

МОДЕЛЮВАННЯ СИСТЕМИ ОПЕРАТИВНОГО ПРОГНОЗУВАННЯ ПАСАЖИРОПОТОКІВ В ПРИМІСЬКОМУ СПОЛУЧЕННІ НА ОСНОВІ ВИКОРИСТАННЯ ІНТЕЛЕКТУАЛЬНИХ ТЕХНОЛОГІЙ

Було розглянуто питання, пов'язане з удосконаленням системи прогнозування залізничних приміських пасажиропотоків на основі використання гібридних систем прогнозування, та розроблено нову нечітко-нейро-генетичну модель прогнозування. Розроблені та описані в статті методи дозволяють з використанням інформаційних технологій реалізувати на їх основі реальну систему прогнозування пасажиропотоків в приміському сполученні

Т.В. Буцько

Доктор технічних наук, професор, завідувача кафедрою
Кафедра "Управління експлуатаційною роботою"*
Контактний тел.: (057) 730-10-89

Д.В. Константінов

Аспірант, асистент*
Кафедра "Управління експлуатаційною роботою"*
Контактний тел.: (057) 730-10-88

Т.О. Дервянко

Студентка*

*Українська державна академія залізничного транспорту
вул. Феєрбаха, 7, м. Харків, Україна

1. Вступ

Характерною особливістю приміських пасажирських перевезень є високий рівень невизначеності і коливань пасажиропотоків, що ускладнює ефективність роботи приміського сектору пасажирського господарства Укрзалізниці. Приміські пасажиропотоки характеризуються значною сезонною та добовою нерівномірністю, а також, притаманною лише їм, значною погодинною нерівномірністю в межах доби. Організація приміських перевезень на сучасному етапі здійс-

нюється на основі застарілих методів управління і не враховує існуючого коливання попиту на перевезення. Це потребує впровадження нових гнучких технологій в організацію перевізного процесу, що засновані на принципах адаптації та раціоналізації використання рухомого складу в умовах дефіциту.

В умовах розвитку інтелектуальних технологій вирішення оперативних завдань управління потребує формування та впровадження системи підтримки прийняття рішень (СППР) оперативного персоналу стосовно задачі оптимізації роботи на основі раціо-

нального використання нових та існуючих транспортних ресурсів.

Реалізація СППР безпосередньо заснована на розробці та впровадженні нової технології реалізації прогнозування пасажиропотоків, що здійснюється на основі стратегії управління за віддаленим горизонтом [9]. Використання стратегії дає змогу реалізації прогнозування на встановлений період, з можливістю подальшої розробки та здійснення оптимальних регулювальних заходів на основі графіку курсування приміських поїздів. Тому, враховуючи особливості нерівномірностей приміських пасажиропотоків, період прогнозування, з метою збереження точності результатів, має бути незначний, строком до 24 годин.

З урахуванням всіх зазначених умов реалізація системи прогнозування має бути заснована на методиці, що враховує нечіткість вхідної інформації і має можливість швидкої адаптації в короткотермінові строки до змін умов в визначений період планування.

Враховуючи складність розв'язання поставленої задачі, необхідним є аналіз можливостей нових інформаційних технологій в області "Soft Computing" [1] і реалізація на їх основі гібридних систем прогнозування [9]. Реалізація цих систем передбачає застосування методики здобуття знань з експериментальних даних на основі нечітко-нейро-генетичного моделювання. Використання генетичних обчислень дозволяє спроектувати базу правил на основі нечіткого висновку, яка представляє собою ядро прогнозуючої системи, з подальшим представленням її у вигляді графа нейро-нечіткої мережі, що має можливість тонкої настройки змінних параметрів, тобто механізм навчання.

2. Зв'язок з науковими програмами

Згідно з Концепцією Державної програми реформування залізничного транспорту від 27 грудня 2006 р. N 651-р та Програмою реструктуризації галузі одним із основних напрямків розвитку організації приміських пасажирських перевезень є формування та подальше

удосконалення гнучких технологій поїздоутворення, що є одним з напрямків вирішення проблеми ресурсозбереження [10].

3. Аналіз останніх досліджень

Запропонований в роботі підхід було раніше представлено в роботі [9] для реалізації прогнозування пасажиропотоків в дальньому та місцевому сполученнях, що не враховує особливостей коливань пасажиропотоків в приміському сполученні, зокрема подинної нерівномірності. Використання даної методики для виконання прогнозів в приміському сполученні потребує перенавчання, зміни та адаптації системи до особливостей приміських пасажиропотоків.

