УДК69:002;69.059

Григоровський П.Є., к.т.н. ДП НДІБВ, м. Київ
Терентьєв О.О., проф., д.т.н., КНУБА, м. Київ

ІНТЕГРОВАНІ МОДЕЛІ, ЯКІ ЗАБЕЗПЕЧУЮТЬ ПРОГНОЗУВАННЯ НАДІЙНОСТІ ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ ЗАДАЧІ СИСТЕМИ ДІАГНОСТИКИ ТЕХНІЧНОГО СТАНУ БУДІВЕЛЬ

АНОТАЦІЯ

Проведені дослідження дозволяють запропонувати підхід до вирішення завдання вибору тієї чи іншої моделі, що описує зміну динаміки виміряних даних внаслідок старіння та зносу конструкцій, дозволяють обґрунтовано вибирати ступінь складності моделі, що забезпечує найвищу точність прогнозу з моменту настання пошкодженого стану.

Ключові слова: інтегровані моделі, діагностика технічного стану, інформаційна технологія, прогнозування надійності прийняття рішень.

Актуальність та аналіз проблеми

В даний час найбільш активно використовуються алгоритми прогнозування термінів настання експлуатаційної придатності будівель, що засновані на застосуванні методів математичної статистики, теорії розпізнавання образів і синергетики. Відмінною особливістю цих алгоритмів є виявлення часових характеристик показників надійності розрахункових параметрів.

Найбільш інформативним параметром, що характеризує рівень технічного стану будівель є напружений стан елементів будівель. У загальному випадку, рівень складності апроксимуючої функції залежить не тільки від самого змінюваного параметра, але і від рівня шумової складової вимірювань і обсягу вибірки.

Вибір тієї чи іншої моделі, що описує зміни показників надійності функціонування технічного стану, є найбільш відповідальним та складним етапом прогностичної процедури.

Спрощення моделі призводить до зменшення точності прогнозу часу настання експлуатаційної придатності будівель. Зайве ускладнення моделі може призвести до нестійкості алгоритму ідентифікації і позбавити ідентифікаційні моделі пророчої сили. Крім того, необхідно враховувати, що ступінь складності моделі залежить не тільки від ідентифікованого параметра, але і від рівня похибки первинних вимірювань.

Мета дослідження

Представляється актуальна багатокритеріальна задача вибору оптимального ступеня складності моделей, що описують зміну показників надійності будівель.

При виборі методу вирішення поставленого завдання постає дві додаткові умови.

Перша модель повинна відповідати прогнозуючим властивостям, тобто при екстраполяції на деякий проміжок часу її значення не повинно "розбовтуватися". Ця умова накладає обмеження на ступінь складності функцій — для надто складної моделі малі помилки вимірювань, не помітні на інтервалі інтерполяції, на етапі прогнозу можуть радикально змінювати поведінку модельної функції.

По-друге, припускаємо, що обсяг вибірки даних, за якою будується модель, невеликий. Це пов'язано з тим, що найбільш достовірна інформація зберігається в базах даних сучасних інформаційних систем, що охоплює часовий інтервал у 5-6 років.

Складність задачі оптимального вибору апроксимуючої функції, що описує ту чи іншу зміну показників експлуатації будівель, посилюється помилками вимірювань, які проявляються у вигляді накладення шуму на координати експериментальних точок.

Одним з найбільш точних та інформативних показників технічного стану, що контролює інформаційні системи, є зміна у часі напруженого стану конструкцій.

Аналіз представлених емпіричних даних показує, що динаміка зміни даних перед різними типами пошкоджень відрізняється кардинальним чином.

Дійсно, використовуючи один або кілька класичних критеріїв (мінімум величини дисперсії адекватності, критерій Тейла) і схему стандартного методу найменших квадратів (МНК), можна побудувати модель з бажаним ступенем точності, не порушуючи при цьому принципу Пуанкаре (точність моделі не може перевершувати точності первинної інформації). Однак це не дає вирішення прогностичності завдання — визначення моменту пошкоджень, так як найкраща на етапі навчання модель не завжди є і більш точною екстраполяцією майбутнього сценарію розвитку.

Приведемо це на прикладі прогнозу моменту настання пошкоджень. Попередня селекція елементарних функцій, що описують таку поведінку експериментальних кривих, показала, що найбільш точні (в сенсі дисперсії адекватності) поліноміальні залежності.

Аналіз отриманих результатів показує, що помилка прогнозу моменту настання пошкоджень лінійною моделлю становить 56%, поліномом 3-го ступеня 14%, поліномом 2-го ступеня 2%. У той же час величина дисперсії адекватності цих моделей на етапі навчання практично однакова. Таким чином, стає очевидною необхідність використання додаткових методів обробки даних, що в повній мірі реалізують інформаційні можливості систем.

Виклад основного матеріалу

Сформулюємо нашу задачу з урахуванням усіх зауважень.

За динамікою даних напруги несучих конструкцій будівель за деякий період часу необхідно побудувати найкращу модель розвитку дефекту по двом критеріям-точність апроксимації плюс точність прогнозу.

