

ПРОГНОЗУВАННЯ ПАРАМЕТРІВ БАГАТОВИМІРНИХ ПРОСТОРОВО-ЧАСОВИХ РЯДІВ З УРАХУВАННЯМ ВЛАСТИВОСТІ НЕСТАЦІОНАРНОСТІ РОЗПОДІЛУ ЧАСОВОЇ СКЛАДОВОЇ

Новожилова М. В. – д-р фіз.-мат. наук, професор, завідувач кафедри комп'ютерних наук та інформаційних технологій, Харківський національний університет міського господарства імені О. М. Бекетова, м. Харків, Україна.

Андронов В. А. – д-р техн. наук, професор, проректор, Національний університет цивільного захисту України, м. Харків, Україна.

Мележик Р. С. – ад'юнкт, Національний університет цивільного захисту України, м. Харків, Україна.

АНОТАЦІЯ

Актуальність. Актуальність роботи полягає в розробці методів збору, аналізу та обробки просторово-часової інформації, а саме рядів даних, розподілених як у просторі, так і у часі та побудові на цій основі комп'ютерної ймовірнісної моделі процесу прогнозування виникнення та розвитку надзвичайних ситуацій техногенного характеру на інженерній інфраструктурі міста. Просторово-часовий характер рядів даних визначає додаткові вимоги до процедур ідентифікації математичної моделі ряду, отже, в статті пропонується низка підходів до виявлення його структури та побудови моделі ряду.

Мета. Створення методична та програмна реалізація комп'ютерної моделі просторово-часового ряду, що використовується для прогнозування майбутніх значень локацій та часу виникнення надзвичайних ситуацій на інженерній інфраструктурі мегаполісу і підвищення ефективності прийняття рішень.

Метод. Розроблено проєкційний підхід, що передбачає незалежне визначення випадкових просторових параметрів аварійних ланок інженерної інфраструктури як послідовності двох одновимірних рівномірних розподілів та визначення часового розподілу моментів виникнення аварій як нестационарного розподілу Пуассона. Запропоновано інтегральний підхід, який включає побудову точок-генераторів, потужність яких (характеристика складності аварії) визначається на основі реалізації підходу порівнювальної статистики з так званим накопичувальним ефектом у межах певного часу. Побудовано релаксаційний підхід до визначення параметрів моделі вихідного просторово-часового ряду, що базується на кластеризації просторової складової за відповідною часовою характеристикою. Проведено формалізацію просторово-часового поля, побудовано процедури інформаційного забезпечення процесу прогнозування параметрів просторово-часового ряду, виконано типізацію вихідних даних для проведення чисельних експериментів на прикладі мережі водопостачання та водовідведення комунального підприємства «Харківводоканал», м. Харків.

Результати. Запропоновано дуальну методологію щодо визначення параметрів імітаційної моделі просторово-часового ряду, що містить як проєкційний так і інтегральний підходи, а також комбіновану методику – релаксаційний підхід. Проведено чисельні експерименти на основі побудованої моделі, що є теоретичною основою для здійснення та побудови прогнозу з використанням великого обсягу історичних даних.

Висновки. В роботі одержав подальший розвиток метод прогнозування параметрів просторово-часових рядів з урахуванням властивості нестационарності розподілу часової складової. Застосування запропонованих інструментальних засобів комп'ютерного моделювання дозволяє підвищити точність прогнозу локації, часу настання та тяжкості можливої аварії на інженерній інфраструктурі мегаполісу.

КЛЮЧОВІ СЛОВА: просторово-часові ряди, нестационарний розподіл Пуассона імітаційна модель, інженерна інфраструктура.

