

УДК 656.2

Горобченко О. М.^{1*}, Неведров О. В.², Незліна О. А.³, Ткаченко В. А.⁴

¹ Професор кафедри електромеханіки та рухомого складу залізниць, Державний університет інфраструктури та технологій, вул. Кирилівська, 9, м. Київ, 04071, Україна

ORCID: <https://orcid.org/0000-0002-9868-3852>

² Аспірант кафедри електромеханіки та рухомого складу залізниць, Державний університет інфраструктури та технологій, вул. Кирилівська, 9, м. Київ, 04071, Україна

ORCID: <https://orcid.org/0000-0001-9347-0973>

³ Доцент кафедри електромеханіки та рухомого складу залізниць, Державний університет інфраструктури та технологій, вул. Кирилівська, 9, м. Київ, 04071, Україна

ORCID: <https://orcid.org/0000-0003-1613-4057>

⁴ Доцент кафедри транспортного права та логістики Національного транспортного університету, вулиця Михайла Омеляновича-Павленка, 1, м. Київ, 02000, Україна

ORCID: <https://orcid.org/0000-0001-5147-0772>

* Автор, відповідальний за листування: gorobchenko.a.n@gmail.com

РОЗРОБКА МЕТОДУ КЛАСТЕРИЗАЦІЇ ПОЇЗНИХ СИТУАЦІЙ

Впровадження інтелектуальних систем керування локомотивом потребує більш якісних підходів до оцінки та контролю поточної поїзної ситуації, ніж ті, які використовуються на сучасному тяговому рухомому складі. Автоматичне виявлення складних нештатних ситуацій в теперішній час не передбачено. Наприклад визначення неефективності дії гальм, перевищення швидкості, наявності переешкод або людей на колії, погіршення тягових властивостей рухомого складу та інше покладається виключно на машиніста локомотива. Враховуючи важливий вплив перелічених факторів на безпеку руху, пропонується включити в функції автоматизованих та інтелектуальних систем керування рухом розпізнавання нештатної ситуації і оповіщення про її виникнення. При керуванні поїздом всі об'єкти класифікації (поїзні ситуації) розбиті на кінцеве число класів. Для кожного класу відомо і вивчено кінцеве число об'єктів – прецедентів. Завдання розпізнавання образів полягає в тому, щоб віднести нову розпізнавану ситуацію до якого-небудь класу. Класифікатором або вирішальним правилом є правило віднесення образу поїзної ситуації до одного з класів на підставі його вектору ознак. Розроблено порядок класифікації поїзних ситуацій, який дозволяє виділяти кластери як завгодно складної форми за умови, що різні частини таких кластерів з'єднані ланцюжками близьких один до одного елементів. Мірою відмінності служить квадрат евклідової відстані

Ключові слова: поїзд, розпізнавання ситуацій, безпека руху, кластеризація, класифікатор.

Вступ. Впровадження інтелектуальних систем керування локомотивом потребує більш якісних підходів до оцінки та контролю поточної поїзної ситуації, ніж ті, які використовуються на сучасному тяговому рухомому складі. В теперішній час основу інформації про поточний стан поїзду системи керування отримують з різноманітних датчиків у вигляді кількісних показників. Однак машиніст локомотива, при керуванні поїздом керується в основному якісними показниками, що найбільш природньо для роботи людського мозку та поведінки людини [1, 2].

Розмаїтість експлуатаційних умов, під якими звичайно розуміють вагу і довжину поїзда, його опір руху, порядок пропуску поїзда по перегонах, погодні умови, установлені попередженнями обмеження швидкості руху та ін., ставить перед машиністом у кожній поїзді завдання вибору і реалізації раціонального режиму ведення поїзда, що відповідає саме даним умовам [3, 4].

