
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНІ ТЕХНОЛОГІЇ ТА МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

УДК [004.7-047.72]:656.2

В. М. ПАХОМОВА^{1*}

¹*Каф. «Електронні обчислювальні машини», Дніпропетровський національний університет залізничного транспорту імені академіка В. Лазаряна, вул. Лазаряна, 2, Дніпро, 49010, Україна, тел. +38 (056) 373 15 89, ел. пошта viknik.p1988@mail.ru, ORCID 0000-0001-8346-0405

ПРОГНОЗУВАННЯ ОБСЯГУ МЕРЕЖНОГО ТРАФІКА В ІНФОРМАЦІЙНО-ТЕЛЕКОМУНІКАЦІЙНІЙ СИСТЕМІ ПРИДНІПРОВСЬКОЇ ЗАЛІЗНИЦІ НА ОСНОВІ НЕЙРОНЕЧІТКОЇ МЕРЕЖІ

Мета. Постійне збільшення обсягу мережного трафіка в інформаційно-телекомунікаційній системі (ІТС) Придніпровської залізниці призводить до необхідності визначення в реальному часі перевантаження в мережі та здійснення контролю потоків даних. Одним із можливих рішень є метод прогнозування обсягу мережного трафіка (вхідного та вихідного) з використанням нейромережної технології, що дозволить уникнути перевантаження сервера та підвищити якість послуг. **Методика.** В роботі виконані аналіз існуючого мережного трафіка в ІТС Придніпровської залізниці та підготовка вибірок: навчальної, тестової, контрольної, а також створення в програмі Matlab нейронечіткої мережі (гібридної системи) та організація на відповідних вибірках таких етапів: навчання, тестування, аналіз адекватності прогнозу. **Результати.** Для фрагмента (Дніпропетровськ – Київ) в ІТС Придніпровської залізниці здійснений прогноз (на добу вперед) обсягу мережного трафіка на основі гібридної системи, що створена в програмі Matlab; значення MAPE складає: 6,9 % для обсягу вхідного трафіка; 7,7 % для обсягу вихідного трафіка. Виявлено, що середня похибка навчання гібридної системи зменшується при збільшенні: кількості входів (від 2 до 4); кількості термів (від 2 до 5) вхідної змінної; потужності навчальної вибірки (від 20 до 100). Значний вплив на середню похибку навчання гібридної системи має кількість термів її вхідної змінної. Визначено, що найменше значення середньої похибки навчання надає чотири-вхідна гібридна система, більш точно здійснюється навчання нейронечіткої мережі за гібридним методом. **Наукова новизна.** Отримані залежності середньої похибки навчання гібридної системи прогнозування обсягу мережного трафіка фрагмента (Дніпропетровськ–Київ) в ІТС Придніпровської залізниці від: кількості її входів, кількості термів вхідної змінної, потужності навчальної вибірки за різними методами навчання. **Практична значимість.** Прогнозування обсягу мережного трафіка в ІТС Придніпровської залізниці дозволить в реальному часі визначити перевантаження в мережі та здійснити контроль потоків даних.

Ключові слова: прогнозування; мережний трафік; обсяг; нейронечітка мережа; гібридна система; терм; функція приналежності; вибірка; адекватність; похибка

Вступ

Для прогнозування параметрів мережного трафіка використовують різні методи та методики, що широко розповсюджені в області аналізу часових рядів економічних показників [9–10]. В загальному, якщо задані n дискретних значень $\{y(t_1), y(t_2), \dots, y(t_n)\}$ в послідовні моменти часу t_1, t_2, \dots, t_n , тоді задача прогнозування полягає в прогнозі значення $y(t_{n+1})$ в майбутній момент часу t_{n+1} . Прогноз, зазвичай, виходить помилковим, але помилка залежить від використаної прогнозуючої системи. Високу ефективність прогнозу надає використання нейронних мереж [1, 11–13]. Задача прогнозування може бути розв'язана на основі таких нейронних мереж: багатошарового перцептронну (MLP), радіально-базисної мережі (RBF), узагальнено-регресійної мережі (GRNN), мережі Вольтерра, мережі Елмана та ANFIS-системи, огляд яких виконано в [7]. Нечіткі нейронні мережі (гібридні системи) покликані об'єднати в собі переваги нейронних мереж і систем нечіткого висновку. Вони дозволяють розробляти та подавати моделі систем у формі правил нечітких продукцій, для побудови яких використовують можливості нейронних мереж [5], зокрема адаптивна мережа нечіткого висновку (Adaptive-Network-Based Fuzzy Inference System, ANFIS), що реалізована в додатку Fuzzy Logic Toolbox програми Matlab [4]. До основних етапів роботи нейронечіткої мережі належать: формування бази правил системи нечіткого висновку; фазифікація вхідних змінних; агрегування; активізація; акумуляція; де-

фазифікація вихідних змінних, алгоритм функціонування такої системи надано в [6]. Зокрема, в [2] пропонується гібридна система прогнозування (на 24 години вперед) приміського пасажиропотоку, в [3] сформовано гібридну модель для прогнозування обсягів навантаження вагонів за дві попередні доби.

Мета

Розробити методику прогнозування обсягу мережного трафіка (вхідного та вихідного) на основі використання нечіткої нейронної мережі (гібридної системи) для розглянутого фрагмента (Дніпропетровськ – Київ) в ІТС Придніпровської залізниці.

Постановка задачі

Постійне збільшення обсягу мережного трафіка в ІТС Придніпровської залізниці породжує необхідність його прогнозування для попередження перевантаження в мережі та підвищення якості послуг. Одним із можливих рішень може бути метод прогнозування обсягу мережного трафіка, який дозволить уникнути перевантаження (зокрема, сервера). Для дослідження використані реальні дані трафіка найбільш важливого фрагменту (Дніпропетровськ – Київ) в ІТС Придніпровської залізниці за період 21.03 – 26.03.2016 р. Виконаний аналіз вхідного та вихідного трафіків за напрямком пошуку довгострокових залежностей (години, доби). Для наочності побудовані графіки обсягу мережного трафіка фрагменту ІТС, що розглядається. Як приклад на рис. 1 наведено вихідний трафік для довжини фрагмента часового ряду 24 години за різні дні тижня.

ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНІ ТЕХНОЛОГІЇ ТА МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

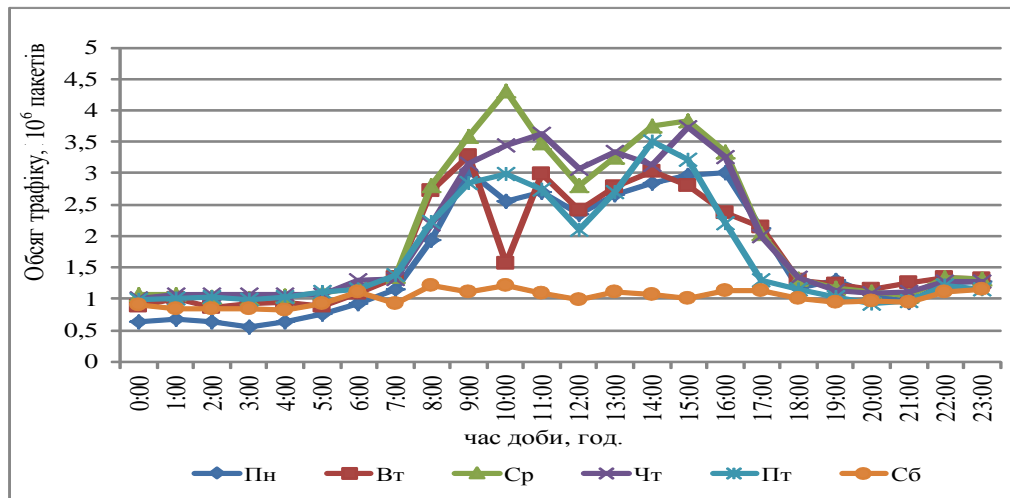


Рис. 1. Обсяг вихідного трафіка в ІТС (Дніпропетровськ – Київ)

На рис.1 проглядається тенденція поведінки обсягу мережного трафіка за тиждень: він приблизно однаковий в понеділок, вівторок, четвер та п'ятницю; існують певні закономірні зміни за визначені періоди. Так зокрема, менший обсяг трафіка та більш-менш стабільний з 00:00 до 7:00, значний та нестабільний обсяг трафіка з 8:00 до 17:00, знову менший та відносно незмінний обсяг трафіка з 18:00 до 23:00. У середу обсяг мережного трафіка найбільш високий, а у вихідні дні – значно менший ніж в будні дні. Із рисунка видно, що обсяг вихідного трафіка в середу порівняно з понеділком, вівторком, четвергом та п'ятницею перевищує приблизно в 1,3 рази. Для побудови прогнозу (на добу вперед) обсягу мережного трафіка обрано інтервал з 8:00 до 17:00, де відбуваються його значні коливання, але за дні тижня (понеділок, вівторок, четвер, п'ятницю), коли характер трафіка приблизно однаковий. Тому прийнято рішення здійснити прогноз (на добу вперед) обсягу трафіка $x(t)$ на базі даних попередніх трьох днів: $x(t-1)$, $x(t-2)$, $x(t-3)$.

Методика

якщо $x(t-1)=min$ та $x(t-2)=min$ та $x(t-3)=min$, то $x(t)=1$;
 якщо $x(t-1)=min$ та $x(t-2)=min$ та $x(t-3)=max$, то $x(t)=2$;
 якщо $x(t-1)=min$ та $x(t-2)=max$ та $x(t-3)=min$, то $x(t)=3$;
 якщо $x(t-1)=min$ та $x(t-2)=max$ та $x(t-3)=max$, то $x(t)=4$;
 якщо $x(t-1)=max$ та $x(t-2)=min$ та $x(t-3)=min$, то $x(t)=5$;

1. *Підготовка вибірок.* Для здійснення прогнозу необхідно підготувати наступні вибірки: навчальну, тестову, контрольну. Від сформованої вибірки залежить наскільки ефективним буде процес навчання та тестування, а також залежить здатність мережі вирішувати поставлені перед нею проблеми під час експлуатації. Для підготовки вибірки зроблений спеціальний масив зі 100 прикладів, що близькі до реальних. При формуванні навчальної вибірки зі створеного масиву використано перші 50 значень, для тестової – наступні 50 значень. Для формування контрольної вибірки використані реальні дані четвертого дня, який не розглядається.

2. *Створення нейронечіткої мережі в Matlab.* Задача прогнозування обсягу трафіка (вхідного, вихідного) на ділянці Дніпропетровськ – Київ зводиться до задачі прогнозування часових рядів, зазвичай для таких завдань вибирається систему типу Сугено. Для лінгвістичної оцінки кожна вхідна змінна має два терми (максимальне та мінімальне значення), функцію приналежності обрано гаусовську (gaussmf), для оцінки результуючої змінної задано функцію приналежності типу linear. У редакторі бази знань задані правила нечіткого висновку, мають наступний вигляд:

ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНІ ТЕХНОЛОГІЇ ТА МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

якщо $x(t-1)=max$ та $x(t-2)=min$ та $x(t-3)=max$, то $x(t)=6$;
 якщо $x(t-1)=max$ та $x(t-2)=max$ та $x(t-3)=min$, то $x(t)=7$;
 якщо $x(t-1)=max$ та $x(t-2)=max$ та $x(t-3)=max$, то $x(t)=8$.