НОМЕНКЛАТУРА

ab – осередок регулярної сітки;
 a – вимір сітки за координатою x ;
 b – вимір сітки за координатою y ;
 A_m – настання події на кластері K_m ;
 D – просторова обмежена область;
 $D\{\bullet\}$ – дисперсія розподілу Пуассона;
 $E\{\bullet\}$ – математичне сподівання;
 EI – опис компонент природного середовища;
 $F(\bullet)$ – рівномірний розподіл просторової координати;
 G – передавальна мережа;
 g_j – ланка передавальної мережі;
 gI_j – інформаційний кортеж характеристик ланки g_j ;
 H – крок розбиття часового періоду;
 I – кортеж інформації про подію;

i – наскрізний номер осередку сітки;
 In_j – множина елементів просторово-часового ряду стосовно ланки g_j ;
 j – індекс ланки передавальної мережі;
 J – кількість ланок передавальної мережі;
 K – множина кластерів розбиття;
 l_m^n – елементи кластерів розбиття;
 M – кількість кластерів;
 n – індекс елементу в m -у кластері;
 m – індекс кластерів розбиття;
 $P_k(t)$ – розподіл ймовірності відмов;
 p_θ – ймовірність реалізації події q_θ ;
 Q – множина несумісних подій;
 q_θ – складна подія;
 R^2 – декартовий простір вимірності 2;
 R_S – назва просторово-часового ряду;

S – кількість осередків регулярної сітки за координатою x ;
 W – кількість осередків регулярної сітки за координатою y ;
 s – індекс осередку регулярної сітки за координатою x ;
 w – індекс осередку регулярної сітки за координатою y ;
 t – змінна часу;
 t_H – період амортизації;
 t_{jc} – дата попереднього ремонту ланки g_j передавальної мережі;
 T – часовий інтервал;
 T_{0j} – початковий рік експлуатації ланки g_j мережі;
 T_F – поточний рік експлуатації ланок мережі;
 T_{FO} – інтервал історичних даних ;
 $v=(x,y)$ – просторова складова елементу ряду;
 v_j – точка-генератор;
 (x_{j-1}, y_{j-1}) – координати початку ланки g_j ;
 (x_j, y_j) – координати кінцевої точки ланки g_j ;
 $Z(v,t)$ – просторово-часове випадкове поле;
 Z_m – рівень збитків на кластері K_m ;
 α_0 – вільний член рівняння авторегресійного процесу;
 α_1 – параметр рівняння авторегресійного процесу;
 $\beta(x,y)$ – параметр закону розподілу просторової складової ряду;
 Δx – розмір осередку регулярної сітки розбиття за координатою x ;
 Δy – розміри осередку регулярної сітки за координатою y ;
 Θ – кількість несумісних подій;
 λ – інтенсивність випадкових відмов;
 $\lambda^{sys}(t)$ – інтенсивність систематичних відмов;
 $\eta(t, t_H)$ – параметр нестационарного закону розподілу часової складової ряду;
 ξ – рівномірно розподілене на $[0, 1]$ випадкове число;
 μ – середнє – параметр нормального закону розподілу;
 σ – стандартне відхилення – параметр нормального закону розподілу;
 Ψ_j – час відновлення аварійної ланки g_j ;
 τ – часова складова елементу ряду;
 \mathcal{G}_j – потужність точки-генератора;
 $\varpi(x,y,t)$ – оціночна функція.

ВСТУП

Проблеми аналізу та моделювання просторово-часових рядів даних виникають в багатьох сферах людської діяльності – геології, економіці, екології, цивільному захисті, медицині тощо. Складність інтерпретації великих обсягів наявних даних, аналізу та подальшого використання у прогнозуванні подій потребує комплексного системного підходу та розробки на цій основі арсеналу різних методів, що характери-

зують ті чи інші властивості процесу, що розглядається.

При побудові комп'ютерної моделі просторово-часового ряду важливим є вивчення особливостей його часової складової, а саме визначення властивості стаціонарності часового ряду. В багатьох галузях застосування моделей та методів обробки просторово-часових рядів, таких, як, наприклад, прогнозування виникнення та розвитку надзвичайних ситуацій техногенного характеру на інженерній інфраструктурі міста, що розглядається як ймовірнісна подія, інформація про яку є розподіленою як у просторі, так і у часі, часова складова є нестационарною.

При цьому також необхідно зважати, що в якості області припустимих значень просторової складової ряду може виступати як неперервна, так і дискретна множина.

Тому важливим є вивчення особливостей таких рядів, які об'єктивно залежать від характеристик перебігу процесу реального світу, що індукує просторово-часовий ряд, та розробка відповідної просторово-часової комп'ютерної моделі як засобу прогнозу параметрів майбутньої ймовірнісної події ряду.