Аналіз останніх досліджень. Існує декілька підходів до формалізації умов, в яких працює людина-оператор. В роботі [5] виконана оцінка надійності людини, представлений новий підхід до оцінки показників людської діяльності щодо залежностей серед відповідних PSF, що називається індексом ефективності роботи залізниць (HuPeROI). Результати показують, що HuPeROI можна ефективно використовувати для оцінки ефективності роботи операторів як функції якості відповідних R-PSF. Однак в цих роботах не в повній мірі проведено оцінку інформаційної складовою факторів, що впливають на безпеку руху. Це не дає розуміння того, який інформаційний вклад в опис та формалізацію небезпечної ситуації надає кожен з факторів. Досвід показує, що виникнення окремих транспортних подій стало можливим тому, що локомотивна бригада невчасно виявила та відреагувала на розвиток НС [6].

Велику низку робіт при описанні поїзних ситуацій спрямовано на дослідження найбільш енергоефективних способів ведення поїзду [7–10]. Тут процес керування поїздом розглядається як вибір оптимальних траєкторій руху з одночасним урахуванням максимальної кількості різних факторів. Причому в різних роботах наведено різний перелік впливів на ефективність керування поїздом. При впровадженні інтелектуальних систем керування рухом потрібно мати повний перелік факторів, що впливають на енергоефективність перевізного процесу [11]. Крім того, для більш якісного їх використання необхідно провести оцінку їх інформативності та впливу на якість керування поїздом [12]. Так в роботі [13] розглянуто системні властивості інформаційних потоків, представлено модель, що дозволяє реєструвати первинну інформацію, яка відображає синхронно тимчасові процеси донормальні, аномальні, постнормальні, а також способи організації синхронно записаних первинних даних та способи їх формування в єдиний інформаційний простір. Але тут не розглянуто інформаційних потоків, що надходять безпосередньо до локомотивної бригади та дозволяють найбільш якісно визначити поточну поїзну ситуацію.

Мета і завдання дослідження. В процесі керування рухом поїзду перед машиністами та автоматизованими системами керування постає важливе завдання – забезпечення безпеки руху та енергоефективності процесу перевезень. Для цього необхідно виявляти нештатні ситуації на етапі їх виникнення та не допускати їх розвитку в аварію або катастрофу. Розпізнавання нештатної ситуації дуже відповідальний процес.

Автоматичне виявлення складних нештатних ситуацій в теперішній час не передбачено. Наприклад визначення неефективності дії гальм, перевищення швидкості, наявності перешкод або людей на колії, погіршення тягових властивостей рухомого складу та інше покладається виключно на машиніста локомотива. Враховуючи важливий вплив перелічених факторів на безпеку руху, пропонується включити в функції автоматизованих та інтелектуальних систем керування рухом розпізнавання нештатної ситуації і оповіщення про її виникнення.

Методи кластерного аналізу при визначення основних підходів до розпізнавання поїзних ситуацій. Завдання кластерного аналізу полягає в тому, щоб на підставі даних, що містяться у множині Ω , розбити множину об'єктів ω на m [14] (m – ціле число, що відповідає кількості простих поїзних ситуацій) кластерів (підмножин)

C_1, C_2, \dots, C_m , так, щоб кожен об'єкт ω_j належав одній і тільки одній підмножині розбиття і щоб об'єкти, що належать одному і тому ж кластеру, були східними, в той час як об'єкти, що належать різним кластерам були різнорідними.

Рішенням задачі кластерного аналізу є розбиття, що задовольняє деякому критерію оптимальності. Цей критерій може представляти собою деякий функціонал, що виражає рівні бажаності різних розбивок і угруповань, який називають цільовою функцією [15]. Наприклад, як цільова функція може бути взята внутрішньогрупова сума квадратів відхилень

$$W = \sum_{j=1}^n (x_j - \bar{x})^2 = \sum_{j=1}^n x_j^2 - \frac{1}{n} \left(\sum_{j=1}^n x_j \right)^2, \quad (1)$$

де x_j – являє собою вимірювання j -го об'єкта.

Для вирішення задачі кластерного аналізу необхідно визначити поняття подібності і різнорідності.

