $P_m$  змінює значення гена в хромосомі на протилежне, тобто з 1 на 0 чи з 0 на 1. Ймовірність мутації  $P_m$  може змалюватись випадковим вибором числа із інтервалу  $[0;1]$  для кожного гена і відбором для виконання цієї операції тих генів, для яких розігране число виявиться меншим або рівним  $P_m$ . Мутація може здійснюватись як над пулом родичів, так і над пулом потомків.

Оператор схрещування складається із двох етапів (рис. 1). На першому етапі формуються підгрупи із  $z$  осіб звідки вибирається найкраща хромосома за критерієм селекції  $R(ch^*) = \min_i R(ch_i)$ . У результаті отримуємо нову популяцію хромосом, до якої застосовують оператор другого етапу.

На другому етапі здійснюється схрещування. Для цього із пулу родичів  $M(k)$  (рис. 1) випадковим чином з ймовірністю  $P_c$  утворюють пари у такий спосіб. Із популяції осіб ви-

падковим чином вибирається пара хромосом. Генерується випадкове число  $P_z$  із інтервалу  $[0; 1]$ , і якщо його значення не більше ніж  $P_c$ , то над парою хромосом здійснюється схрещування. У протилежному випадку пара хромосом залишається без зміни. Потім для кожної пари родичів розігрується позиція гена (локус) в хромосомі, яка визначає точку схрещування. Якщо хромосома кожного із родичів включає у себе  $M$  генів, то точка схрещування  $L_c$  – це натуральне число, яке менше  $M$ . Тому фіксація точки схрещування зводиться до випадкового вибору цілого числа із інтервалу  $[1; L_c - 1]$ . Дія оператора схрещування призводить до того, що із пари родичів утворюється нова пара потомків так: перший потомок у парі, хромосома якого на позиціях від 1 до  $L_c$  складається із ген першого родича, а на позиція від  $L_c + 1$  до  $M$  – із ген другого родича; другий потомок у парі, хромосома, якого на позиціях від 1 до  $L_c$  складається із ген другого родича, а на позиція від  $L_c + 1$  до  $M$  із ген першого родича.

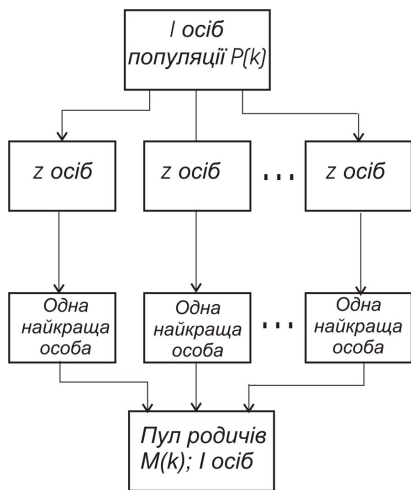


Рисунок 1 – Схема турнірної селекції для підгруп, що складаються із  $z$  осіб

Після виконання оператора схрещування відбувається перехід до К2.

Як приклад застосування розробленого методу розглянемо ідентифікацію статичних характеристик газоперекачувальних агрегатів (ГПА) природного газу.

Для вирішення задачі оптимального керування процесом компримування природного газу актуальною є побудова адекватних процесу математичних моделей на основі експериментального спостереження за роботою ГПА на режимах їх нормальної експлуатації.

Розглянемо  $i$ -й ГПА, як об'єкт, який характеризується керуючою дією  $x_i$  та впливами зовнішнього середовища  $z_{ij}$ , де  $j$  – номер зовнішнього впливу для  $i$ -го нагнітача;  $j = \overline{1, r}$ ;  $r$  – кількість зовнішніх впливів (рис. 2).