Об'єкт дослідження – просторово-часові ряди даних з урахуванням властивості нестационарності розподілу часової складової.

Предмет дослідження – методи ідентифікації та реалізації імітаційної моделі просторово-часових рядів.

Метою дослідження є розробка, а також методична та програмна реалізація імітаційної моделі багатовимірних просторово-часових рядів з урахуванням властивості нестационарності розподілу часової складової.

1 ПОСТАНОВКА ЗАДАЧІ

В якості вихідних даних виступають статистичні ряди R_S , що мають розглядатися у просторово-часовому просторі

$$(v, t): v \in D \subset R^2, t \in T, \quad (1)$$

де область D може бути як неперервною, так і дискретною.

В даному дослідженні просторово-часовий ряд даних R_S формують результати спостережень за відмовами елементів інженерної інфраструктури мегаполісу – складної технічної системи, що є просторово розподіленою і містить передавальну мережу, яка є скінченною дискретною множиною ланок

$$G = \{g_1, g_2, \dots, g_j\}. G \subset D. \quad (2)$$

При цьому часова складова просторово-часового ряду R_S характеризується інтенсивністю виникнення відмов (випадкових подій), що є функцією часу $\lambda(t)$, тобто часова складова є нестационарною. Таке узагальнення є адекватним описом стану інженерної ін-

фраструктури міст України, що є джерелом випадкових – інтенсивність λ – і систематичних – інтенсивність $\lambda^{st}(t)$ – відмов елементів.

Отже, необхідно вирішити наступні задачі.

1. Розробити методики побудови аналітичної ймовірнісної моделі просторово-часового ряду із нестационарною часовою складовою на основі формалізації просторово-часового випадкового поля

$$Z(v, \tau); v \in D, \tau \in T, \quad (3)$$

де величина $Z(v, \tau)$ – визначена характеристика елементів ряду. При розгляді елементів інженерної інфраструктури величина $Z(v, \tau)$ оцінюється сумарними збитками від наслідків відмов.

2. Розробити програмне забезпечення, що реалізує дані методики.

3. Визначити працездатність та особливості застосування розроблених методик для опрацювання статистичної інформації при дослідженні інженерної інфраструктури міст.

2 ОГЛЯД ЛІТЕРАТУРИ

Системологічний аналіз вітчизняної та закордонної наукової літератури щодо існуючих методів та інструментальних засобів математичного та комп'ютерного моделювання просторово-часових рядів та розробки на цій основі прогнозних моделей дозволив провести певну класифікацію наукових напрямів щодо обробки просторово-часових рядів, що розвиваються, та показав широкий спектр практичних застосувань.

В [1] проведено порівняння множини просторово-часових умовних авторегресійних моделей на основі набору даних про епідемію в Колумбії. Досліджено пуассоно-логнормальні та біноміальні моделі з різними методами байєсівського просторово-часового моделювання.

Проблема побудови прогнозної моделі просторово-часового міського трафіку, яка може бути адаптована в декількох просторово-часових умовах на основі створення складної глибокої нейронної системи, розглянута в [2]. Комп'ютерна модель, що пропонується, базується на інтерпретованій нейромережі, певні складові якої у сукупності охоплюють ключові просторово-часові компоненти рядів даних щодо функціонування транспортної інфраструктури міста.

В [3] наведено метод дослідження просторово-часового розподілу впливу екологічного стану околиць міст на екологічну ефективність останніх. На основі застосування множини підходів до ідентифікації параметрів та економетричних тестів щодо просторово-часового розподілу впливу ефекту такого сусідства на екологічну ефективність міст у якості основної визначено просторово-часову модель Дарбіна та метод часткового диференціалу.

Методи розв'язання проблеми формалізації просторово розподілених даних розвиваються у межах

такої сучасної наукової дисципліни як геостатистика [4]. Згідно з [5], геостатистичні моделі просторово-часової кореляційної структури можна розділити на два види: такі, що передбачають розподіл на просторову та часову компоненти (метрична модель, лінійна модель, модель добутку, модель добуток-сума тощо), та такі, що такого розподілу не припускають. Зважаючи на складність задачі, одним з прийнятних підходів першого типу є методологія порівняльної статистики, тобто дискретизація наявної сукупності статистичних даних за часом та послідовний розгляд множини часових перерізів без опису переходу від одного перерізу до іншого.