Об'єкти i -й та j -й потрапляли б в один кластер, коли відстань (віддаленість) між точками X_i і X_j була б досить маленькою і потрапляли б в різні кластери, коли ця відстань була б досить великою. Таким чином, попадання в один чи різні кластери об'єктів визначається поняттям відстані між X_i і X_j з E_p , де E_p – p -мірне евклідовий простір. Невід'ємна функція $d(X_i, X_j)$ називається функцією відстані (метрикою), якщо:

а) $d(X_i, X_j) \geq 0$, для всіх X_i і X_j з E_p

б) $d(X_i, X_j) = 0$, тоді і тільки тоді, коли $X_i = X_j$

в) $d(X_i, X_j) = d(X_j, X_i)$

г) $d(X_i, X_j) \leq d(X_i, X_k) + d(X_k, X_j)$, де X_i, X_j і X_k – будь-які три вектори з E_p .

Значення $d(X_i, X_j)$ для X_i і X_j називається відстанню між X_i і X_j і еквівалентно відстані між G_i і G_j відповідно обраним характеристикам $(f_1, f_2, f_3, \dots, f_p)$.

Найбільш часто вживаються наступні функції відстаней:

1. Евклідова відстань
$$d_2(X_i, X_j) = \left[\sum_{k=1}^p (x_{ki} - x_{kj})^2 \right]^{\frac{1}{2}}.$$

2. l_1 – норма
$$d_1(X_i, X_j) = \left[\sum_{k=1}^p |x_{ki} - x_{kj}| \right].$$

3. Сюзремум-норма
$$d_\infty(X_i, X_j) = \sup \left\{ |x_{ki} - x_{kj}| \right\}$$

 $k = 1, 2, \dots, p$

4. l_p – норма
$$d_p(X_i, X_j) = \left[\sum_{k=1}^p |x_{ki} - x_{kj}|^p \right]^{\frac{1}{p}}.$$

Евклідова метрика є найбільш популярною. Метрика l_1 найбільш легка для обчислень. Сюзремум-норма рахується легко і включає в себе процедуру упорядкування, а l_p -норма охоплює функції відстаней 1, 2, 3

Нехай n вимірювань X_1, X_2, \dots, X_n подані у вигляді матриці даних розміром $p \times n$:

$$x = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{p1} & x_{p2} & \dots & x_{pn} \end{pmatrix} = (X_1, X_2, \dots, X_n). \quad (2)$$

Тоді відстань між парами векторів $d(X_i, X_j)$ може бути представлена у вигляді симетричної матриці відстаней:

$$D = \begin{pmatrix} 0 & d_{12} & \dots & d_{1n} \\ d_{21} & 0 & \dots & d_{2n} \\ \dots & \dots & \dots & \dots \\ d_{n1} & d_{n2} & \dots & 0 \end{pmatrix}. \quad (3)$$

Поняттям, протилежним відстані, є поняття подібності між об'єктами ω_i та ω_j . Не негативна функція $S(X_i, X_j) = S_{ij}$ є мірою подібності, якщо:

- 1) $0 \leq S(X_i, X_j) < 1$ для $X_i \neq X_j$
- 2) $S(X_i, X_j) = 1$
- 3) $S(X_i, X_j) = S(X_j, X_i)$

Пари значень мір подібності можна поєднати в матрицю подібності:

$$S = \begin{pmatrix} 1 & s_{12} & \dots & s_{1n} \\ s_{21} & 1 & \dots & s_{2n} \\ \dots & \dots & \dots & \dots \\ s_{n1} & s_{n2} & \dots & 1 \end{pmatrix}. \quad (4)$$

Величину S_{ij} називають коефіцієнтом подібності.

Розробка алгоритму кластеризації поїзних ситуацій. Розглянемо $I = (I_1, I_2, \dots, I_n)$ як множину кластерів $\{I_1\}, \{I_2\}, \dots, \{I_n\}$. Виберемо два з них, наприклад, I_i та I_j , які в деякому сенсі більш близькі один до одного й об'єднаємо їх в один кластер. Нова множина кластерів, що складається вже з $n-1$ кластерів, буде $\{I_1\}, \{I_2\}, \dots, \{I_i, I_j\}, \dots, \{I_n\}$.