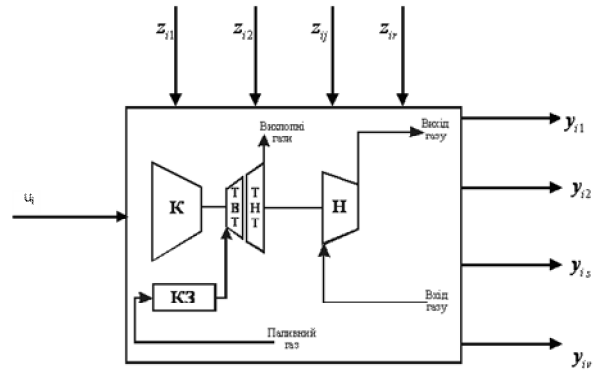


Рисунок 2 – Газоперекачувальний агрегат як об'єкт керування

Сукупність величин  $x_i$  та  $z_{ij}$  утворюють групу вхідних величин ГПА (об'єкта керування). Реакція керованого об'єкта на вхідні величини характеризується певними показниками роботи нагнітача – вихідними величинами  $y_{is}$ ,  $s = \overline{1, v}$ , де  $v$  – кількість вихідних величин.

Для нагнітачів з газотурбінним приводом (ГТП) вихідними величинами будуть [7]:

- витрата паливного газу;
- температура газу на виході із нагнітача;
- температура газів на виході ТНТ;
- продуктивність нагнітача.

Кожна із перелічених вихідних величин є функцією керуючої дії  $u_i$  та зовнішніх впливів  $z_{ij}$ ,

$$y_{is} = f_{is}(u_i, \bar{z}_i), \quad (15)$$

де  $\bar{z}_i$  – вектор зовнішніх впливів  $i$ -го нагнітача.

Керуючою дією є частота обертання вала нагнітача  $n_n$ .

Проведений аналіз літературних джерел і роботи компресорних станцій виявив [7], що для нагнітачів з газотурбінним приводом (ГТП) зовнішніми впливами є: температура газу на вході в нагнітач  $T_{in}$ , ступінь підвищення тиску газу  $\varepsilon$ , тиск газу на вході в нагнітач  $P_{in}$ , температура  $T_c$  та тиск  $P_c$  навколишнього середовища.

Тоді, для витрати паливного газу  $G$ , температури газу на виході нагнітача  $T_{out}$ , продуктивності нагнітача  $Q$  та температури газів на виході турбіни низького тиску  $T_v$  отримаємо:

$$G = f_1(n_n, T_{in}, \varepsilon, P_{in}, T_c, P_c),$$

$$T_{out} = f_2(n_n, T_{in}, \varepsilon, P_{in}, T_c, P_c),$$

$$Q = f_3(n_n, T_{in}, \varepsilon, P_{in}, T_c, P_c),$$

$$T_v = f_4(n_n, T_{in}, \varepsilon, P_{in}, T_c, P_c)$$

або

$$\{G, T_{out}, Q, T_v\} = f_i(n_n, T_{in}, \varepsilon, P_{in}, T_c, P_c),$$

$$i = \overline{1, 3}. \quad (16)$$

Для ідентифікації моделей (16) у роботі [5] запропонований комбінаторний метод самоорганізації. Використання цього методу передбачає, що число членів у моделі (9) визначається формулою (10), а загальне число моделей претендентів це  $2^M - 1$ . У табл. 1 наведена кількість моделей-претендентів для співвідношення (16), яке є поліномом (9) степеня  $m$ .

**Таблиця 1 – Число моделей-претендентів для  $n=7$**

Степінь полінома, $m$	Кількість членів у поліномі, $M$	Кількість моделей-претендентів, $2^M - 1$
1	8	2.55e+002
2	36	6.871947673500000e+010
3	120	1.329227995784916e+036
4	330	2.187250724783012e+099
5	792	2.604693137843693e+238

Як видно із табл. 1 число моделей претендентів дуже швидко зростає із ростом степеня полінома. Допустимо, що для вибору однієї моделі із всіх можливих затрачається 0,001 с, а степінь полінома (9)  $m = 5$ , тоді для повного перебору затрати машинного часу становитимуть приблизно 3,2 роки, що є нереальним (це при тому, що величина 0,001 с є заниженою!). З цієї причини у роботі [5] для ідентифікації моделей ГПА з застосуванням комбінаторного методу степінь полінома (9) не перевищувала 2.