Проведене дослідження традиційних і новітніх конструктивних засобів аналізу та обробки просторово розподілених даних дозволило визначити в якості основних: інтерполяційні моделі, а саме лінійну інтерполяцію на основі триангуляції [5], метод обернених середньозважених відстаней (inverse distance weighting) тощо; власне моделі, що базуються на статистичній інтерпретації даних; алгоритми, засновані на навчанні – штучні нейронні мережі [6], генетичні алгоритми [7].

Одним з підходів до прогнозування часових рядів є стохастична теорія черг [8]. Поряд з цим, класичною методологією дослідження потоку відмов обладнання потенційно небезпечних виробничих та інфраструктурних об'єктів є теорія надійності [9].

Розроблені методи опрацювання просторово-часових рядів застосовуються в багатьох галузях діяльності людини.

В [10] побудовано просторову модель геоінформаційної підтримки державних пожежно-рятувальних частин м. Харкова за рахунок врахування та застосування критерію розподілу доріг відповідно до категорій, що дозволило розробити програмну платформу для швидкого реагування.

Методи дослідження ефективності складних технічних систем, зокрема стійкість інфраструктури до можливих ризикованих подій, запропоновані в [11]. Модель, що пропонується, застосована до дослідження випадку інфраструктури водопостачання з гіпотетичним сценарієм відновлення.

При цьому необхідно враховувати факт нестационарності статистичних часових рядів даних щодо аварійних ситуацій на об'єктах інженерної інфраструктури [12]. Аналіз наявних методів економічної оцінки збитків від аварій та надзвичайних ситуацій, ними обумовлених, що застосовуються у світі й Україні, проведено в [13].

Методичні підходи до аналізу та прогнозування природно-техногенної безпеки великих міст України в рамках функціонування механізмів державного регулювання на основі застосування нейромереж і карт Кохонена розглянуто в [14]. Зокрема, за цією методикою м. Харків віднесений до кластера із вищим за середній рівнем частоти прояву ТНС.

У [15] апарат нечітких множин застосований до класифікації надзвичайних ситуацій в мегаполісі, що

грунтується на оцінці впливу параметрів ТНС на громаду та можливості здійснення дорожнього руху на транспортній інфраструктурі міста.

Різноманіття підходів до математичного та комп'ютерного моделювання просторово-часових рядів даних різної природи свідчить про складність проблеми, що розглядається.

Однак переважна кількість наукових публікацій присвячена розгляду задачі прогнозування режимів функціонування складних інфраструктурних систем тільки у часі, без урахування просторової складової.

Ще одним важливим недоліком наявних інструментальних засобів моделювання є практична відсутність методів обробки багатовимірних нестационарних статистичних рядів, і, як наслідок, відсутність адекватних засобів аналізу та синтезу прийнятних управлінських рішень.

3 МАТЕРІАЛИ І МЕТОДИ

Комп'ютерне моделювання багатовимірних просторово-часових рядів має враховувати основні їх особливості, як розподіленість у часі та просторова розподіленість; випадковість моментів настання; недостатній обсяг та зашумленість статистичної інформації, багатомірність та наявність різних факторів впливу зовнішнього середовища на характеристики ряду, у тому числі керування у «ручному» режимі.

Інформація щодо параметрів просторово-часового ряду є відмінною за змістом, формою та вимірністю, тому найбільш прийнятним її поданням у моделі (3) є кортеж I вигляду

$$I = \langle v=(x,y), \tau, \{\beta(x,y), \lambda(t)\} \rangle. \quad (4)$$

Розглянемо послідовно три підходи до побудови моделі просторово-часового ряду за складовими кортежу I .

Перші два з цих підходів: проєкційний та інтегральний – складають так звану дуальну методологію щодо визначення просторових параметрів просторово-часового ряду.

Як проєкційний, так і інтегральний підходи спираються на постановку та вирішення задачі просторового оцінювання.