Повторюючи процес, отримаємо послідовні множини кластерів, що складаються з $(n-2)$, $(n-3)$, $(n-4)$ і т. д. кластерів. В кінці процедури можна отримати кластер, що складається з n об'єктів і збігається з початковою множиною $I = (I_1, I_2, \dots, I_n)$.

Як міру відстані візьмемо квадрат евклідової метрики d_{ij}^2 і обчислимо матрицю $D = \{d_{ij}^2\}$, де d_{ij}^2 - квадрат відстані між I_i та I_j (табл. 1).

Таблиця 1. Попередня матриця відстаней

	I_1	I_2	I_3	I_n
I_1	0	d_{12}^2	d_{13}^2	d_{1n}^2
I_2		0	d_{23}^2	d_{2n}^2
I_3			0	d_{3n}^2
....			
I_n					0

Нехай відстань між I_i та I_j буде мінімальною:

$$d_{ij}^2 = \min \{d_{ij}^2, i \neq j\}. \quad (5)$$

Утворимо за допомогою I_i і I_j новий кластер $\{I_i, I_j\}$. Побудуємо нову $((n-1), (n-1))$ матрицю відстаней (табл. 2)

Таблиця 2. Матриця відстаней

	{I _i , I _j }	I ₁	I ₂	I ₃	I _n
{I _i ; I _j }	0	d _{ij} ² ₁	d _{ij} ² ₂	d _{ij} ² ₃	d _{ij} ² _n
I ₁		0	d ₁₂ ²	d ₁₃	d _{1n} ²
I ₂			0	d _{ij} ² ₁	d _{2n}
I ₃				0	d _{3n}
I _n						0

У ній (n-2) рядки для останньої матриці взяті з попередньої, а перший рядок обчислено заново. Обчислення можуть бути зведені до мінімуму, якщо вдасться виразити $d_{ij}^2, k = 1, 2, \dots, n; (k \neq i \neq j)$ через елементи первісної матриці.

Початково визначено відстань лише між одноелементними кластерами, але треба визначати відстані між кластерами, що містить більш ніж один елемент. Це можна зробити різними способами, в залежності від обраного способу ми отримаємо алгоритми кластерного аналізу з різними властивостями. Можна, наприклад, покласти відстань між кластером $i + j$ та деяким іншим кластером k , рівним середньому арифметичному з відстаней між кластерами i і k і кластерами j і k :

$$d_{i+j,k} = \frac{1}{2} (d_{ik} + d_{jk}). \quad (6)$$

Але можна також визначити $d_{i+j,k}$ як мінімальне з цих двох відстаней:

$$d_{i+j,k} = \min (d_{ik} + d_{jk}). \quad (7)$$

Таким чином, описаний перший крок роботи агломеративного ієрархічного алгоритму. Наступні кроки аналогічні.

Досить широкий клас алгоритмів може бути отриманий, якщо для перерахунку відстаней використовувати таку загальну формулу:

$$d_{i+j,k} = A(w) \min(d_{ik} d_{jk}) + B(w) \max(d_{ik} d_{jk}), \quad (8)$$

де

$$A(w) = \frac{wn_i}{wn_i + n_j}, \text{ якщо } d_{ik} \leq d_{jk}$$

$$A(w) = \frac{wn_j}{n_j + wn_i}, \text{ якщо } d_{ik} > d_{jk}$$

$$B(w) = \frac{n_i}{wn_i + n_j}, \text{ якщо } d_{ik} \leq d_{jk}$$

$$B(w) = \frac{n_j}{wn_j + n_i}, \text{ якщо } d_{ik} > d_{jk}$$

де n_i і n_j – кількість елементів у кластерах i та j , а w – вільний параметр, вибір якого визначає конкретний алгоритм. Наприклад, при $w = 1$ ми одержуємо так званий алгоритм «середнього зв'язку», для якого формула перерахунку відстаней приймає вигляд:

$$d_{i+j,k} = \frac{n_i}{n_i + n_j} d_{ik} + \frac{n_j}{n_i + n_j} d_{jk}. \quad (9)$$

В даному випадку відстань між двома кластерами на кожному кроці роботи алгоритму виявляється рівною середньому арифметичному з відстаней між усіма такими парами елементів, що один елемент пари належить до одного кластера, інший – до іншого.