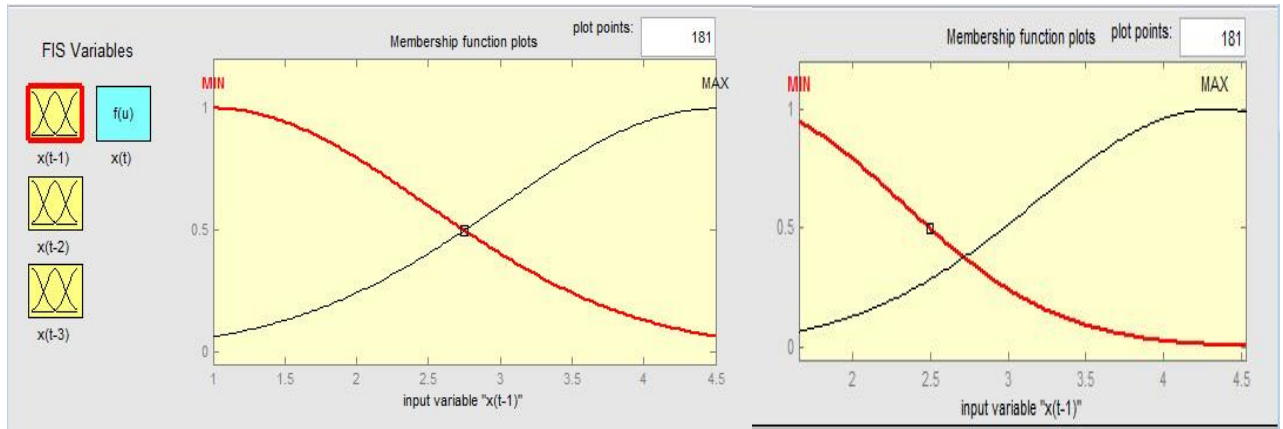


Рис. 2. Функція приналежності першої вхідної змінної до та після навчання системи

Структура спроектованої системи нечіткого висновку наведена на рис. 3.

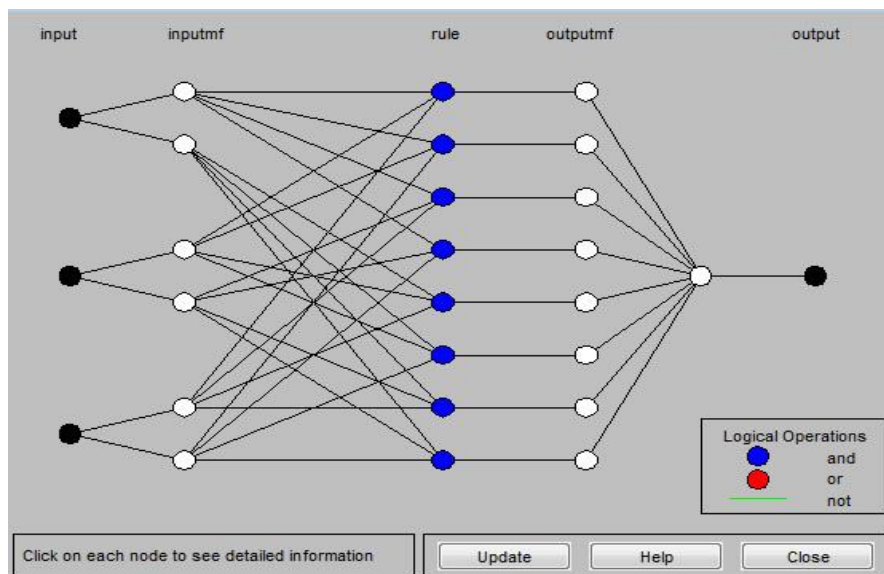


Рис. 3. Структура спроектованої гібридної системи

Як видно з рис. 3, система має 5 шарів. Перший шар (input) – має три вузли ($x(t-3)$, $x(t-2)$, $x(t-1)$), на які подаються вхідні дані. Перший шар виконує роздільну фазифікацію кожної змінної, визначаючи для кожного j -го правила висновку значення коефіцієнта приналежності відповідно до застосованої функції фазифікації. Другий шар (inputmf) складається з $3 \cdot 2 = 6$ вершин, оскільки кожній вхідній змінній відповідає 2 терми, виконує агрегування окремих змінних x_i , визначаючи результуюче значення

коефіцієнта приналежності для вектора x (рівень активізації правила висновку); цей шар непараметричний. Третій шар (rule) являє собою генератор функції TSK; це параметричний шар, в якому адаптації підлягають лінійні ваги, що визначають функцію слідства моделі TSK. Четвертий шар (outputmf) складається з функцій приналежності для кожного правила нечіткого висновку (кількість вершин цього шару відповідає кількості правил $2^3 = 8$); цей шар непараметричний. П'ятий шар (output) – нормалі-

ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНІ ТЕХНОЛОГІЇ ТА МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

зуючий, він має єдину вершину, яка відповідає виходу системи; цей шар непараметричний.

3. *Навчання нейронечіткої мережі.* Під час навчання як метод оптимізації (optim. method) вибраний гібридний метод (hybrid), що поєднує в собі комбінацію методу найменших квадратів та методу зменшеного зворотного градієнта; кількість ітерацій навчання (epochs) – 40. Для прикладу графік функції приналежності першої вхідної змінної до та після навчання системи наведено на рис. 2.

4. *Тестування гібридної системи.* Тестування гібридної системи виконано на тестовій вибірці. Результат тестування по рівняно з результатом навчання системи наведено на рис. 4.

5. *Аналіз адекватності гібридної системи.* Для оцінки якості та точності прогнозу створе-

ної гібридної системи розраховано MAPE (Mean Absolute Percentage Error) за формулою:

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|Z(t) - Z_1(t)|}{Z(t)} \cdot 100\%, \quad (1)$$

де $Z(t)$ – реальні дані в момент часу t ; $Z_1(t)$ – прогнозовані дані в момент часу t ; N – кількість годин.