Проведемо моделювання просторових параметрів – вектору $v=(x,y)$ – багатовимірною ряду (2) у певному часовому перерізі часового параметру t .

У загальному вигляді постановка задачі просторового оцінювання передбачає наявність неперервної двовимірної області D , на якій проведений ряд вимірювань заданої величини Z . Ці вимірювання проводяться на скінченній множині точок $(x, y) \in D$, які є довільним чином розподіленими по області D .

За постановкою задачі необхідно враховувати дискретність області D , яка, власне, перетворюється множини G .

Для моделювання просторової складової ряду, що розглядається, проводиться розбиття контрольованої території D (відповідно, мережі G) сіткою вимірів на © Новожилова М. В., Андронов В. А., Мележик Р. С., 2021
 DOI 10.15588/1607-3274-2021-1-7

осередки. В роботі використовується стандартна прямокутна (регулярна) сітка. При цьому виникає ефект просторової роздільної здатності, тобто задача визначення адекватного розміру осередку $(\Delta x, \Delta y)$.

При використанні стандартної прямокутної сітки просторова роздільна здатність характеризується розміром комірки $\{(a_s, b_w), (a_{s+1}, b_{w+1})\}$:

$$\Delta x = |a_{s+1} - a_s|, \Delta y = |b_{w+1} - b_w|, s=1,2,\dots,S, w=1,2,\dots,W.$$

В цілому область D є вписаною у прямокутник $\{(a_0, a_S), (b_0, b_W)\}$.

Геометричні параметри ланок мережі G визначаються таким чином: складова gI_i задає ланку мережі, що належить осередку регулярної сітки $ab_i = \{(a_s, b_w), (a_{s+1}, b_{w+1})\}$ так, що індекс i є функцією індексів $\{s, w\}$:

$$gI_i = G \cap ab_i.$$

Тоді параметри розміщення (x_{i-1}, y_{i-1}) g_i -ї ланки множини G пов'язуються з лівою нижньою точкою ланки в межах обраного дискрету.

Таким чином, початкове подання (2) мережі G уточнюється стосовно розбиття сіткою вимірів, і складова gI_i кортежу (5) у загальному випадку є множиною, тобто є мультиланкою.

Однак цей спосіб визначення геометричних параметрів ланок множини G має певні недоліки, які необхідно враховувати при виборі кроку дискретизації. Завеликий крок формує дискрет $(\Delta x, \Delta y)$, що містить кілька ланок множини G з принципово різними характеристиками, замалий крок може завдати не виправдано великих втрат часу опрацювання моделі.

На даній сітці вибудовується SxW -вимірна випадкова функція $Z(v)$ (3) з функцією розподілу

$$F(v_1, v_2, \dots, v_{SxW}, z_1, z_2, \dots, z_{SxW}) = P\{Z(v_1) \leq z_1, Z(v_2) \leq z_2, \dots, Z(v_{SxW}) \leq z_{SxW}\}, \quad (5)$$

у кожному часовому перерізі τ , яка характеризує просторову ймовірність реалізації елемента ряду та визначається як набір випадкових змінних $Z(v_i)$, по одній для кожної локації $v_i, i=1,2,\dots, SxW$, області D .

На основі вирішення задачі просторового оцінювання розглянемо підходи до прогнозування параметрів просторово-часового ряду.

Проєкційний підхід. Передбачає незалежне визначення випадкових параметрів (x, y) розміщення аварійних ланок інженерної інфраструктури як послідовності двох одновимірних рівномірних розподілів $\{x\}, \{y\}$ на $[a_0, a_S], [b_0, b_W]$ із функціями розподілу

$$F(x) = \begin{cases} 0, & x \leq a_0; \\ \frac{x-a_0}{a_S-a_0}, & x \in (a_0, a_S]; \\ 1, & x > a_S. \end{cases} \quad F(y) = \begin{cases} 0, & y \leq b_0; \\ \frac{y-b_0}{b_W-b_0}, & y \in (b_0, b_W]; \\ 1, & y > b_W. \end{cases}$$

відповідно. В результаті двократного застосування процедур генерації випадкових величин за цими розподілами отримується пара просторових координат (x, y) , що визначає комірку $v_i, i=1,2,\dots, SxW$, області D та відповідну ланку g_j множини $G, g_j \in v_i$.