Найбільш ефективним інструментом вирішення подібного роду завдань є метод структурної мінімізації середнього ризику. Адаптуємо цей метод до умов нашої задачі.

В інформаційній базі системи зберігається безліч локальних баз даних, $\{x_i\}$, кожна з яких представляє собою ретроспективний часовий ряд зміни показника експлуатації в часі $i=1,2,..., L$, де L визначається частотою опитування первинних датчиків.

Припустимо, що на підставі аналізу цих масивів даних будуються моделі виду $y=y(x)$ (у розглянутому випадку $Q = Q(t)$, де $Q(t)$ — зміна напруженого стану внаслідок зносу конструкцій, t — час). В такому разі, у розпорядженні є вибірка $\{xi\ yj\}$, де y_j y, — модельне значення функції, що відповідає експериментально виміряному значенню параметра x_i.

Враховуючи, що експериментальні дані завжди вимірюються з деякою похибкою, введемо в розгляд перешкоду вимірювання ε_i. Тоді шукана модель прийме остаточний вигляд:

$$y=F(t_i)+ \varepsilon_i . \qquad (1)$$

Припускаючи, що клас функцій, у якому шукається регресія $y(x)$, є параметричним з параметрами a, задачу можна звести до мінімізації функціоналу емпіричного ризику:

$$I_o(a) = \frac{1}{L}\sum_{i=1}^{L} (y_i - F(x_i, a))^2 , \qquad (2)$$

де y_i — модельне значення параметра з урахуванням перешкоди виміру;

$F(x_i, a)$ — моделююча функція;

L — обсяг вибірки вимірювань, визначається частотою опитування первинних датчиків.

В роботі показано, що для критерію (3.36) можуть бути отримані верхні оцінки виду:

$$I(a) \le I_m(a) = I_o(a)\Omega(\frac{1}{h}, \frac{l_{n\eta}}{L}) , \qquad (3)$$

справедливі з імовірністю $1-\eta$. Величина h являє собою ємність класу функцій $F(x,a)$ і визначає складність ідентифікованої моделі. Зокрема, якщо розглядається клас лінійних за параметрами функцій:

$$F(x,a) = \sum_{i=1}^{n} a_i\psi(x) , \qquad (4)$$

де $h=n$, тобто ємність класу функцій (складність моделі) дорівнює числу шуканих параметрів n.

Величина $1/\eta$ визначає відносний обсяг вибірки. Структура другого множника (3) така, що з ростом $1/h$ величина зменшується, прагнучи до одиниці.

Функціонал (2) зі збільшенням $1/h$, як правило, збільшується. Тому існує деяке оптимальне значення $1/h$, при якому верхня оцінка середнього ризику (його гарантоване значення) досягає мінімуму, це значення $1/\Omega$ і визначає оптимальну складність шуканої функції.

У відповідності з рекомендаціями при відновленні регресії в класі функцій (4) в якості критерію Ω використовуємо величину:

$$\Omega = \left[\frac{1}{1-\sqrt{\dfrac{n(l_n \frac{1}{n} - l_{n\eta})}{L}}} \right. , \qquad (5)$$

$$[Z]_\infty = \begin{cases} z, z \ge 0 \\ \infty, z \le 0 \end{cases}$$

В роботі відмічено, що вирішення поставленої двоїстої задачі вдається отримати у разі використання досить великих вибірок експериментальних даних (обсяг $L>20$ вимірювань). У разі подальшого розвитку дефектів і при побудові відповідних моделей ця вимога не виконується, і метод СМСР стає надто грубим, що свідомо надає перевагу більш простим моделям. Найбільш ефективні результати для подолання подібного роду труднощів у ряді випадків можуть бути досягнуті шляхом залучення методів теорії нечітких множин.

Стосовно задачі під поняттям приналежності до того або іншого об'єкта будемо розуміти значення *{yi}*, обчислені за допомогою різних моделей (*i* — кількість розглянутих моделей).

Нечіткими множинами *A* в *U* називається сукупність пар виду $(u, \mu_A, (u))$ де — $u \in U$, $\mu_A, (u)$ функція належності нечіткої множини *A*. Близькість функції $\mu_A, (u)$ до 1 є кількісною мірою впевненості в тому, що елемент належить множині А.

Використання понять теорії нечітких множин дозволяє звести пошук сталого вирішення багатокритеріальної задачі до задачі пошуку екстремуму функції належності, яка визначається як:

$$\mu(a,n) = (\mu_o(I_o(a,n))\mu_o(n))^{0,5}, \quad (6)$$

де $\mu_0\,(I_0)$ і $\mu_0(n)$ — функції належності нечітких множин "малі значення емпіричного ризику" та "-мала складність моделі". Ці функції можуть бути визначені наступним чином:

$$\mu_o(I_o) = \psi(\frac{I_o}{I_i}, m_1),$$

$$\mu_o(n) = \psi(\frac{n}{0{,}5L}, m_2),$$

$$\psi(t,m) = \begin{cases} 1 - t^m, 0 \le t \le 0 \\ 0, t \succ 0 \end{cases}, \quad (7)$$

де L_i — значення функціоналу емпіричного ризику, що відповідає числу параметрів *m1* і *m2* — показники ступеня, що визначають ставлення алгоритму до зменшення емпіричного ризику і збільшення складності моделі.