Прогнозування обсягу мережного трафіка виконувалося з 8:00 до 17:00 (загальна кількість годин $N=10$). Значення MAPE складає: 6,9 % для прогнозу обсягу вхідного трафіка, 7,7 % для прогнозу обсягу вихідного трафіка. Як приклад фактичний та прогнозований обсяги вихідного трафіків в ІТС Придніпровської залізниці (Дніпропетровськ – Київ) наведено на рис. 5.

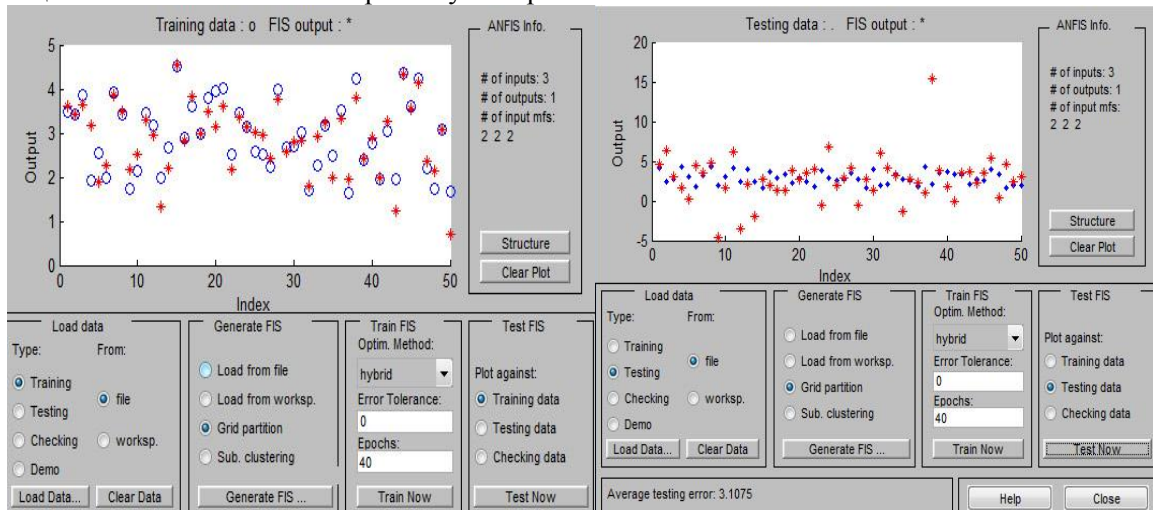


Рис. 4. Результати навчання та тестування нейронечіткої мережі

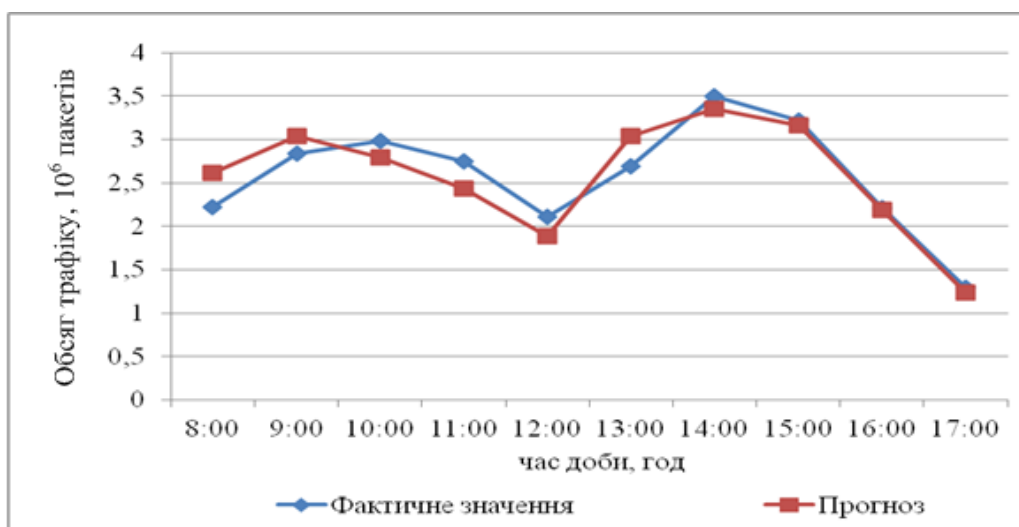


Рис. 5. Фактичний та прогнозований обсяги вихідного трафіків

ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНІ ТЕХНОЛОГІЇ ТА МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

Результати

1. Дослідження залежності середньої похибки навчання гібридної системи від кількості її входів. Виконано дослідження середньої похибки навчання створеної гібридної системи при різній кількості її входів: 2, 3, 4. У всіх експериментах довжина навчальної вибірки складала 50 прикладів, кількість епох – 40,

навчання системи відбувалось за гібридним методом. За отриманими даними побудовані графіки залежності середньої похибки навчання гібридної системи від кількості її входів для вхідного (вихідного) трафіків в ІТС Придніпровської залізниці для розглянутого фрагменту Дніпропетровськ – Київ та наведені на рис. 6.

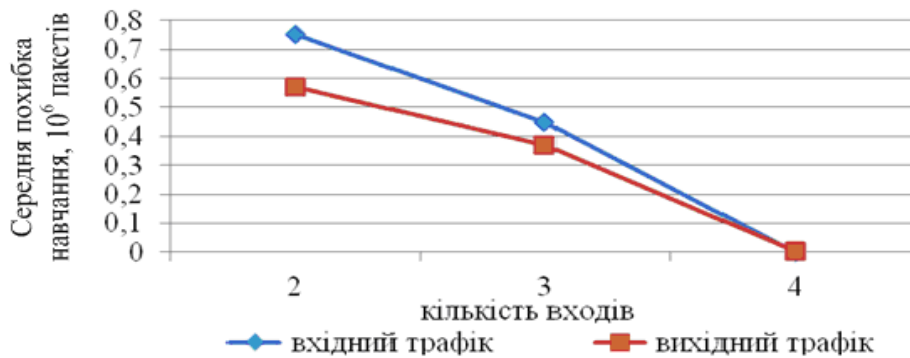


Рис. 6. Залежність середньої похибки навчання гібридної системи від кількості її входів

Із рисунку видно, що найменше значення середньої похибки навчання гібридної системи: $0,27 \cdot 10^{-3} \cdot 10^6 = 2,7 \cdot 10^2$ пакетів для вхідного трафіка; $0,19 \cdot 10^{-2} \cdot 10^6 = 19 \cdot 10^2$ для вихідного трафіка надає чотири-вхідна гібридна система при довжині навчальної вибірки із 50 прикладів.