Оскільки ГПА є складною системою з багатьма входами, то складність моделі такої системи визначатиметься саме кількістю аргументів і найвищим степенем полінома (9). Для таких об'єктів модель оптимальної складності лежить за межами лінійних, квадратичних і, навіть, кубічних поліноміальних моделей [4].

Співвідношення (16) інтерпретуються як поліноми (9) степеня  $m$  [5]. Тому методика побудови всіх чотирьох моделей ГПА є однотиповою. З цієї причини була вибрана модель, яка описує залежність температури на виході турбіни низького тиску від технологічних параметрів  $n_n, T_{in}, \varepsilon, P_{in}, T_c$  та  $P_c$ . Для побудови моделі були використані експериментальні дані, які отримані у ході експлуатації КС-3 Долинського лінійного виробничого управління магістральними трубопроводами ДП «Прикарпаттрансгаз». Вимірювання і реєстрація технологічних параметрів нагнітача природного газу здійснювались за допомогою штатних технічних засобів, якими оснащена компресорна станція КС-3. Вимірювались такі параметри, як тиск і температури на вході та виході ВЦН, частота обертання вала нагнітача, температура та тиск навколишнього середовища перепад тиску на конфузори, температури газів на виході тур-

біни низького тиску, витрата паливного газу. На КС-3 встановлені ГПА з приводним двигуном ДГ-90Л2 потужністю 16 МВт та відцентровим нагнітачем (ВЦН) ГПА-Ц1-16С/76-1,5, продуктивність нагнітача якого обчислювалась за такою формулою:

$$Q = A \sqrt{\frac{\Delta P_k Z_k R T_{in}}{P_{in} + 0,101}},$$

де:  $A$  – коефіцієнт витрати конфузора;  
 $\Delta P_k$  – перепад тиску на конфузори, МПа;  
 $Z_k$  – коефіцієнт стисливості газу;  
 $R$  – газова стала,  $\frac{Дж}{кг \cdot ^\circ K}$ .

Якщо величини  $T_{in}$  і  $P_{in}$  вимірювати відповідно у  $^\circ K$  та МПа, то об'ємна продуктивність нагнітача за параметрами входу буде мати

розмірність  $\frac{м^3}{хв}$ . Значення величин  $Z_k$  і  $R$

обчислювались за методикою, яка наведена у [8]. Залежність  $T_v = f_4(n_n, T_{in}, \varepsilon, P_{in}, T_c, P_c)$  апроксимувалась поліномом (9) степеня  $m = 4$ . Аналіз експериментальних даних, який здійснювався за допомогою розробленої авторами програми у інтерактивному режимі, засвідчив, що зміна тиску навколишнього середовища незначно впливає на зміну величини  $T_v$ ; значно сильніший вплив на  $T_v$  мають попередні значення ступеня підвищення тиску  $\varepsilon$ . Це пояснюється тим, що ГПА разом з прилеглими трубопроводами є динамічною системою, у якій відбувається накопичення газу, що спричинює не миттєву, а поступову зміну тиску природного газу на виході ВЦН. Тому кінцевою моделлю була вибрана модель (9), для якої, а  $X_1 = n_n^{(i)}, X_2 = \varepsilon^{(i)}, X_3 = \varepsilon^{(i-1)}, X_4 = \varepsilon^{(i-2)}, X_5 = \varepsilon^{(i-3)}, X_6 = P_{in}^{(i)}, X_7 = T_c^{(i)}$  і  $Y_i = T_v^{(i)}$ , де  $i = \overline{1, N}$ . Таким чином, у відповідності з формулою (10), кількість членів початкової моделі (9) буде складати ( $m = 4, n = 7$ ) –  $M = 330$ . Оскільки у процесі розв'язання задачі з синтезу оптимальної структури полінома за допомогою розробленого методу багаторазово доводиться розв'язувати нормальне рівняння Гауса, то для кращої обумовленості задачі вхідні та вихідні змінні рекомендується [3] приводити до безрозмірного виду. Був вибраний діапазон зміни безрозмірних як вхідних, так і вихідних величин від нуля до одиниці. Тоді