В якості інформаційного масиву для побудови найкращої моделі прогнозу моменту настання пошкодження будемо використовувати 30-ти добове вимірювання даних, а в якості конкуруючих гіпотез розглянемо поліноми. Результати проведених розрахунків представлені в табл. 1.

МНК рекомендує максимальний ступінь складності апроксимуючої функції. Це цілком зрозуміло, оскільки МНК прагне мінімізувати відхилення експериментальних точок від апроксимуючої залежності, а це реалізується тільки при максимальній складності полінома. Метод СМСР допускає застосування інтерполюючого полінома зі ступенями *n=1* та *n=2*, тоді як методи теорії нечітких множин однозначно вказують, що оптимальним є ступінь полінома *n=2*, що повністю підтверджує достовірність результатів.

Чисельна оцінка "прогнозуючої здатності" розглянутих моделей проводилася на підставі визначення величин середньоквадратичних відхилень (СКВ) експериментальних точок від відповідних модельних функцій, що і визначає точність прогнозу. У нашому випадку величини СКВ рівні 1.24, 0,26 та 2.31 для поліномів 1-ої, 2-ої і 3-го ступеня відповідно. Тому очевидно, що здатність до прогнозу найбільш висока у полінома другого ступеня, що збігається з висновком, отриманим на основі теорії нечітких множин 100.

Аналіз отриманих результатів свідчить про те, що пропонований метод визначення оптимальної складності моделі дозволяє отримувати найбільш високу точність прогнозу моменту настання пошкоджень. У всіх розглянутих випадках збільшення точності прогнозу становить 20 — 30%.

При цьому слід зауважити, що різниця у виборі моделі, рекомендованої методом СМСР і мето-

Таблиця 1 **Обґрунтування вибору найбільш прийнятної прогностичної моделі визначення моменту настання пошкодженого стану**

Складність моделі	Критерій вибору моделі оптимальної складності		
	I0	*1m*	$\mu 10^4$
n=1	*0,024*	*0,0124*	*3,44*
n=2	*0,020*	*0,0124*	*2,48*
n=3	*0,014*	*0,0126*	*2,63*

дами теорії нечітких множин, збільшується зі зменшенням обсягу вибірки вимірювань даних. При досить великих вибірках (як правило, L>20) результати розрахунків за обома методами практично збігаються.

Висновок. Проведені дослідження дозволяють зробити висновок про те, що пропонований підхід до вирішення завдання вибору тієї чи іншої моделі, що описує зміну динаміки виміряних даних внаслідок старіння та зносу конструкцій, дозволяє обґрунтовано вибирати ступінь складності моделі, що забезпечує найвищу точність прогнозу з моменту настання пошкодженого стану.

ЛІТЕРАТУРА

1. Терентьєв О.О. Моделі визначення фізичного зношення конструктивних елементів будівлі для задач діагностики технічного стану / Баліна О.І., Шабала Є.Є.// — К.: Управління розвитком складних систем, збірник наукових праць, випуск 26/2016, КНУБА, 2016. — С. 153-157.

2. Терентьєв О.О. Основи організації нечіткого виведення для задачі діагностики технічного стану будівель та споруд [Текст] //О.О. Терентьєв, Є.Є. Шабала, Б.С. Малина// — К.: Управління розвитком складних систем, збірник наукових праць. — КНУБА, 2015. — №22. — С. 138 — 143.

3. Терентьєв О.О. Інформаційна технологія системи діагностики технічного стану будівель на основі дослідження мікросейсмічних коливань /Шабала Є.Є, Малина Б.С.// — К.: Управління розвитком складних систем, збірник наукових праць, випуск 23/2015, КНУБА, 2015. — С.133 — 139.

4. Olexander Terentyev Development of models and methods for determining the physical deterioration of items for the task of diagnostics of technical condition of buildings and structures /Olexander Poltorak// — Scientific Journal "ScienceRise" №8/2(25), August 2016. — P. 14-19.

5. Olexander Terentyev The Method of Direct Grading and the Generalized Method of Assessment of Buildings Technical Condition [Text] // Olexander Terentyev, Mykola Tsiutsiura// — International Journal of Science and Research (IJSR), Volume 4 Issue 7, July 2015. — P. 827-829.

АННОТАЦИЯ

Проведенные исследования позволяют предложить подход к решению задачи выбора той или иной модели, описывающей изменение динамики измеренных данных вследствие старения и износа конструкций, позволяет обоснованно выбирать степень сложности модели, что обеспечивает высочайшую точность прогноза с момента наступления поврежденного состояния.

Ключевые слова: интегрированные модели, диагностика технического состояния, информационная технология, прогнозирования надежности принятия решений.

ANNOTATION

The conducted research allows to propose an approach to solving the problem of choosing a model describing the change in the dynamics of the measured data due to the ageing and deterioration of structures, allow to choose the degree of complexity of the model that provides the highest accuracy of the forecast since the onset of the corrupted state.

Key words: integrated model, diagnostics of the technical condition, information technology, forecasting the reliability of decision-making.