2. Дослідження залежності середньої похибки навчання гібридної системи від кількості термів її вхідної змінної. Дослідження виконувалися на створеній гібридній системі, що має 3

вхідні змінні; у всіх експериментах довжина навчальної вибірки складалась із 50 прикладів. Проаналізуємо значення середньої похибки навчання гібридної системи від кількості термів її вхідної змінної: 2, 3, 5.

На основі отриманих значень побудовані графіки залежності середньої похибки навчання гібридної системи від кількості термів її вхідної змінної за різними методами навчання та наведені на рис. 7.

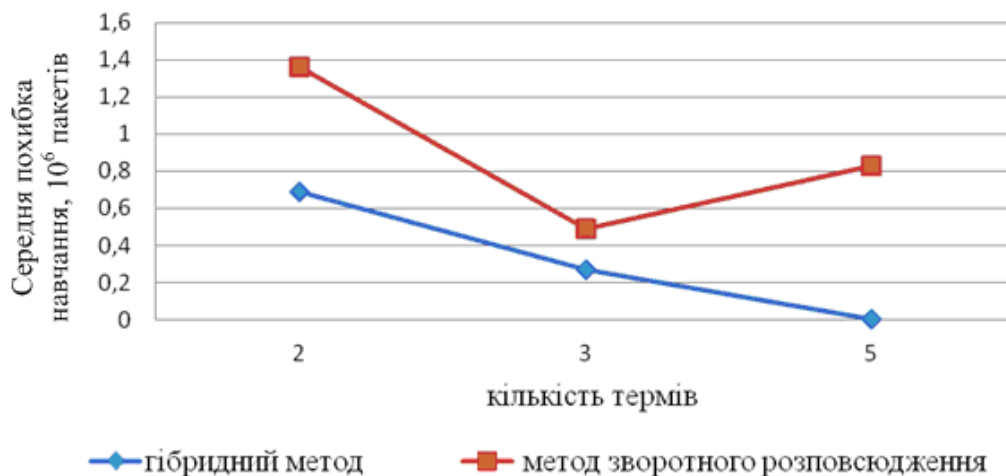


Рис. 7. Залежність середньої похибки навчання гібридної системи від кількості термів її вхідної змінної

ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНІ ТЕХНОЛОГІЇ ТА МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

Із рисунку видно, що при збільшенні кількості термів (з 2 до 5) середня похибка навчання гібридної системи зменшується: з $0,69 \cdot 10^6$ до $0,45 \cdot 10^{-5} \cdot 10^6 = 4,5$ пакетів за гібридним методом навчання; з $1,36 \cdot 10^6$ до $0,83 \cdot 10^6$ пакетів за методом зворотного розповсюдження. Таким чином, навчання три-вхідної гібридної системи (5 термів для кожної вхідної змінної) проходить більш точно за гібридним методом, ніж за методом зворотного розповсюдження.

3. Дослідження залежності середньої похибки навчання гібридної системи від потужності навчальної вибірки. Для виконання дослідження взято навчальну вибірку різної довжини: 20, 50, 100. Дослідження виконувалися на гібридній системі, що має 3 вхідні змінні, цикл навчання склав 100 епох. За результатами експериментів побудовані графіки залежності середньої похибки навчання гібридної системи від потужності навчальної вибірки за алгоритмами навчання та наведені на рис. 8.

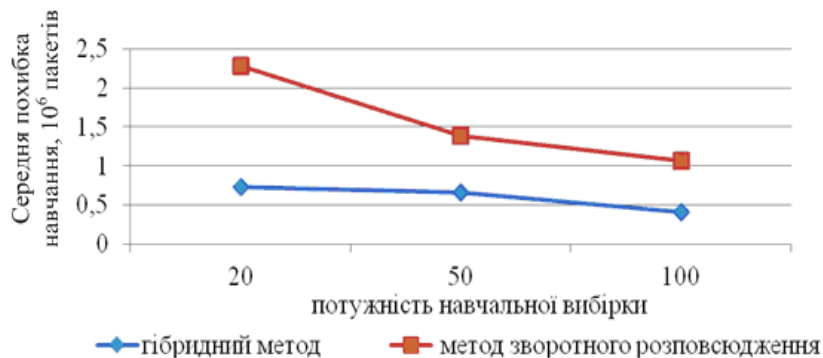


Рис. 8. – Залежність середньої похибки навчання гібридної системи від потужності навчальної вибірки

Із рисунку видно, що при збільшенні потужності навчальної вибірки (від 20 до 100 прикладів) на 3-вхідну гібридну систему її середня похибка навчання зменшується: з $0,72 \cdot 10^6$ до $0,41 \cdot 10^6$ пакетів за гібридним методом навчання; з $2,28 \cdot 10^6$ до $1,07 \cdot 10^6$ пакетів за методом зворотного розповсюдження. Таким чином, більш точно здійснюється навчання гібридної системи за гібридним методом при потужності навчальної вибірки із 100 прикладів.

Наукова новизна та практична значимість

Наукова новизна полягає в тому, що знайдені залежності середньої похибки навчання гібридної системи прогнозування обсягу мережного трафіка фрагмента (Дніпропетровськ – Київ) в ІТС Придніпровської залізниці від: кількості її входів, кількості термів вхідної змінної, потужності навчальної вибірки за різними методами навчання. Практична значимість полягає в тому, що прогнозування обсягу мережного трафіка в ІТС дозволить в реальному часі

визначити перевантаження в мережі та здійснити контроль потоків даних.