Для подальшої обробки інформації про часову складову вихідного багатовимірного ряду зауважимо наступне.

Періоду нормальної експлуатації складних технічних систем відповідає стаціонарний пуассонівський (найпростіший) потік [16] числа k випадкових відмов протягом часу t вигляду

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad (6)$$

якому притаманні властивості стаціонарності, відсутності післядії і ординарності.

Факт розподілу ймовірності відмов $P_k(t)$ за законом Пуассона з постійною інтенсивністю λ в період нормального функціонування, що визначається періодом амортизації $[0, t_H]$ складної технічної системи, підтверджується також результатами аналізу наявних статистичних даних та досвіду експлуатації технічних систем.

Поряд з цим у життєвому циклі будь-якої складної технічної системи неминуче настає фаза старіння. На даному етапі особливе значення набувають систематичні відмови, обумовлені закономірними і неминучими явищами, що викликають поступове накопичення ушкоджень, втому, зношення обладнання.

Аналіз історичних рядів статистичних даних та досвіду експлуатації складних технічних систем показує, що в період старіння об'єкта дослідження інтенсивність систематичних відмов $\lambda^{syst}(t)$ зростає лінійно або квадратично.

Тоді параметр $\eta(t, t_H)$ розподілу Пуассона є функцією $\lambda^{syst}(t)$ та приймає вигляд:

$$\eta(t, t_H) = \lambda_{t \geq t_H}(t) = \lambda + \lambda^{syst}(t) \Big|_{t \geq t_H}. \quad (7)$$

Таким чином, при перевищенні періоду амортизації – в період старіння – пуассонівський потік θ відмов отримує властивість нестационарності, тобто перестає бути найпростішим

$$P_\theta(t, t_H) = \frac{\eta(t, t_H)^\theta}{k!} e^{-\eta(t, t_H)}. \quad (8)$$

Зауваження 1. Для подальшого моделювання функцію інтенсивності систематичних відмов $\lambda^{syst}(t)$ на інтервалі $[t_H, t]$ доцільно замінити кусково-постійною функцією вигляду

$$\lambda^{syst} = \lambda_{c-1}^{syst}, \text{ якщо } t_{c-1} \leq \tau < t_c, c=1,2,\dots,C, t_c \in [t_H, t].$$

Таким чином, завдання номеру ланки g_j множини $G, g_j \in v_i$, обумовлює конкретне значення параметру $\lambda(t)$ – нестационарного розподілу Пуассона часової складової ряду в залежності від вищезазначених параметрів та урахуванням Зауваження 1.

Величина $Z(v, \tau)$ є зваженою зростаючою функцією інтенсивності $\lambda(t)$.

Інтегральний підхід. Дана концепція містить побудову точок-генераторів, потужність яких визначається на основі реалізації підходу порівнювальної статистики з так званим накопичувальним ефектом у межах певного часу.

Введемо наступні визначення.

Визначення 1. Назвемо точкою-генератором v_i певну точку комірки $\{(a_s, b_w), (a_{s+1}, b_{w+1})\}$ регулярної сітки території D таку точку, яка інкапсулює в собі властивості цілої комірки.

Наприклад, це може бути точка перетину діагоналей комірки $\{(a_s, b_w), (a_{s+1}, b_{w+1})\}$.

Визначення 2. Назвемо потужністю ϑ_i точки-генератору v_i інтегральну характеристику, що розглядається в певному часовому інтервалі (часовому вікні) $[T_0, T_F]$ та обчислюється за формулою:

$$\vartheta_i = \int_{T_0}^{T_F} \int_{b_m}^{b_{m+1}} \int_{a_s}^{a_{s+1}} \varpi(x, y, t) dx dy dt, \quad (9)$$

де функція $\varpi(x, y, t)$ приймає значення $\{0,1\}$, де $\varpi(x, y, t) = 1$, якщо елемент $(x,y,t) \in R_S$, де $t \in [T_0, T_F]$, $(x,y) \in \{(a_s, b_w), (a_{s+1}, b_{w+1})\}$ та $\varpi(x, y, t) = 0$ в іншому випадку.

