УДК 681.84.087.4

А. Ковальчук
Національний університет “Львівська політехніка”,
кафедра автоматизованих систем управління

ЗАСТОСУВАННЯ ГЕНЕТИЧНОГО АЛГОРИТМУ В НЕЧІТКИХ МОДЕЛЯХ ПРОГНОЗУВАННЯ ЗАБРУДЕННЯ ДОВКІЛЛЯ

© Ковальчук А. 2008

Показано можливість застосування генетичного алгоритму в нечітких моделях прогнозу забруднення довкілля, використовуючи менший об’єм необхідної інформації, незначно втрачаючи в якості і точності прогнозу. Запропонований підхід може бути застосовний в різних сферах соціального і природничого передбачення.

The feasibility of genetical algorithm in indistinct fuzzy models of contamination of medium is rotined,using a smaller volume of the indispensable information,a little bit losing in quality and accuracy of the forecast. The offered approach can is applied in miscellaneous orbs of a social naturalists' prediction.

Вступ

Метою роботи є опис методу прогнозу забруднення довкілля в нечітких умовах, коли інформація неповна. Основу методу становить використання методів теорії нечітких множин, зокрема нечітких моделей систем з логічною структурою [1–8]. Використовуючи ідеї генетичного алгоритму [9] в таких методах, можна будувати передбачення, використовуючи мінімальну вхідну множину емпіричних даних.

1. Моделі систем з логічною структурою

У дослідженнях природничих і соціальних систем пари “вхід – вихід” задаються нечіткими правилами висловлювань типу “якщо $\boldsymbol{A}$, то $\boldsymbol{B}$”, де $\boldsymbol{A}$ і $\boldsymbol{B}$ – нечіткі підмножини вхідного універсума $\boldsymbol{U}$ і вихідного універсума $\boldsymbol{V}$ відповідно. Сукупність таких висловлювань можна розглядати як вербальне задання нечіткої системи. У висловленні “якщо $\boldsymbol{A}$, то $\boldsymbol{B}$”, які в нечіткій логіці записуються у вигляді $\boldsymbol{A}\to\boldsymbol{B}$, вважатимемо множину $\boldsymbol{A}$ нечітким вхідним, множину $\boldsymbol{B}$ – нечітким вихідним. Тому запис $\boldsymbol{A}\to\boldsymbol{B}$ інтерпретуємо як пара “вхід – вихід” ($\boldsymbol{A}$, $\boldsymbol{B}$).

За цим підходом наголос робиться на отримання нечіткої системи, закріпленної на заданій логічній структурі. Розглядаються два різні типи системних моделей, які відповідають двом різних типам логічних структур. Моделі систем названі модель I і модель II. В логічній структурі моделі I приймається , що більшій вхідній множині відповідає більша вихідна множина. Для логічної структури моделі II, навпаки, вважають, що більшій вхідній множині відповідіє невелика вихідна множина.

Розглянута проблема полягає в тому, щоб отримати системне представлення $\boldsymbol{R}$ з якимось оператором *, таким, щоб:

1) $\boldsymbol{B}_i = \boldsymbol{A}_i * \boldsymbol{R}$ для цієї пари вхід – вихід $\boldsymbol{A}_i \to \boldsymbol{B}_i$;

2) система $\boldsymbol{R}$ була наділена логічною структурою.

2. Нечіткі правила виведення

Модель виведення основується на нечітких висловлюваннях. Прикладом моделі виведення може бути правило:

нечітка вхідна множина $\boldsymbol{A}_1$: $\boldsymbol{x}$ мале, нечітке відношення $\boldsymbol{R}$: x і y приблизно однакові (нечітка система).

нечіткі вихідні множини $\boldsymbol{B}_1 = \boldsymbol{A}_1 * \boldsymbol{R}$: y більш-менш малі.Тут оперція * – мінімаксний добуток.