4. Постановка завдання

Метою роботи є створення на принципах [9] нової моделі прогнозування приміських пасажиропотоків у реальному часі, адаптованої до особливостей їх динамічних подинних коливань протягом доби. Це обумовлює необхідність перенавчання та адаптації раніше запропонованої методики до особливостей більш значної невизначеності притаманної приміським перевезенням з подальшим створенням нової моделі, що буде здатна адекватно працювати у відповідному середовищі.

5. Основний матеріал

Задача прогнозування в загальному вигляді зводиться до оцінки майбутніх значень впорядкованих в часі даних на основі аналізу та виявлення складних залежностей у вже існуючих даних. Враховуючи значні коливання пасажиропотоків протягом доби та нерівномірність графіку відправлення приміських поїздів на напрямку, що розглядається (як і по всіх інших на-

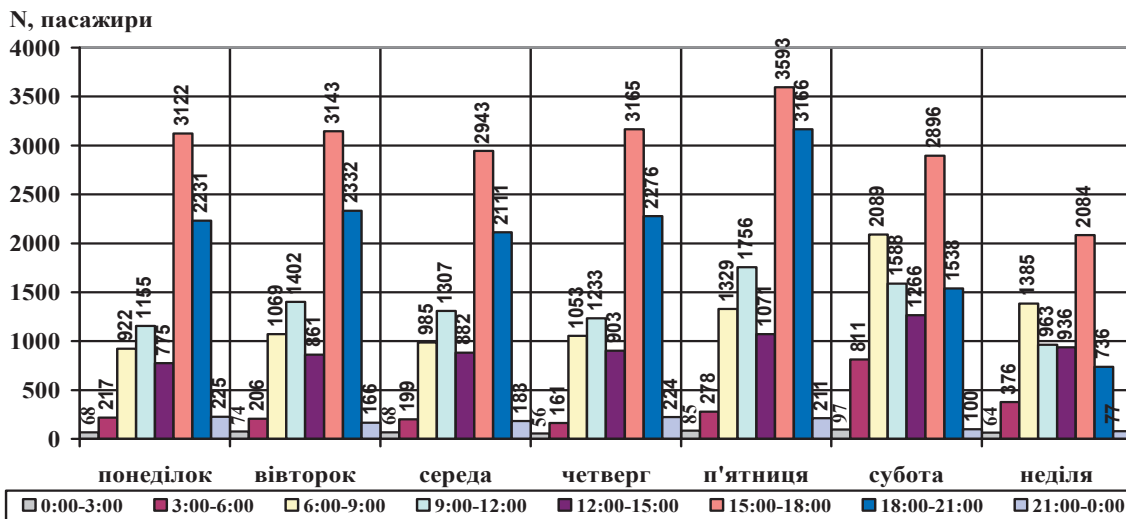


Рисунок 1. Динаміки відправлення пасажирів в приміському сполученні по годинам доби в Казачанському напрямку ст. Харків-Пас

прямах), з метою оптимізації процесу прогнозування та адаптації до коливань і нечіткості вхідної інформації вирішено розділити вихідну вибірку (дані відправлення за добу) на періоди (8 періодів по 3 години) (см. рис.1). Це дасть змогу наблизитись до більш чітких значень і, на основі спрогнозованих даних, отримати достатній період часу на проведення оперативних регулювальних заходів.

Вихідною інформацією для проведення прогнозування є тимчасовий ряд, що складається з d значень ряду в послідовні моменти часу:

$$X_{t-d} = (X_{t-1}, \dots, X_{t-d}) \quad (1)$$

де d - глибина занурення.

Для проведення прогнозування необхідно трансформувати вихідні дані до ковзаючого вікна [9]. Дана методика припускає використання двох вікон W_d і W_q з фіксованими розмірами d та q , що переміщуються з одинарним кроком по тимчасовій послідовності історичних даних, починаючи з першого елементу. Виникаюча на кожному кроці пара

$$W_d \rightarrow W_q \quad (2)$$

використовується як елемент навчальної вибірки виду "вхід-вихід" $\langle \bar{X}_r, \bar{y}_r \rangle$, $r = \overline{1, M}$, де $\bar{X}_r = \{ \bar{x}_i^r \}$, $i = \overline{1, n}$ - вхідний вектор в парі r та відповідний вихід \bar{y}_r .