$$x_k^{(i)} = \frac{X_k^{(i)} - X_{k,max}}{X_{k,max} - X_{k,min}}, \quad k = \overline{1, n}; \quad (17)$$

$$y_i = \frac{Y_i - Y_{min}}{Y_{max} - Y_{min}}, \quad i = \overline{1, N}. \quad (18)$$

Перехід до розмірних величин здійснюється за допомогою співвідношення

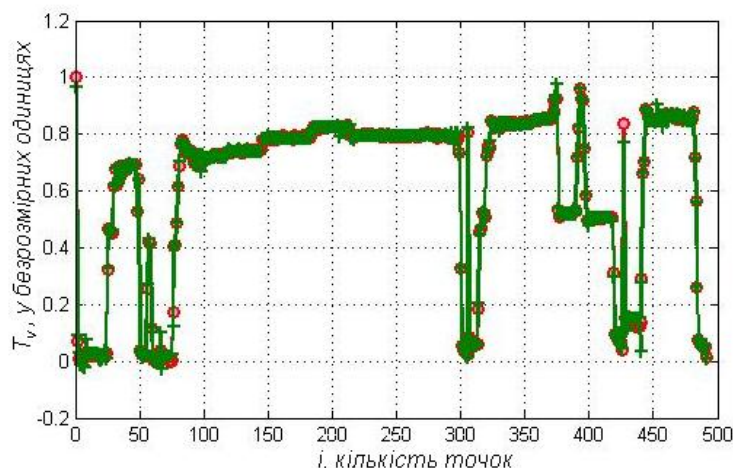


Рисунок 3 – Залежність температури на виході ТНТ від технологічних параметрів

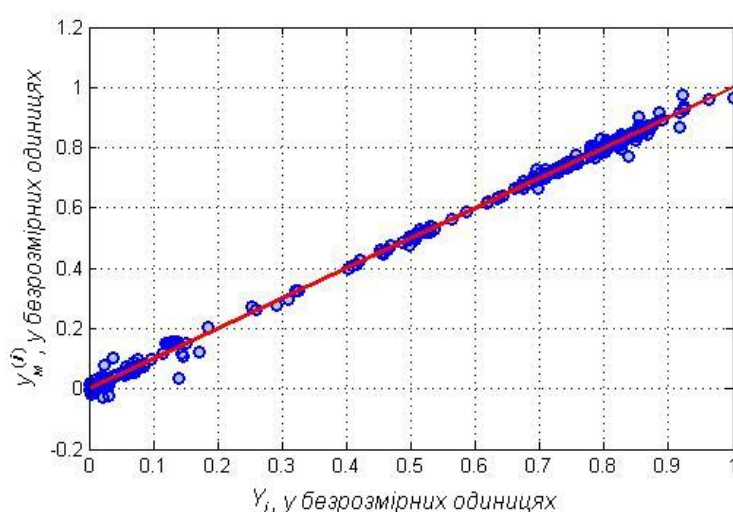


Рисунок 4 – Результати перевірки моделі на адекватність

$Y_i = y_i (Y_{\max} - Y_{\min}) + Y_{\min}$ ,  
яке легко отримати із формули (18).

З використанням розробленої програми синтезована модель, яка вміщує 154 ненульових і  $330 - 154 = 176$  нульових параметрів  $a_i$ ,  $i = \overline{0, M - 1}$  полінома (9). Результати роботи програми відтворює рисунок 3, де через «○» позначені експериментальні дані, а через «+» – значення температури  $T_v$ , які обчислені як вихід синтезованої моделі.