Оскільки таке нечітке висловлювання, як $\boldsymbol{A} \rightarrow \boldsymbol{B}$, задає нечіткі відношення між вхідними і вихідними множинами, то нечітку систему $\boldsymbol{R}$ можна визначити правилом:

$$\boldsymbol{R} = (\boldsymbol{A} \times \boldsymbol{B}) \cup (\bar{\boldsymbol{A}} \times \boldsymbol{B}), \tag{1}$$

де $\boldsymbol{A} \times \boldsymbol{B}$ – декартів добуток множин $\boldsymbol{A}$ і $\boldsymbol{B}$, а $\bar{\boldsymbol{A}}$ – доповнення $\boldsymbol{A}$. В еквівалентному формулюванні маємо

$$\boldsymbol{R}(u, \boldsymbol{v}) = (\mu_A(\boldsymbol{u}) \max \mu_B(\boldsymbol{v})) \min (\mu_{\bar{A}}(\boldsymbol{u}) \max \mu_V(\boldsymbol{v})), \tag{2}$$

де μ_A, μ_B, і μ_V– функція належності множин $\boldsymbol{A}$, $\boldsymbol{B}$ і $\boldsymbol{V}$, причому $\mu_V(\boldsymbol{v})=1$ для всіх $\boldsymbol{v}$, і $\mu_{\bar{A}} = 1 - \mu_A$.

Більш загальне висловлювання "якщо A, то $\boldsymbol{B}$, інакше $\boldsymbol{C}$" можна записати як

$$\boldsymbol{R} = \boldsymbol{A} \times \boldsymbol{B} \cup \bar{\boldsymbol{A}} \times \boldsymbol{C}. \tag{3}$$

Якщо задана така нечітка система $\boldsymbol{R}$, для вхідної $\boldsymbol{A}$ і вихідної $\boldsymbol{B}$ множин виконується

$$\boldsymbol{B} = \boldsymbol{A} \bullet \boldsymbol{R}, \tag{4}$$

звідки випливає, що система $\boldsymbol{R}$ задає неточне представлення висловлювання $\boldsymbol{A} \rightarrow \boldsymbol{B}$. Однак якщо використовувати наступне визначення R,

$$\mu_R(\boldsymbol{u}, \boldsymbol{v}) = \begin{cases} 1, \text{ якщо } \mu_A(\boldsymbol{u}) \le \mu_B(\boldsymbol{v}) \\ \mu_B(\boldsymbol{v}), \text{ якщо } \mu_A(\boldsymbol{u}) > \mu_B(\boldsymbol{v}), \end{cases} \tag{5}$$

то отримаємо рівність

$$\boldsymbol{A} \bullet \boldsymbol{R} = \boldsymbol{B}, \tag{6}$$

точне представлення умови висловлювання "якщо $\boldsymbol{A}$, то $\boldsymbol{B}$, інакше $\boldsymbol{V}$".

Треба зазначити, що "інакше $\boldsymbol{C}$" заміняється на "інакше $\boldsymbol{V}$".

3. Загальна модель нечіткої системи

Досі розглядалися системи $\boldsymbol{R}$ з одним висловлюванням $\boldsymbol{A} \rightarrow \boldsymbol{B}(\boldsymbol{C})$. Тепер припустимо, що задано висловлювання типу:

$$\left.\begin{array}{l} \boldsymbol{A}_1 \rightarrow \boldsymbol{B}_1(\boldsymbol{C}_1), \\ \boldsymbol{A}_2 \rightarrow \boldsymbol{B}_2(\boldsymbol{C}_2), \\ \quad \ldots\ldots\ldots\ldots\ldots \\ \boldsymbol{A}_n \rightarrow \boldsymbol{B}_n(\boldsymbol{C}_n). \end{array}\right\} \ldots\ldots \tag{7}$$

Побудуємо представлення системи $\boldsymbol{R}$, яке відповідає системі висловлювань. Тут виникає проблема, як встановити зв'язки між нечіткими висловлюваннями. Розглянемо як висловлювання: $\boldsymbol{A}_1 \rightarrow \boldsymbol{B}_1(\boldsymbol{C}_1)$ і $\boldsymbol{A}_2 \rightarrow \boldsymbol{B}_2(\boldsymbol{C}_2)$ і ... і $\boldsymbol{A}_n \rightarrow \boldsymbol{B}_n(\boldsymbol{C}_n)$. Наступне відношення:

$$\boldsymbol{C}_i \supset \bigcup_j \boldsymbol{B}_j \tag{8}$$

допустиме, оскільки $\bar{\boldsymbol{A}}_i$ відповідає умові

$$\boldsymbol{B} \supset \boldsymbol{B}_1 \cup \ldots \cup \boldsymbol{B}_{i-1} \cup \boldsymbol{B}_{i-1} \cup \boldsymbol{B}_{i+1} \cup \ldots \cup \boldsymbol{B}_n$$