Наочний сенс параметра w стає зрозумілим, якщо припустити $w \rightarrow \infty$. Формула перерахунку відстаней приймає вигляд:

$$d_{i+j,k} = \min(d_{ik}, d_{jk}). \quad (10)$$

Це буде так званий алгоритм «найближчого сусіда», який дозволяє виділяти кластери як завгодно складної форми за умови, що різні частини таких кластерів з'єднані ланцюжками близьких один до одного елементів. В даному випадку відстань між двома кластерами на кожному кроці роботи алгоритму виявляється рівною відстані між двома самими близькими елементами, що належать до цих двох кластерів.

Досить часто припускають, що початкові відстані (відмінності) між елементами, що групуються, задані. У деяких задачах це дійсно так. Однак, задаються тільки об'єкти та їх характеристики і матрицю відстаней будують виходячи з цих даних. Залежно від того, обчислюються відстані між об'єктами або між характеристиками об'єктів, використовуються різні способи.

У разі кластер аналізу об'єктів найбільш часто мірою відмінності служить квадрат евклідової відстані

$$d_{ij}^2 = \sum_{h=1}^m (x_{ih} - x_{jh})^2, \quad (11)$$

де x_{ih}, x_{jh} – значення h -ї ознаки для i -го та j -го об'єктів;
 m – число характеристик.

Якщо ознакам приписується різну вагу, то ці ваги можна врахувати при обчисленні відстані

$$\delta_{ij}^2 = \sum_{h=1}^m w_h (x_{ih} - x_{jh})^2. \quad (12)$$

Іноді як міра відмінності використовується відстань, що обчислюється за формулою

$$\Delta_{ij} = \sum_{h=1}^m |x_{ih} - x_{jh}|, \quad (13)$$

які називають: "хеммінговим", "манхеттенським" або "сіті-блок" відстанню.

Природною мірою подібності характеристик об'єктів у багатьох задачах є коефіцієнт кореляції між ними

$$r_{ij} = \frac{\sum_{h=1}^N (x_{hi} - m_i)(x_{hj} - m_j)}{\delta_i \delta_j}, \quad (14)$$

де $m_i, m_j, \delta_i, \delta_j$ - відповідно середні і середньоквадратичні відхилення для характеристик i та j . Мірою відмінності між характеристиками може служити величина $1 - r$. У деяких завданнях знак коефіцієнта кореляції неважливий і залежить лише від

вибору одиниці вимірювання. В цьому випадку в якості міри відмінності між характеристиками використовується $|1 - r_{ij}|$.

Метод визначення поїзної ситуації на основі кластерного аналізу. Отримання вектору характеристик поточної поїзної ситуації у вигляді:

$$C_{nom} = \begin{pmatrix} f_{1пот} \\ f_{2пот} \\ \dots \\ f_{iпот} \\ \dots \\ f_{nпот} \end{pmatrix}, \quad (15)$$

де $f_{iпот}$ – значення i -го фактору, що характеризує поїзну ситуацію.

Нормалізація отриманих величин факторів згідно з виразом

$$f'_{iпот} = \frac{f_{iпот} - f_{imin}}{f_{imax} - f_{imin}}, \quad (16)$$

де f_{imax} (f_{imin}) – максимально (мінімально) можливе значення i -го фактора в експлуатації.

Визначити евклідову відстань між першою ситуацією та поточною C_{nom}

$$D(c_1) = \sqrt{\sum_{h=1}^n (f(c_1)_h - f_{hпот})^2}, \quad (17)$$

де $f(c_1)_h$ – значення h -го фактора, що характеризує ситуацію c_1 .