Розглянемо прогнозування тимчасового ряду як задачу апроксимації невідомої функції багатьох змінних:

$$y_{t+1} = f(X_t, X_{t-d}), \quad (3)$$

що засновано на деяких експериментальних даних $\langle \bar{X}_r, \bar{y}_r \rangle$, які задані історією даного тимчасового ряду та приведені до ковзаючого вікна.

Для визначення глибини занурення в роботі проведено аналіз тимчасової послідовності даних про кількість відправлених пасажирів по кожному напрямку курсування приміських поїздів ст. Харків-Пас та встановлено наявність добової періодичності. Відповідно проведених досліджень з метою зменшення розмірності моделі прогнозування була прийнята структура математичної моделі, що має п'ять входів x_i , $i = \overline{1, 5}$, на які подаються попередні значення тимчасового ряду в момент часу $t, t-1, t-2, t-6, t-7$ відповідно, та один вихід y який представляє задачу визначення значення ряду в момент $t + 1$.

З метою пристосування моделі до нечіткості вхідної інформації застосовується теорія нечітких множин, яка передбачає представлення кількісних значень параметрів моделі у вигляді лінгвістичних змінних [2, 3, 9], які оцінюються нечіткими термами. Внаслідок цього втрачається точність вхідних даних, проте виникає можливість пристосування до невизначених умов, що присутні в задачі прогнозування пасажиропотоків. Виходячи з цього, формування моделі прогнозування припускає завдання вхідних змінних x_i ($i = \overline{1, d}$) у вигляді лінгвістичних змінних N_i , що задані в універсальній множині X_i . Для оцінки лінгвістичних змінних використовуються якісні терми T_i^p , $p = \overline{1, l}$ із

наступної терм-множини T_i , $T_i^p \in T_i$, де кожен терм T_i^p описується нечіткою множиною $\tilde{C}_i^p = \{ \mu_{\tilde{C}_i^p}(x_i) / x_i \}$, $x_i \in X_i$, де $\mu_{\tilde{C}_i^p} : X_i \rightarrow [0, 1]$ - функція приналежності, що приймає значення в інтервалі $[0, 1]$ та $x_i \in X_i$, $\tilde{C}_i^p \subset X_i$. Представлення функції приналежності $\mu_{\tilde{C}_i^p}$ базується на дотриманні умови нормальності:

$$\sup_{x_i \in X_i} \mu_{\tilde{C}_i^p}(x_i) = 1 \quad (4)$$

Розглянемо ряд додаткових умов, яким повинна відповідати терм-множина T_i обраної лінгвістичної змінної N_i . Припустимо, що C_i носій нечіткої множини \tilde{C}_i^p та $X \subseteq R_1$, де R_1 - дійсна вісь. Упорядкуємо множину T в відповідності [2] за виразом:

$$(\forall T_i \in T)(\forall T_j \in T)(i > j \leftrightarrow (\exists x \in C_i)(\forall y \in C_j)(x > y)). \quad (5)$$

Відповідно до розглянутих вимог та обмежень, представимо входи моделі як лінгвістичні змінні, значення яких визначаються на єдиній шкалі з п'яти термів для всіх x_i , $i = \overline{1, 5}$: Н - низький, НС - нижчий за середнє, С - середній, ВС - вищий за середнє, В - високий [9].

На основі неперервності множини базових значень вхідних змінних використано функціональний метод визначення форми представлення нечітких термів [3, 9]. Для відображення функцій приналежності (ФП) термів вхідних змінних у функціональній формі була обрана крива Гауса, що має наступний аналітичний запис:

$$\mu_{\tilde{C}_i^p} = \exp[(x_i - b_i^p / \sigma_i^p)], \quad (6)$$

де вектор x_i - елемент універсальної множини; параметри b_i^p та σ_i^p - параметри настройки, що дозволяють змінювати положення і структуру нечітких множин: b_i^p - координата максимуму функції; σ_i^p - коефіцієнт концентрації функції [9].