Через це в загальному випадку $\boldsymbol{C}_i$ розглядається як $\boldsymbol{V}$, $\boldsymbol{C}_i = \boldsymbol{V}$. Якщо ж (7) представити як висловлювання $\boldsymbol{A}_1 \rightarrow \boldsymbol{B}_1(\boldsymbol{C}_1)$ або $\boldsymbol{A}_2 \rightarrow \boldsymbol{B}_2(\boldsymbol{C}_2)$ або, ... або $\boldsymbol{A}_n \rightarrow \boldsymbol{B}_n(\boldsymbol{C}_n)$, то при цьому стає справедливим відношення

$$\boldsymbol{C}_i \supset \bigcap_j \boldsymbol{B}_j \tag{9}$$

оскільки $\bar{\boldsymbol{A}}_i$ має властивість $\boldsymbol{B} \supset \boldsymbol{B}_1 \text{U} \ldots \text{U}\, \boldsymbol{B}_{i-1} \text{U} \boldsymbol{B}_{i-1} \text{U} \boldsymbol{B}_{i+1} \text{U} \ldots \text{U} \boldsymbol{B}_n$. Тому в загальному випадку пропонується, що $\boldsymbol{C}_i = \boldsymbol{\varnothing}$. Тоді для спрощення вважаємо, що $\boldsymbol{C} = \boldsymbol{V}$ або $\boldsymbol{C} = \boldsymbol{\varnothing}$. Виходячи із сказаного, будемо користуватися наступною інтерпретацією: запис $\boldsymbol{C} = \boldsymbol{V}$ означає „довільний елемент в $\boldsymbol{V}$", а запис $\boldsymbol{C} = \boldsymbol{\varnothing}$ – „ невідомий як елемент $\boldsymbol{V}$".

Розглянута задача полягає в отриманні системи

$$\boldsymbol{Y} = \boldsymbol{X} * \boldsymbol{R} \tag{10}$$

веденням наступних двох логічних структур.

4. Модель – І нечіткої системи

Модель – І. Моделлю-І називається система $\boldsymbol{R}$, що задається такими умовами:

1) для всіх заданих пар "вхід – вихід" ($\boldsymbol{A}_i$, $\boldsymbol{B}_i$)
2) $\boldsymbol{B}_i = \boldsymbol{A}_i * \boldsymbol{R}$;
3) для вхідних множин $\boldsymbol{A}'_i$ і $\boldsymbol{A}''_i$ (за винятком цієї вхідної множини $\boldsymbol{A}_i$)

$$\boldsymbol{A}'_i \subset \boldsymbol{A}_i \subset \boldsymbol{A}''_i \Rightarrow \boldsymbol{B}'_i = \boldsymbol{A}'_i * \boldsymbol{R} \subset \boldsymbol{B}'_i \subset \boldsymbol{B}''_i = \boldsymbol{A}''_i * \boldsymbol{R}.$$

Модель-І являє таке відношення між вхідною і вихідною множинами, при якому чим менша нечітка вхідна множина, тим менша нечітка вихідна.