Адекватність моделі перевірялась за допомогою коефіцієнта кореляції  $K_{y_y}$  між значеннями  $Y_i$  та її виходом  $y_m^{(i)}$

$$K_{y_y} = \frac{\sum_{i=1}^N (Y_i - \bar{Y})(y_m^{(i)} - \bar{y}_m)}{\sqrt{\sum_{i=1}^N (Y_i - \bar{Y})^2} \cdot \sqrt{\sum_{i=1}^N (y_m^{(i)} - \bar{y}_m)^2}},$$

де  $\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i$ ,  $\bar{y}_m = \frac{1}{N} \sum_{i=1}^N y_m^{(i)}$  – оцінки математичних сподівань для величин  $Y_i$  і  $y_m^{(i)}$ . Було отримано:  $K_{y_y} = 0,999$ , що свідчить про високу ступінь кореляції між величинами  $Y_i$  і  $y_m^{(i)}$ .

На рис. 4 зображено залежність між виходом моделі  $y_m^{(i)}$  і експериментальними значеннями  $Y_i$ . У разі виконання умови  $y_m^{(i)} = Y_i$  на площині  $Y_i 0 y_m^{(i)}$  матимемо пряму лінію, яка засвідчує про повний збіг експериментальних результатів і виходу емпіричної моделі, побудованої за такими результатами. Пряма лінія, яка зображена на рис. 4, побудована з використанням МНК-метода і вона вказує на досить мале відхилення експериментальних точок  $Y_i$  від розрахункових значень  $y_m^{(i)}$ , що свідчить про адекватність синтезованої емпіричної моделі на засадах генетичних алгоритмів.

Таким чином, розроблено метод синтезу емпіричних моделей на засадах генетичних алгоритмів, який, на відміну комбінаторного методу, значно скорочує кількість обчислювальних операцій, що відкриває можливості для побудови складних моделей технологічних процесів. Такі моделі можуть бути використані у задачах оптимізації та діагностування технічних систем. Ефективність розробленого методу перевірена на прикладі синтезу емпіричних моделей процесу компримування природного газу.

### *Література*

- 1 Вержбицкий В.М. Основы численных методов: учебник для вузов / В.М.Вержбицкий. – М.: Высшая школа, 2002. – 840 с.
- 2 Ермаков С.М. Математическая теория оптимального эксперимента: учебное пособие / С.М.Ермаков, А.А.Жиглявский. – М.: Наука, 1987. – 320 с.
- 3 Ивахненко А.Г. Индуктивный метод самоорганизации моделей сложных систем / А.Г.Ивахненко. – К.: Наукова думка, 1981. – 296 с.
- 4 Ивахненко А.Г. Справочник по типовым программам моделирования / А.Г.Ивахненко, Ю.В.Коппа, В.С.Степашко и др.; под ред. А.Г.Ивахненко. – К.: Техніка, 1980. – 180 с.

5 Горбійчук М.І. Індуктивний метод побудови математичних моделей газоперекачувальних агрегатів природного газу / М.І.Горбійчук, М.І.Когутяк, Я.І.Заячук // Нафтова і газова промисловість. – 2008. – № 5. – С. 32 – 35.

6 Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы / Д. Рутковская, М. Пилиньский, Л. Рутковский; пер. с польск. И. Д. Рудинского. – М.: Горячая линия-Телеком, 2004. – 452 с.

7 Горбійчук М.І. Математичне моделювання процесу компримування природного газу / Горбійчук М. І., Когутяк М. І., Ковалів Є. О. // Розвідка та розробка нафтових і газових родовищ. – 2003. – №3 (8). – С. 21 – 26.

8 Компресорні станції. Контроль тепло-технічних та екологічних характеристик газоперекачувальних агрегатів: СОУ 60.03-30019801-011:2004. – Офіц. вид. – К.: ДК «Укртрансгаз», 2004. – 117 с.

*Стаття поступила в редакційну колегію*  
29.10.09

*Рекомендована до друку професором*  
**Юрчишиним В.М.**