Реалізація моделювання на основі нечітких баз знань здійснюється через нечіткий логічний висновок по алгоритму Такагі-Сугено [3, 9]. Структура взаємозв'язку між вхідними і вихідною змінними в такому алгоритмі описується правилами r^k , $k = \overline{1, N}$, які нечіткі тільки в частині умови (антедент), тоді як висновок являє чітку лінійну функцію від входів. Множина правил представляє собою нечітку базу правил R , в якій нечітке правило r^k для змінних x_i можна представити як логічну комбінацію в наступному вигляді:

$$r^k : \text{ЯКЩО } x_1 \in F_1^k(x_1) \text{ та...та } x_n \in F_n^k(x_n), \text{ ТО } y^k = p_0^k + p_1^k x_1 + \dots + p_n^k x_n, [w^k] \quad (7)$$

де F_i^k - нечіткий терм з ФП $\mu_{\tilde{C}_i^k}(x_i)$, ($k = \overline{1, N}$, $i = \overline{1, n}$); y^k - вихід кожного правила, що представлений як поліном першого порядку з коефіцієнтами p_1^k, \dots, p_n^k та вільною складовою p_0^k ; $w^k \in [0, 1]$ - вага правила, що характеризує ступінь впевненості у кожному k -му правилі бази знань; "та" - операція логічного зв'язування.

Для всіх вхідних змінних x_i виконується операція введення нечіткості на основі визначення фактичних

значень функцій приналежності $\mu_{C_i^k}^k(x_i)$ по кожному з лінгвістичних термів F_i^k для представлення ступеня істинності передумови $x_n \in F_n^k(x_n)$ кожного правила.

Таким чином, для вхідного вектору $x = [x_{1,k}, \dots, x_{n,k}]^T$ виходом $y(\bar{x})$ нечіткої системи є зважене середнє від розрахованих значень вихідних змінних кожного правила \tilde{y}^k :

$$y(\bar{x}) = \frac{\sum_{k=1}^N \tau^k \tilde{y}^k}{\sum_{k=1}^N \tau^k} \tag{8}$$

де \tilde{y}^k розраховується за виразом $\tilde{y}^k = p_0^k + p_1^k x_1 + \dots + p_n^k x_n$, в якому значення вхідних змінних x_i підставляються до етапу введення нечіткості, а вага τ^k представляє узагальнену величину істинності при застосуванні до входу правила r^k , $k = \overline{1, N}$ і обчислюється на основі використання оператора розрахунку алгебраїчного добутку за виразом:

$$\tau^k = w^k \prod_{i=1}^n \mu_{F_i^k}(x_i) \tag{9}$$

Позначимо через $\zeta^k = \tau^k / \sum_{k=1}^N \tau^k$ твірний ступінь виконання висновку k -го правила для вхідного вектора \bar{x} . Тоді (9) можна переписати у вигляді:

$$y(\bar{x}) = \sum_{k=1}^N \zeta^k \tilde{y}^k = \sum_{k=1}^N \zeta^k (p_0^k + p_1^k x_1 + \dots + p_n^k x_n) \tag{10}$$

Співвідношення (9)-(11) [9] визначають нечітку модель прогнозування пасажиропотоків. Побудови її дає можливість представлення нечіткого логічного висновку моделі у вигляді нейроподібної структури, що називається адаптивною нейро-нечіткою системою висновків (adaptive network based Fuzzy Inference System, ANFIS) [4](Рис.2).

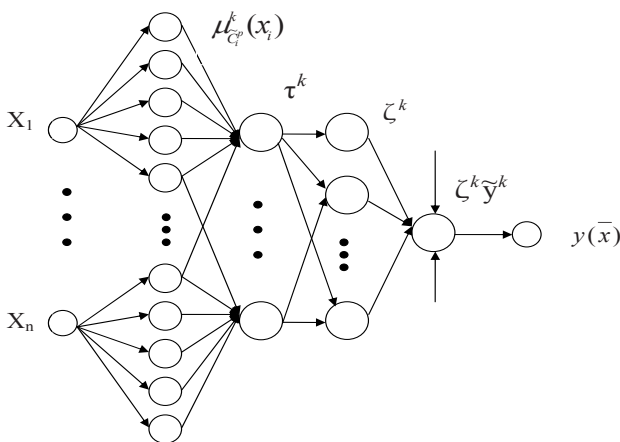


Рисунок 2. Нейроструктурна схема гібридної системи ANFIS

Для генерування та уточнення нечітких правил нейро-нечіткої мережі на основі спеціальних навчальних методик доцільним є використання сучасних моделей генетичних алгоритмів.