Визначити всі відстані між C_{nom} та $D(c_i)$, при $i \in [1;12]$ згідно (17).

Перерахувати значення відстаней $D(c_i)$ у відповідності до вагових коефіцієнтів

$$D'(c_i) = D(c_i) \cdot \gamma_i, \quad (18)$$

де $D'(c_i)$ – приведені значення відстані між поточною ситуацією та i -м кластером (елементарною поїзною ситуацією);

γ_i – ваговий коефіцієнт, що визначає відносну важливість кожної i -ї ситуації.

Таким чином отримана множина D , що складається з відстаней від поточних факторів, що формують поточну ситуацію керування локомотивом, від базових поїзних ситуацій.

Елементи D ранжуються від мінімального до максимального значення $D = \{D_{min}, \dots, D_i, \dots, D_{max}\}$.

Перший елемент визначає основну ситуацію, другий і декілька наступних елементів визначають додаткові впливи на основну поїзну ситуацію. Знання основної та впливаючих поїзних ситуацій необхідно для вироблення керуючого рішення, що враховує всі наслідки для тієї або іншої ситуації.

Висновки. Розглядаючи перелік поїзних ситуацій можна сказати, що під час керування поїздом машиніст часто знаходиться в складних обставинах, що обумовлені одночасним знаходженням у декількох поїзних ситуаціях. І якщо для людини постійний моніторинг та визначення обстановки це природній процес, то для інтелектуальних та автоматизованих систем керування необхідно розробити формалізоване описання цієї процедури.

В разі керування поїздом всі об'єкти класифікації (поїзні ситуації) розбиті на кінцеве число класів. Для кожного класу відомо і вивчено кінцеве число об'єктів – прецедентів. Завдання розпізнавання образів полягає в тому, щоб віднести нову розпізнавану ситуацію до будь-якого класу. Класифікатором або вирішальним правилом є правило віднесення образу поїзної ситуації до одного з класів на підставі його вектору ознак.

Розроблено порядок класифікації поїзних ситуацій, який дозволяє виділяти кластери як заведено складної форми за умови, що різні частини таких кластерів з'єднані ланцюжками близьких один до одного елементів. Мірою відмінності служить квадрат евклідової відстані.

ЛІТЕРАТУРА

1. Kuric I. et al. Research of vehicle control informative functioning capacity //IOP Conference Series: Materials Science and Engineering. – IOP Publishing, 2020. Vol. 776. №. 1. P. 012036.
2. Butko, T., Babanin, A., Gorobchenko, A. Rationale for the type of the membership function of fuzzy parameters of locomotive intelligent control systems //Eastern-European Journal of Enterprise Technologies. 2015. №. 3 (73). P. 4-8.
3. Горобченко О. М. Моделювання виникнення нештатної ситуації в ергатичній системі локомотивна бригада-поїзд //Збірник наукових праць Донецького інституту залізничного транспорту. 2014. №. 38. С. 144-147.
4. Горобченко О. М. Розробка методу оцінки факторів, що впливають на дії локомотивних бригад в нештатних ситуаціях //Сборник научных трудов Донецкого института железнодорожного транспорта. 2010. №. 24. С. 113-121
5. Kyriakidis M., Majumdar A., Ochieng W. Y. The human performance railway operational index—a novel approach to assess human performance for railway operations //Reliability engineering & system safety. 2018. Val. 170. P. 226-243.
6. Горобченко О., Слободянюк М., Неведров О. Формалізація поїзних ситуацій при керуванні локомотивом на основі методів нечіткої логіки. Транспортні системи і технології, № 34, С. 65-70, doi:10.32703/2617-9040-2019-34-1-5.
7. Gorobchenko O. et al. Study of the influence of electric transmission parameters on the efficiency of freight rolling stock of direct current. Eastern-European Journal of Enterprise Technologies, 1(3-91), 60-67. doi:10.15587/1729-4061.2018.121713.
8. Zhou Y. et al. Integrated optimization on train control and timetable to minimize net energy consumption of metro lines //Journal of Advanced Transportation. 2018. Vol. 2018.
9. Jia J. et al. Designing train-speed trajectory with energy efficiency and service quality. Engineering Optimization. 2018. Val. 50. №. 5. P. 797-818.
10. Фалендиш А., Гатченко В., Возненко С., Клецька, О., Барибін, М. Математичне моделювання основних параметрів в тягових розрахунках. Транспортні системи і технології, (35), С. 102-112. <https://doi.org/10.32703/2617-9040-2020-35-11>
11. Gorobchenko O. et al. Intelligent locomotive decision support system structure development and operation quality assessment //2018 IEEE 3rd International Conference on Intelligent Energy and Power Systems (IEPS). IEEE, 2018. P. 239-243.
12. Gorobchenko O., Nevedrov O. Development of the structure of an intelligent locomotive DSS and as-sessment of its effectiveness. Archives of Transport. 2020. Val. 56. №. 4. P. 47-58.
13. Holub, H., Kulbovskyi, I., Skok, P., Melnychenko, O., Kharuta, V., Bambura, O., Tretynychenko, Y. System model of information flows in networks of the electric supply system in transport infrastructure projects. Transport Means - Proceedings of the International Conference. 2020, 132-135.
14. Wierzchoń S. T., Kłopotek M. A. Modern algorithms of cluster analysis. Springer, 2018. 421 p.
15. Anderberg M. R. Cluster analysis for applications: probability and mathematical statistics: a series of monographs and textbooks. – Academic press, 2014. Vol. 19.