Розглянемо спочатку нечіткий випадок, для якого відношення вхід – вихід можна подати у вигляді

$$\boldsymbol{X}_{\mathbf{Y}} = \sup_{u \in U} \{ 1\text{-}\boldsymbol{X}_X(\boldsymbol{u}, \boldsymbol{\upsilon}), \min \boldsymbol{X}_R(\boldsymbol{u}, \boldsymbol{\upsilon})\} = \sup_{u \in U} \{ \boldsymbol{X}\, \bar{X}\ (\text{u}), \min \boldsymbol{X}_R(\boldsymbol{u}, \upsilon)\}.$$

Використовуючи процес розмивання, отримаємо таке подання нечіткої системи:

$$\begin{aligned}
\boldsymbol{\mu}_Y &= \sup_{\alpha \in [0,1]} [\, \boldsymbol{\alpha} \max \sup_{u \in U} \{ \text{X}\, \bar{X}^{\alpha}(\text{u}) \max \text{X}_R^{\alpha}(\text{u}, \upsilon) \}] = \\
&= \sup_{\alpha \in [0,1]} [\alpha \min \sup_{u' \in \{u | X_X \alpha(u) = 1\}} \{ \text{X}_R(\text{u}', \upsilon)\}] = \\
&= \sup_{\alpha \in [0,1]} [\alpha \min \sup_{u' \in \{u | X_X \alpha(u) > X_R \alpha(u, \upsilon)\}} \{ \boldsymbol{X}_R^{\alpha}(\text{u}', \upsilon)\}] = \\
&= \sup_{u' \in \{u | X_X \alpha(u) > X_R(u, \upsilon)\}} \{ \mu_R(\text{u}', \upsilon)\},
\end{aligned} \tag{11}$$

де sup $\{(\mu_R(\boldsymbol{u}', \boldsymbol{\upsilon})\} = 0$, якщо множина $\{\boldsymbol{u} | \mu_X(\boldsymbol{u}) > \mu_R(\boldsymbol{u}, \boldsymbol{\upsilon})\}$ пуста.

Нехай на основі пар вхід – вихід вже складено відношення R^*_0 і тепер отримуємо додаткову інформацію $\boldsymbol{A}_1 \rightarrow \boldsymbol{B}_1$. Оскільки у моделі-ІІ відношення "вхід – вихід" зв'язані відношенням $\boldsymbol{A}_1 \rightarrow \boldsymbol{B}_1$ або $\boldsymbol{A}_2 \rightarrow \boldsymbol{B}_2$ або...або $\boldsymbol{A}_n \rightarrow \boldsymbol{B}_n$, то представлення системи $\boldsymbol{R}^*$ визначимо як

$$\boldsymbol{R}^* = (\boldsymbol{A}_1 \otimes \boldsymbol{B}_1)\, \text{U} \boldsymbol{R}^*_0. \tag{12}$$

вважаючи, що $\boldsymbol{R}^*_0$ не несе інформації відносно $\boldsymbol{A}_1$, а систему $\boldsymbol{R}^*_0$ на виході, відповідну $\boldsymbol{A}_1$, можна розглянути як „елемент, не відомий в $\boldsymbol{V}$", причому для двох нечітких множин $\boldsymbol{A} \rightarrow \boldsymbol{B}$ операція $\boldsymbol{A} \otimes \boldsymbol{B}$ визначається так:

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{cases} \mu_B, \mu_A > \mu_B \\ 1, \mu_A \le \mu_B \end{cases}$$

Тепер розглянемо дві пари вхід – вихід $A_1 \to B_1$ і $A_2 \to B_2$. Запишемо умови несуперечливості

$$A_1 R (A_2 \otimes B_2) = \varnothing,$$
$$A_2 R (A_1 \otimes B_1) = \varnothing. \quad (13)$$

Очевидно, що

$$B = (A_1 \cap A_2) R \supseteq B_1 \cap B_2 .$$

Отже, маємо

$$\left.\begin{array}{l} A_1 \to B_1, \\ A_2 \to B_2, \\ A_1 \cap A_2 \to B \subset B_1 \cap B_2 . \end{array}\right\} \quad (14)$$

У загальному випадку, якщо є умови несуперечливості на n парах $(A_i \to B_i)$, аналогічні (13), то маємо

$$\left.\begin{array}{l} A_i R = B_i, \\ (A_i \cap A_j) R \subset B_i \cap B_j, \end{array}\right\}$$

5. Використання генетичного алгоритму

Нехай для визначення екологічних факторів діяльності людини обрані n впливів людини на середовище й набір з m індикаторів стану, найважливіших, на думку ряду експертів. Як тестовий приклад використовуватимемо, відповідно до вищенаведеного принципу 3, дані роботи [6]. Вплив, що відповідає кожній дії й кожному факторові, описується через амплітуду й важливість. Амплітуда – це міра загального рівня, масштабу впливу.