З метою підвищення швидкості та точності знаходження відповідних рішень поставленої задачі доцільним є використання генетичного алгоритму з дійсним кодуванням (англ.: Real-coded Genetic Algorithm, RGA)

[9,5], в якому рішення поставленої задачі передбачає представлення набору параметрів моделі у вигляді хромосоми фіксованої довжини C_h , що складається з трьох частин:

$$C_h = (C^1 C^2 C^3) \tag{11}$$

де $h = \overline{1, K}$ – номер хромосоми C .

Перша частина хромосоми представлена генами, які визначаються параметрами ФП:

$$C^1 = (b_{11}^p, c_{11}^p, \dots, b_{n_1}^p, c_{n_1}^p) \tag{12}$$

де l_i – кількість термів-оцінок вхідної змінної

$$x_i, l_1 + l_2 + \dots + l_n = q, i = \overline{1, n};$$

q – загальна кількість термів. На можливі значення параметрів ФП накладено обмеження відповідно до допустимого відхилення $\pm \Delta b$, $\pm \Delta c$, що дозволяє знайти оптимальну форму і положення функцій приналежності на граничному інтервалі враховуючи недопустимість зміни позиції в умовах кожного правила.

Друга частина хромосоми реалізує гени, що визначають коефіцієнти висновків відповідних правил:

$$C^2 = (p_0^k, p_1^k, \dots, p_n^k), k = \overline{1, N} \tag{13}$$

Остання частина хромосоми реалізує гени, що визначають структуру анцеденту кожного правила:

$$C^3 = (1, 2, \dots, q), q = \overline{1, L} \tag{14}$$

Умова зміни положення елементів терм-множини в анцеденті правил виконується прямим способом (номер позиції в умові правила завжди відповідає номеру ФП), що дозволяє використовувати кожний терм в різних можливих позиціях анцедента правил.

Для оптимізації структури заданою мірою якості є функція пристосованості, що забезпечує мінімальне відхилення між теоретичними і експериментальними результатами моделювання на навчальній вибірці (MSE):

$$FF(Ch_j) = \frac{1}{M_{r=1, M}} \sum (y_r - y_r^f)^2 \rightarrow \min \tag{15}$$

де y_r - вихідний вектор навчальної вибірки, y_r^f - результат виводу нечіткої моделі прогнозування при значенні

входів із r -ї строки вибірки $\langle \bar{X}_r, \bar{y}_r \rangle$.

Для виконання операції схрещування та мутації запропоновано використовувати кросовер BLX- α з параметром $\alpha = 0,5$ [7,9] та нерівномірну мутацію Михалевича [8,9].

Алгоритм формування моделі прогнозування розділений на дві фази: фаза навчання на основі самоорганізації, що дозволяє за допомогою RGA визначити структуру нейроподібної моделі – знайшовши рішення достатньо близьке до оптимального та фаза навчання з вчителем, що потім розглядає знайдене рішення як відповідний крок для проведення “тонкої” настройки змінних параметрів моделі, за класичними методами навчання нейро-нечіткої мережі [3,4].

Для отримання комплексної моделі прогнозування, що дозволяє враховувати послідовність інтервалів доби

у періоді прогнозування відповідно вибраному типу доби та сезону використано орієнтований граф $G = (X, F)$. Вершини графа це запропоновані нейро-нечіткі моделі прогнозування, що навчені на вибірці даних окремого інтервалу доби відповідного типу дня тижня та сезону року, тобто: $F_{1,1}$ $F_{1,2}$ - моделі, що відповідають нічному інтервалу, $F_{1,1}$ - відповідно від 0-00 до 3-00, $F_{1,2}$ - від 3-00 до 6-00; $F_{2,1}$ $F_{2,2}$ - моделі, що відповідають ранковому інтервалу за відповідні періоди з 6-00 до 9-00, та з 9-00 до 12-00; F_3 - модель дня в період від 12-00 до 15-00; $F_{4,1}$ $F_{4,2}$ - моделі вечірнього інтервалу доби відповідно за періоди з 15-00 до 18-00 та з 18-00 до 21-00; F_5 - модель переднічного періоду з 21-00 до 00-00.