REFERENCES

1. Kuric, I., Gorobchenko, O., Litikova, O., Gritsuk, I., Mateichyk, V., Bulgakov, M., & Klackova, I. (2020). Research of vehicle control informative functioning capacity. *Paper presented at the IOP Conference Series: Materials Science and Engineering*, 776(1) doi:10.1088/1757-899X/776/1/012036.
2. Butko, T., Babanin, A. & Gorobchenko, A. (2015). Rationale for the type of the membership function of fuzzy parameters of locomotive intelligent control systems. *Eastern-European Journal of Enterprise Technologies*, 1(3), 4-8. doi:10.15587/1729-4061.2015.35996.
3. Горобченко, О. М. (2014). Моделювання виникнення нештатної ситуації в ергатичній системі локомотивна бригада-поїзд. *Збірник наукових праць Донецького інституту залізничного транспорту* (38). С. 144-147.
4. Горобченко, О. М. (2010). Розробка методу оцінки факторів, що впливають на дії локомотивних бригад в нештатних ситуаціях. *Збірник наукових праць Донецького інституту залізничного транспорту*, (24). С. 113-121.
5. Kyriakidis, M., Majumdar, A., & Ochieng, W. Y. (2018). The human performance railway operational index—a novel approach to assess human performance for railway operations. *Reliability engineering & system safety*, 170, 226-243.
6. Gorobchenko, O., Slobodianiuk, M., & Nevedrov, O. (2019). Formalization situations at train locomotive management on the basis of fuzzy logic. *Transport Systems and Technologies*, (34), 65-70. <https://doi.org/10.32703/2617-9040-2019-34-1-5>