Вхідні дані:

1. Контроль над ерозією; 2. Споруди для відпочинку; 3. Іригація; 4. Спалювання відходів;
5. Будівництво мостів і доріг; 6. Штучні канали; 7. Греблі; 8. Тунелі й підземні споруди;
9. Підривні й бурові роботи; 10. Відкриті розробки; 11. Вирубка лісів;
12. Комерційне полювання й рибальство; 13. Рослинництво; 14. Розведення й випас худоби;
15. Хімічна промисловість; 16. Лісопосадки; 17. Добрива;
18. Розведення й регулювання популяції диких тварин; 19. Автомобільний рух;
20. Трубопроводи; 21. Сховища відходів; 22. Використання отрутохімікатів;
23. Течії і розливи.

Вихідні дані (ухвалення рішення):

1. Стан ґрунту; 2. Стан поверхневих вод; 3. Якісний склад вод; 4. Якісний склад повітря;
5. Температура повітря; 6. Ерозія; 7. Дерева й чагарники; 8. Трави; 9. Сільгоспкультури;
10. Мікрофлора; 11. Тварини суші; 12. Риби й молюски; 13. Комахи;
14. Заболочування території; 15. Курорти на суші; 16. Парки й заповідники;
17. Здоров'я й безпека; 18. Зайнятість людей; 19. Щільність населення; 20. Солоність води;
21. Солончаки; 22. Хащі; 23. Зсуви грунту.

Нехай задано нечіткі множини $S_i \subset U$, $i = 1, 2, 3$, який описують деякий реальний процес. Якщо прийняти, що $A_1 = S_1$, $B_1 = S_2$, $A_2 = S_2$, $B_2 = S_3$, $A^* = S_3$, то використовуючи співвідношення (11) і (12), можемо отримати нечітку множину $S_4 = R^*(S_3)$, яку можна інтерпретувати як результат існування нечітких мнжин $S_i \subset U$, $i = 1, 2, 3$.

Використовуючи генетичний алгоритм [17], можна обійтися тільки двома множинами $A_1 = S_1$, $B_1 = S_2$. Дійсно, нехай $m \geq 1$ – номер ітерації. Розглянемо вектор $\vec{a} = \left\{a_1^{(m)}, \ldots, a_n^{(m)}\right\}$ і числа $s^{(m)} = \max\ a_i^{(m)}$, $t^{(m)} = \sum_{i=1}^{n} a_i^{(m)}$, $c_i^{(m)} = a_i^{(m)} / t^{(m)}$, $r_i^{(m)} = a_i^{(m)} / s^{(m)}$, $x_i^{(m)} = c_i^{(m)} \cdot r_i^{(m)} + (1 - c_i^{(m)}) \cdot a_i^{(m)}$, $y_i^{(m)} = (1 - c_i^{(m)}) \cdot r_i^{(m)} + c_i^{(m)} \cdot a_i^{(m)}$, і виберемо

$a_i^{(m+1)} = \mathbf{max}\ (x_i^{(m)}, y_i^{(m)})$, $1 \le i \le n$. Застосувавши до $\boldsymbol{A_1 = S_1, B_1 = S_2}$ вказаний алгоритм, знайдемо $\boldsymbol{A_2 = S_1^{(1)}, A^* = S_1^{(2)}, B_2 = S_2^{(1)}}$, і отримаємо $\boldsymbol{S_4 = R}$ *($\boldsymbol{S_1^{(2)}}$).