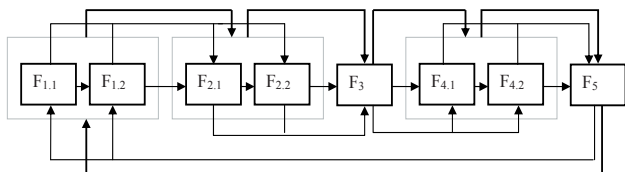


Рисунок 3. Орієнтований граф $G = (X, F)$ формування послідовності моделей відповідних періодів прогнозування

В режимі запиту до користувача алгоритм визначає початковий період доби прогнозування та пропонує кроки формування структури загальної моделі прогнозування відповідно до графу $G = (X, F)$ (ставить мету до 5 кроків).

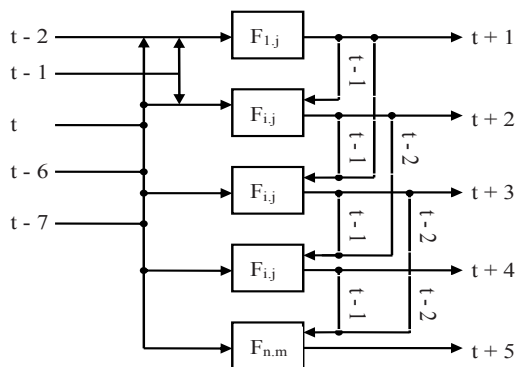


Рисунок 4. Загальна структура моделі прогнозування

6. Отримані результати

Запропонований і адаптований до умов нерівномірності приміських пасажиропотоків алгоритм прогнозування стійкий до нечіткості вхідної інформації та забезпечує похибку не більше 4–5%. В подальшому є можливість підвищення точності прогнозу завдяки донавчання в процесі роботи.

7. Висновок

У результаті комплексного застосування запропонованої методики прогнозування в системі підтримки

прийняття рішень оперативного персоналу стосовно раціоналізації та ефективності використання рухомого складу стає можливим отримання своєчасних і більш точних рішень. Це дозволить оперативно реагувати і формувати управляючі дії на перевізний процес, розробити найбільш ефективну та раціональну технологію курсування рейкових автобусів, удосконалити графік обороту приміських поїздів на основі принципів не маятникового руху та організувати загальною роботу приміського комплексу на принципах пасажирської логістики.

Література

1. Zadeh L. A., "Fuzzy logic and soft computing: Issues, contentions and perspectives," in Proc. IIZUKA'94: 3rd Int. Conf. Fuzzy Logic, Neural Nets and Soft Computing, Iizuka, Japan, 1994, pp. 1–2.
2. Заде Л. "Понятие лингвистической переменной и его применение к принятию приближенных решений." -М.: Мир, 1976. -162 с.
3. Рутковская Д., Пилинский М., Рутковский Л. "Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польск. И.Д.Рудинского.-М.:Горячая линия – Телеком, 2004. –452 с.: ил.
4. Бутко Т.В., Прохорченко А.В. "Застосування нейро-нечіткого моделювання в системах підтримки прийняття рішень для оперативного корегування поїздоутворенням пасажирських составів"//Восточно-европейский журнал передовых технологий. –Харьков,2006. – Вип.1/2(-19). – С.32-36.
5. Wright A."Genetic algorithms for real parameter optimization"// Foundations of Genetic Algorithms, V. 1. – 1991. – P. 205-218.
6. O. Cord'ón, F. Herrera, A two-stage evolutionary process for designing TSK fuzzy rule-based systems, IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics 29:6 (1999) 703–715.
7. Eshelman LJ, Schaffer JD Real-coded genetic algorithms and interval-schemata. In: Foundations of genetic algorithms 2, Whitley LD (ed) Morgan Kaufmann Publishers, San Mateo, CA, pp 187–202 (1993).
8. Michalewicz Z. "Genetic Algorithms, Numerical Optimization and Constraints, Proceedings of the 6th International Conference on Genetic Algorithms", Pittsburgh, July 15-19, 1995. - P. 151-158.
9. Бутко Т.В., Прохорченко А.В. Удосконалення системи оперативного прогнозування пасажирських потоків на основі використання інтелектуальних технологій // Зб.наук.праць. – Харків:УкрДАЗТ, 2007. – Вип.85. – С.161-171.
10. Концепція Державної програми реформування залізничного транспорту від 27 грудня 2006 р. N 651-р. <http://www.uz.gov.ua/>