Висновки

1. Виконаний аналіз обсягу мережного трафіка (вхідного та вихідного) в ІТС Придніпровської залізниці (Дніпропетровськ – Київ) на основі реальних даних. Для одержання прогнозу (на добу вперед) обсягу мережного трафіка обрано інтервал з 8 до 17 години, де відбуваються значні коливання, але за той період тижня (понеділок, вівторок, четвер, п'ятниця), коли характер трафіка приблизно однаковий.

2. Підготовлені навчальна, тестова та контрольна вибірки на основі фактичних даних за період 21.03. – 26.03.2016 р. Прогноз обсягу мережного трафіка в ІТС Придніпровської залізниці (Дніпропетровськ – Київ) здійснений з використанням нейронечіткої мережі (гібридної системи), що створена в програмі Matlab. На вхід гібридної системи подається обсяг мережного трафіка за попередні три доби; прогнозування обсягу мережного трафіка виконувалось з 8 по 17 (загальна кількість годин $N=10$);

ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНІ ТЕХНОЛОГІЇ ТА МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

значення $MARE$ складає: 6,9 % для вхідного трафіка; 7,7 % для вихідного трафіка.

3. Виконані експериментальні дослідження залежності середньої похибки навчання гібридної системи від: кількості її входів (перше дослідження), кількості терм вхідної змінної (друге дослідження), потужності навчальної вибірки (третє дослідження) за різними методами навчання: гібридним, зворотного розповсюдження. Значний вплив на середню похибку навчання гібридної системи має кількість термів вхідної змінної. В ІТС Придніпровської залізниці (Дніпропетровськ – Київ):

– за результатами першого дослідження найбільш точний прогноз обсягу вхідного трафіка (похибка навчання $2,7 \cdot 10^2$) та вихідного трафіка (похибка навчання $19 \cdot 10^2$) надає чотири-вхідна гібридна система при довжині навчальної вибірки із 50 прикладів;

– за результатами другого дослідження при збільшенні кількості термів (від 2 до 5) її вхідної змінної середня похибка навчання зменшується: з $0,69 \cdot 10^6$ до $4,5$ пакетів за гібридним методом; з $1,36 \cdot 10^6$ до $0,83 \cdot 10^6$ пакетів за методом зворотного розповсюдження. Таким чином, навчання 3-вхідної гібридної системи, що має 5 термів для кожної вхідної змінної, проходить більш точно за гібридним методом;

– за результатами третього дослідження при збільшенні потужності навчальної вибірки (від 20 до 100 прикладів) на три-вхідну гібридну систему середня похибка навчання зменшується: з $0,72 \cdot 10^6$ до $0,41 \cdot 10^6$ пакетів за гібридним методом; з $2,28 \cdot 10^6$ до $1,07 \cdot 10^6$ пакетів за методом зворотного розповсюдження. Таким чином, більш точно здійснюється навчання за гібридним методом при потужності навчальної вибірки із 100 прикладів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Герасина, А. В. Адаптивное нечеткое прогнозирование трафика в информационных телекоммуникационных сетях / А. В. Герасина // Системи обробки інформації : зб. наук. пр. / Харк. ун-т повітр. сил ім. Івана Кожедуба. – Харків, 2013. – Вип. 9 (116). – С. 141–145.
2. Константінов, Д. В. Формування адаптивної технології приміських залізничних перевезень : автореф. дис. ... канд. техн. наук : 05.22.01 / Константінов Денис Володимирович ; Укр. держ. акад. залізн. трансп. – Харків, 2010. – 20 с.
3. Костенніков, О. М. Удосконалення технології формування місцевого вагонопотоку на дільниці в умовах сезонного коливання обсягів навантаження : автореф. дис. ... канд. техн. наук : 05.22.01 / Костенніков Олексій Михайлович ; Укр. держ. акад. залізн. трансп. – Харків, 2012. – 20 с.
4. Леоненков, А. В. Нечеткое моделирование в среде MatLAB и fuzzy TECH / А. В. Леоненков. – Санкт-Петербург : БХВ-Петербург, 2003. – 736 с.
5. Манусов, В. З. Краткосрочное прогнозирование электрической нагрузки на основе нечеткой нейронной сети и ее сравнение с другими методами / В. З. Манусов, Е. В. Бирюков. – Изв. Томск. политехн. ун-та. – 2006. – Т. 309, № 6. – С. 153–158.
6. Мещеряков, В. А. Моделирование адаптивной системы нейронечеткого управления рабочим процессом стрелового крана / В. А. Мещеряков, И. В. Денисов // Проектирование инженер. и науч. приложений в среде MatLAB : материалы V Междунар. науч. конф. – Харьков, 2011. – С. 367–375.
7. Пахомова, В. М. Розробка підсистеми оперативного прогнозування простоїв прибуваючих поїздів на основі ANFIS-системи / В. М. Пахомова, С. Ю. Дмитрієв // Інформ.-керуючі системи на залізн. трансп. – 2013. – № 4. – С. 46–55.
8. Пахомова, В. М. Дослідження інжинірингу трафіка в комп'ютерній мережі УЗ за технологією MPLS TE / В. М. Пахомова // Наука та прогрес транспорту. – 2015. – № 1 (55). – С. 139–147. doi: 10.15802/STP2015/38262.
9. Покровская, М. А. Метод прогнозирования изменения трафика с использованием нейросетевой модели / М. А. Покровская // T-Comm – Телекоммуникации и Транспорт. – 2012. – № 6. – С. 27–30.
10. Сравнительный анализ методов прогнозирования трафика в телекоммуникационных системах [Электронный ресурс] / К. М. Руккас, Ю. В. Соляник, К. А. Овчинников, О. О. Давид // Проблемы телекоммуникаций. – 2014. – № 1 (13). – С. 84–95. – Режим доступа: http://pt.journal.kh.ua/2014/1/1/141_rukkas_analys.pdf. – Назва з екрана. – Перевірено : 22.11.16.
11. Chabaa, S. Identification and prediction of internet traffic using artificial neural networks / S. Chabaa, A. Zeroual, J. Antari // J. of Intelligent Learning Systems and Applications. – 2010. – Vol. 02. –