6. Аналіз результатів

Числові експерименти проводилися для вказаних вхідних і вихідних даних, значення яких знаходилися в межах між 0 і 1. Для побудови прогнозу бралися три множини $\boldsymbol{S_i} \subset \boldsymbol{U}$, $\boldsymbol{i = 1, 2, 3}$ в припущенні, що $\boldsymbol{A_1 = S_1, B_1 = S_2, A_2 = S_2}$, $\boldsymbol{B_2 = S_3, A^* = S_3}$. З множин $\boldsymbol{A_1} \rightarrow \boldsymbol{B_1}$, $\boldsymbol{A_2} \rightarrow \boldsymbol{B_2}$ будується відношення $\boldsymbol{R}$, таке, що $\boldsymbol{B^* = R(A^*)}$. Результати наведені на рис. 1.

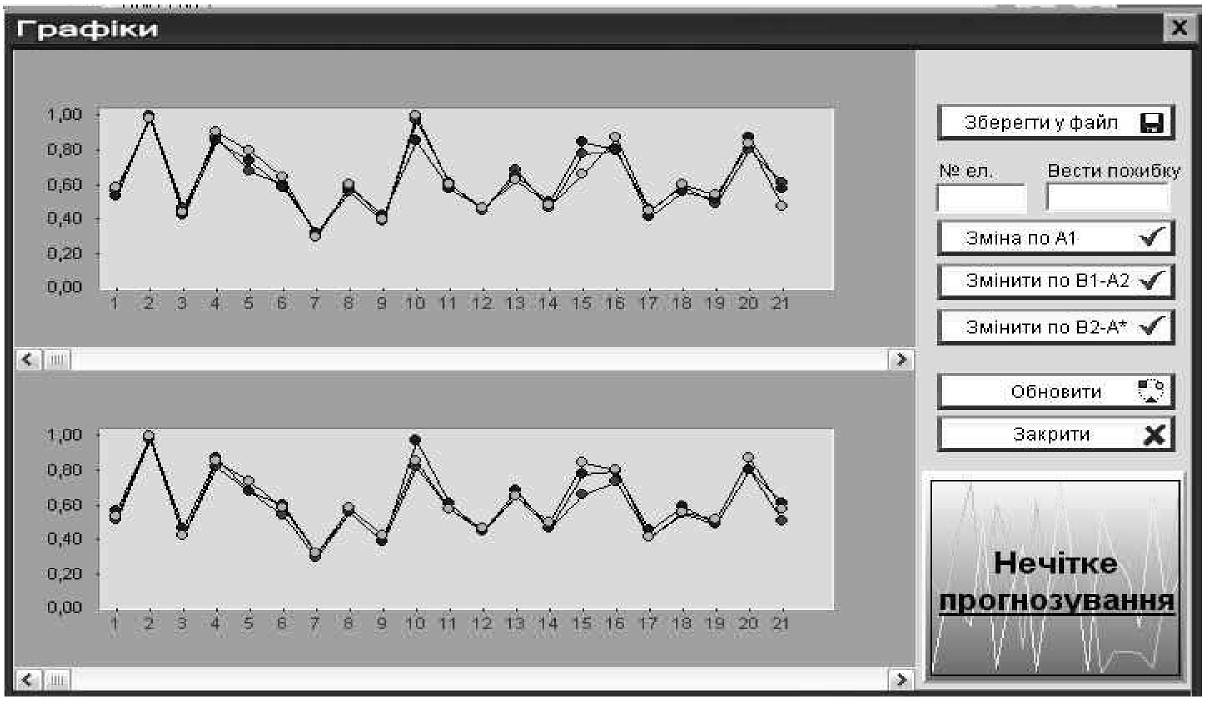

Рис. 1. Нечіткий прогноз забруднення

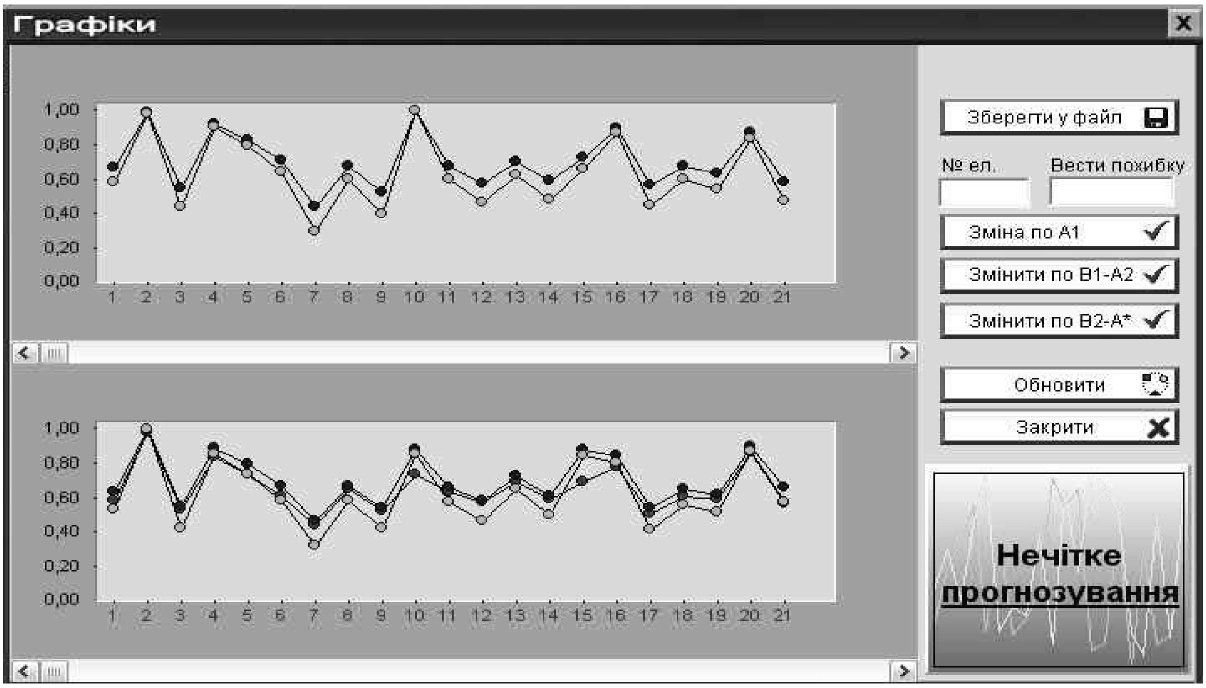

Рис. 2. Нечіткий прогноз забруднення з використанням генетичного алгоритму

На рис. 2 показано результати прогнозу забруднення з використанням генетичного алгоритму за наявності тільки двох множин $\boldsymbol{A_1} = \boldsymbol{S_1}$, $\boldsymbol{B_1} = \boldsymbol{S_2}$.

Висновок

На розглянутих прикладах показано можливість застосування генетичного алгоритму в нечітких моделях прогнозу забруднення середовища з використанням меншого об'єму необхідної інформації, незначно втрачаючи в якості і точності. Запропонований підхід може бути застосовний в різних сферах соціального і природничого передбачення.