7. Gorobchenko, O., Fomin, O., Fomin, V., & Kovalenko, V. (2018). Study of the influence of electric transmission parameters on the efficiency of freight rolling stock of direct current. *Eastern-European Journal of Enterprise Technologies*, 1(3-91), 60-67. doi:10.15587/1729-4061.2018.121713.
8. Zhou, Y., Bai, Y., Li, J., Mao, B., & Li, T. (2018). Integrated optimization on train control and timetable to minimize net energy consumption of metro lines. *Journal of Advanced Transportation*, 2018.
9. Jia, J., Yang, K., Yang, L., Gao, Y., & Li, S. (2018). Designing train-speed trajectory with energy efficiency and service quality. *Engineering Optimization*, 50(5), 797-818.
10. Falendish, A., Hatchenko, V., Voznenko, S., Kletska, O., & Varybin, M. (2020). Математичне моделювання основних параметрів в тягових розрахунках. *Транспортні системи і технології*, (35), 102-112. <https://doi.org/10.32703/2617-9040-2020-35-11>
11. Gorobchenko, O., Fomin, O., Gritsuk, I., Saravas, V., Grytsuk, Y., Bulgakov, M., . . . Zinchenko, D. (2018). Intelligent locomotive decision support system structure development and operation quality assessment. Paper presented at the 2018 IEEE 3rd International Conference on Intelligent Energy and Power Systems, IEPS 2018 - Proceedings, 2018-January 239-243. doi:10.1109/IEPS.2018.8559487.
12. Gorobchenko, O., Nevedrov, O., (2020). Development of the structure of an intelligent locomotive DSS and assessment of its effectiveness. *Archives of Transport*, 56(4), 47- 58. DOI: <https://doi.org/10.5604/01.3001.0014.5517>.
13. Holub, H., Kulbovskiy, I., Skok, P., Melnychenko, O., Kharuta, V., Bambura, O., & Tretynychenko, Y. (2020). System model of information flows in networks of the electric supply system in transport infrastructure projects. Paper presented at the Transport Means - Proceedings of the International Conference, 132-135.
14. Wierzchoń, S. T., & Kłopotek, M. (2018). *Modern algorithms of cluster analysis* (421 p.). Berlin, Germany: Springer.
15. Anderberg, M. R. (2014). *Cluster analysis for applications: probability and mathematical statistics: a series of monographs and textbooks* (Vol. 19). Academic press.

Gorobchenko O,^{1*}, Nevedrov O², Nezlina O³, Tkachenko V.⁴

¹ Professor of the Department of Electromechanics and Railway Rolling Stock, State University of Infrastructure and Technology, vul. Kyrylivska, 9, Kyiv, 04071, Ukraine

ORCID: <https://orcid.org/0000-0002-9868-3852>

² Postgraduate student of the Department of Electromechanics and Railway Rolling Stock, State University of Infrastructure and Technologies, vul. Kyrylivska, 9, Kyiv, 04071, Ukraine

ORCID: <https://orcid.org/0000-0001-9347-0973>

³ Associate Professor of Electromechanics and Railway Rolling Stock, State University of Infrastructure and Technology, st. Kyrylivska, 9, Kyiv, 04071, Ukraine

ORCID: <https://orcid.org/0000-0003-1613-4057>

⁴ Associate Professor of the Department of Transport Law and Logistics, National Transport University, M. Omelianovycha-Pavlenka Str., 1, Kyiv, 01010, Ukraine

* Author responsible for correspondence: gorobchenko.a.n@gmail.com

DEVELOPMENT OF THE METHOD OF CLUSTERIZATION OF TRAIN SITUATIONS

The introduction of intelligent locomotive control systems requires better approaches to assessing and monitoring the current train situation than those used in modern traction rolling stock. Automatic detection of complex abnormal situations is currently not provided. For example, determining the inefficiency of the brakes, speeding, the presence of obstacles or people on the track, the deterioration of the traction properties of rolling stock, etc. relies solely on the driver of the locomotive. Given the important impact of these factors on traffic safety, it is proposed to include in the functions of automated and intelligent traffic control systems recognition of abnormal situations and notification of its occurrence. When driving a train, all objects of classification (train situations) are divided into a finite number of classes. A finite number of precedent objects are known and studied for each class. The task of pattern recognition is to assign a new recognizable situation to a class. The classifier or decisive rule is the rule of assigning the image of a train situation to one of the classes on the basis of its vector of features. An order of classification of train situations has been developed, which allows to allocate clusters of any complex shape, provided that different parts of such clusters are connected by chains of close to each other elements. The measure of difference is the square of the Euclidean distance.

Keywords: train, situation recognition, traffic safety, clustering, classifier.