- Iss. 03. – P. 147–155. doi: 10.4236/jilsa.2010.23018.
12. Gowrishankar, S. A time series modeling and prediction of wireless network traffic / S. Gowrishankar, P. S. Satyanarayana // Intern. J. of Interactive Mobile Technologies (IJIM). – 2009. – Vol. 3. – Iss. 1. – P. 53–62. doi: 10.3991/ijim.v3i1.284.
13. Multi-scale Internet traffic forecasting using neural networks and time series methods / P. Cortez, M. Rio, M. Rocha, P. Sousa // Expert Systems. – 2010. – Vol. 29. – Iss. 2. – P. 143–155. doi: 10.1111/j.1468-0394.2010.00568.x.

В. Н. ПАХОМОВА^{1*}

^{1*}Каф. «Електронні вычислительные машины», Днепропетровский национальный университет железнодорожного транспорта имени академика В. Лазаряна, ул. Лазаряна, 2, Днипро, Украина, 49010, тел. +38 (056) 373 15 89, эл. почта viknik.p1988@mail.ru, ORCID 0000-0001-8346-0405

ПРОГНОЗИРОВАНИЕ ОБЪЕМА СЕТЕВОГО ТРАФИКА В ИНФОРМАЦИОННО-ТЕЛЕКОМУНИКАЦИОННОЙ СИСТЕМЕ ПРИДНЕПРОВСКОЙ ДОРОГИ НА ОСНОВЕ НЕЙРОНЕЧЕТКОЙ СЕТИ

Цель. Постоянное увеличение объема сетевого трафика в информационно-телекоммуникационной системе (ИТС) Приднепровской железной дороги приводит к необходимости определения в реальном времени перегрузки в сети и осуществления контроля потоков данных. Одним из возможных решений является метод прогнозирования объема сетевого трафика (входного и выходного) с использованием нейросетевой технологии, что позволит избежать перегрузки сервера и повысить качество услуг. **Методика.** В работе проведен анализ существующего сетевого трафика в ИТС Приднепровской железной дороги и подготовка выборки: учебной, тестовой, контрольной, а также создание в программе Matlab нейронечеткой сети (гибридной системы) и организация на соответствующих выборках следующих этапов: обучение, тестирование, анализ адекватности прогноза. **Результаты.** Для фрагмента (Днепропетровск – Киев) в ИТС Приднепровской железной дороги осуществлен прогноз (на сутки вперед) объема сетевого трафика на основе гибридной системы, созданной в программе Matlab; значение MAPE составляет: 6,9 % для объема входящего трафика; 7,7 % для объема исходящего трафика. Выявлено, что средняя ошибка обучения гибридной системы уменьшается при увеличении: количества входов (от 2 до 4); количества термов (от 2 до 5) входной переменной; мощности обучающей выборки (от 20 до 100), большое влияние на среднюю ошибку обучения гибридной системы оказывает число термов ее входной переменной. Определено, что наименьшее значение ошибки обучения дает 4-входная гибридная система, более точно осуществляется обучение нейронечеткой сети по гибричному методу. **Научная новизна.** Получены зависимости средней ошибки обучения гибридной системы прогнозирования объема сетевого трафика фрагмента (Днепропетровск–Киев) в ИТС Приднепровской дороги от: количества входов, количества термов входной переменной, мощности обучающей выборки при различных методах обучения. **Практическая значимость.** Прогнозирование объема сетевого трафика в ИТС Приднепровской дороги позволит в реальном времени определить перегрузки в сети и осуществить контроль потоков данных.

Ключевые слова: прогнозирование; сетевой трафик; объем; нейронечеткая сеть; гибридная система; терм; функция принадлежности; выборка; адекватность; ошибка

V. M. PAKHOMOVA^{1*}

^{1*}Dep. «Electronic Computing Machines», Dnipropetrovsk National University of Railway Transport named after Academician V. Lazaryan, Lazaryan St., 2, Dnipro, Ukraine, 49010, tel. +38 (056) 373 15 89, e-mail viknik.p1988@mail.ru, ORCID 0000-0001-8346-0405

NETWORK TRAFFIC FORCASTING IN INFORMATION-TELECOMMUNICATION SYSTEM OF DNIEPER RAILWAYS BASED ON NEURO-FUZZY NETWORK

ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНІ ТЕХНОЛОГІЇ ТА МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

Purpose. Continuous increase in network traffic in the information-telecommunication system (ITS) of Dnieper Railways leads to the need to determine the real-time network congestion and to control the data flows. One of the possible solutions is a method of forecasting the volume of network traffic (inbound and outbound) using neural network technology that will prevent from server overload and improve the quality of services. **Methodology.** Analysis of current network traffic in ITS of Dnieper Railways and preparation of sets: learning, test and validation ones. Creation of neuro-fuzzy network (hybrid system) in Matlab program and organization of the following phases on the appropriate sets: learning, testing, forecast adequacy analysis. **Findings.** For the fragment (Dnipropetrovsk-Kyiv) in ITS of Dnieper Railways we made a forecast (day ahead) for volume of network traffic based on the hybrid system created in Matlab program; MAPE values are as follows: 6.9% for volume of inbound traffic; 7.7% for volume of outbound traffic. It was found that the average learning error of the hybrid system decreases in case of increase in: the number of inputs (from 2 to 4); the number of terms (from 2 to 5) of the input variable; learning sample power (from 20 to 100); a significant impact on the average learning error of the hybrid system is caused by the number of terms of its input variable. It was determined that the lowest value of the average learning error is provided by 4-input hybrid system, it ensures more accurate learning of the neuro-fuzzy network by the hybrid method. **Originality.** The work resulted in the dependences for the average hybrid system error of the network traffic volume forecasting for the fragment (Dnipropetrovsk – Kyiv) in ITS of Dnieper Railways on: the number of its inputs, the number of input variable terms, the learning sample power for different learning methods. **Practical value.** Forecasting of network traffic volume in ITS of Dnieper Railways will allow for real-time identification of the network congestion and control of data flows.