1. *Казиев В.М. Математические и компьютерные модели экологических систем / Тезисы докладов региональной научной конференции "Современные проблемы экологии". Ч. 2. – Краснодар–Анапа, 1996. – С. 87. 2. Большаков В.Н., Криницин С.В., Кряжимский Ф.М., Мартинес Рика Х.П. Проблемы восприятия современным обществом основных понятий экологической науки // Экология. – 1996. – № 3. – С. 165–170. 3. Пых Ю.А., Малкина-Пых И.Г. Об оценке состояния окружающей среды. Подходы к проблеме // Экология. – 1996. – № 5. – С. 323–329. 4. Казиев В.М. Некоторые оптимизационные задачи управления экосистемами // Доклады А(Ч)М АН. – 1994. – № 1. 5. Прикладные нечеткие системы / Под ред. Т. Тэрано, К. Асаи, М. Сугэно. – М.: Мир, 1993. 6. Экологические системы. Адаптивная оценка и управление / Под ред. К.С. Холинга. – Мир, 1981. 7. Обработка нечеткой информации в системах принятия решений. – М.: Радио и связь, 1989. 8. Прикладные нечеткие системы / Пер. с япон. К. Асан, Д. Ватада, С. Иваи и др. – М.: Мир, 1993. 9. Баодінг Лю. Теория и практика неопределенного программирования. – М.: БИНОМ. Лаборатория знаний, 2005. – 416 с.*