Keywords: forecasting; network traffic; volume; neuro-fuzzy network; hybrid system; term; membership function; set; adequacy; error

REFERENCES

1. Gerasina A.V. Adaptivnoye nechetkoye prognozirovaniye trafika v informatsionnykh telekommunikatsionnykh setyakh [Adaptive fuzzy prediction of traffic in information and telecommunication networks]. *Systemy obrobky informatsii – Information Processing Systems*, 2013, issue 9 (116), pp. 141-145.
2. Konstantinov D.V. *Formuvannya adaptivnoi tekhnologii prymskykh zaliznychnykh perevezhen*. Avtoreferat Diss. [Formation of adaptive technology of commuter rail transportation. Author's abstract]. Kharkiv, 2010. 20 p.
3. Kostiennikov O.M. *Udoskonalennia tekhnologii formuvannia mistsevoho vahonopotoku na dilnytsi v umovakh sezonnoho kolyvannia obsiahiv navantazhennia*. Avtoreferat Diss. [Improving the technology of forming a local car traffic volume at the section in terms of seasonal fluctuations in load. Author's abstract]. Kharkiv, 2012. 20 p.
4. Leonenkov A.V. *Nechetkoye modelirovaniye v srede MatLAB i fuzzy TECH* [Fuzzy modeling in MatLAB and fuzzy TECH environment]. Saint-Petersburg, BKhV-Peterburg Publ., 2003. 736 p.
5. Manusov V.Z., Biryukov Ye.V. *Kratkosrochnoye prognozirovaniye elektricheskoy nagruzki na osnove nechetkoy neyronnoy seti i yeye sravneniye s drugimi metodami* [Short-term forecasting of electric load based on fuzzy neural network and its comparison with other methods]. *Izvestiya Tomskogo politekhnicheskogo universiteta – Bulletin of the Tomsk Polytechnic University*, 2006, vol. 309, no. 6, pp. 153-158.
6. Meshcheryakov V.A., Denisov I.V. *Modelirovaniye adaptivnoy sistemy neyronechetkogo upravleniya rabochim protsessom strelovogo krana* [Adaptive system modeling of neuro-fuzzy control of operational process for jib crane]. *Materialy V Mezhdunarodnoy nauchnoy konferentsii «Proyektirovaniye inzhenernykh i nauchnykh prilozheniy v srede MatLAB»* [Proc. of Vth Intern. Sci. Confererence «Design of Engineering and Scientific Applications in Matlab Environment»]. Kharkov, 2011, pp. 367-375.
7. Pakhomova V.M., Dmitriiev S.Yu. *Rozrobka pidsystemy operativnoho prohozuvannia prostoiu prybuvaiuchykh poizdiv na osnovi ANFIS-systemy* [Subsystem development of operational forecasting of inactive coming trains based on ANFIS-system]. *Informatsiino-keruiuchi systemy na zaliznychnomu transporti – Information and Control Systems at Railway Transport*, 2013, no. 4, pp. 46-55.
8. Pakhomova V.M. *Doslidzhennia inzhynirynhu trafika v kompiuternii merezhi UZ za tekhnolohiieiu MPLS TE* [Research of engineering traffic in computer of Ukrzaliznytsia network using MPLS TE technology]. *Nauka ta prohres transportu – Science and Transport Progress*, 2015, no. 1 (55), pp. 139-147. doi: 10.15802/stp2015/38262.

ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНІ ТЕХНОЛОГІЇ ТА МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ

9. Pokrovskaya M.A. Metod prognozirovaniya izmeneniya trafika s ispolzovaniyem neurosetevoy modeli [Prediction method of traffic change with the use of neural network model]. *T-Comm – Telekommunikatsii i Transport – T-Comm – Telecommunications and Transport*, 2012, vol. 6, no. 6, pp. 27-30.
10. Rukkas K.M., Solyanik Yu.V., Ovchinnikov K.A., David O.O. Sravnitelnyy analiz metodov prognozirovaniya trafika v telekommunikatsionnykh sistemakh (Comparative analysis of traffic prediction techniques in telecommunication systems). *Problemy telekommunikatsiy – Problems of Telecommunications*, 2014, no. 1 (13), pp. 84-95. Available at: http://pt.journal.kh.ua/2014/1/1/141_rukkas_analysis.pdf (Accessed 22 November 2016).
11. Chabaa S., Zeroual A., Antari J. Identification and prediction of internet traffic using artificial neural networks. *Journal of Intelligent Learning Systems and Applications*, 2010, vol. 02, issue 03, pp. 147-155. doi: 10.4236/jilsa.2010.23018.
12. Gowrishankar S., Satyanarayana P.S. A time series modeling and prediction of wireless network traffic. *Intern. Journal of Interactive Mobile Technologies (iJIM)*, 2009, vol. 3, issue 1, pp. 53-62. doi: 10.3991/ijim.v3i1.284.
13. Cortez P., Rio M., Rocha M., Sousa P. Multi-scale internet traffic forecasting using neural networks and time series methods. *Expert Systems*, 2010, vol. 29, no. 2, pp. 143-155. doi: 10.1111/j.1468-0394.2010.00568.x.

Стаття рекомендована до публікації д. т. н., проф. В. В. Ткачовим (Україна);
д. т. н., проф. В. В. Скалозубом (Україна)

Надійшла до редколегії: 26.09.2016

Прийнята до друку: 05.12.2016