Визначення точок-генераторів проводиться із застосуванням кластерного аналізу із визначенням розміру кластеру, а саме (2–5) % від максимального лінійного розміру контрольованого району. При цьому покладаємо, що всі точки відповідної комірки сітки вимірів мають потужність ϑ_i точки-генератору v_i .

Побудова точок-генераторів уможливило вирішення задачі моделювання просторово-часового ряду R_S та здійснення подальшого прогнозування на основі побудови двовимірної апроксимаційної поверхні $\vartheta(x,y)$ із застосуванням підходу порівнювальної статистики з так званим накопичувальним ефектом.

Зауважимо, що потужність ϑ_i залежить від довжини часового вікна $[T_0, T_F]$.

Розглянемо ще один метод моделювання багатовимірного просторово-часового ряду R_S , а саме релаксаційний підхід.

Цей підхід базується на зведенні багатовимірного ряду R_S до множини незалежних одновимірних нестационарних (зокрема стаціонарних) часових рядів.

Узагальнено метод складається з двох етапів.

На першому етапі формуються кластери ланок множини $G=\{g_1, g_2, \dots, g_j\}$ за показником T_{0j} . Крок H розбиття на кластери покладається 5 (7) років. В результаті одержуємо множину $K=\{K_1, \dots, K_M\}$ кластерів

ланок, кожен з цих кластерів геометрично є незв'язною множиною.

Для елементів кластеру K_m фіксується параметр розподілу Пуассона $\eta(t, t_H)$, тобто потік моментів настання подій на ланках мережі певного кластеру приймається стаціонарним.

Другий етап релаксаційного підходу містить кроки створення так званого імітаційного симулятора, що генерує певну кількість випробувань.

Зауваження 2. Вважаємо, що результат випробування є складною подією: можуть виникнути події на різних кластерах мережі одночасно (наприклад, протягом тижня – одиницю виміру часу обирає особа, що приймає рішення).

Таке припущення щодо урахування кількох джерел настання подій в цілому, як зазначалось, порушує вимогу ординарності процесу. Тому для подальшого моделювання застосовується наступний підхід.

Позначимо через A_m – настання події на кластері K_m у одиницю часу, відповідно \bar{A}_m – ненастання події на кластері K_m .

На цій основі визначимо повну дискретну множину Q несумісних подій вигляду

$$Q = \left\{ \prod_{m=1}^M A_m, \bar{A}_1 \prod_{m=2}^M A_m, \dots, \bar{A}_M \prod_{m=1}^{M-1} A_m, \bar{A}_1 \bar{A}_2 \prod_{m=3}^M A_m, \dots, A_M \prod_{m=1}^{M-1} \bar{A}_m, \prod_{m=1}^M \bar{A}_m \right\}. \quad (10)$$

При цьому кількість елементів множини Q : $\text{card } Q = \Theta$.

Для кожного елемента множини (9) визначимо ймовірність p_θ реалізації відповідної складної події q_θ , $q_\theta \in Q$, $\theta = 1, 2, \dots, \Theta$. Далі отримані значення ймовірностей p_θ відкладаються на відрізку $[0, 1]$. Генерується рівномірно розподілене на $[0, 1]$ випадкове число ξ і перевіряється умова

$$\sum_{\mu=1}^{\theta-1} p_\mu \leq \xi \leq \sum_{\mu=1}^{\theta} p_\mu. \quad (11)$$

За виконання умови (11) вважають, що в процесі випробування настала подія q_θ з множини (10).

Реалізація запропонованих методів уможливується в рамках інформаційної підсистеми, функціональну модель якої наведено на рис. 1.

У блоці геоінформаційної системи виконується аналіз і візуалізація графічної інформації у вигляді набору електронних карт і прив'язаних до них даних про просторові параметри ряду R_S .

Інформація у підсистемі зберігання та обробки даних організується як реляційна база даних, що підтримує властивість розподіленої обробки даних. В якості базового інструментарію реалізації функцій аналітичної підсистеми служить методичне забезпечення (математичні моделі, методи імітаційного програмування, засоби обробки експертної інформації з

метою отримання знань) задач побудови прогнозних значень ряду.

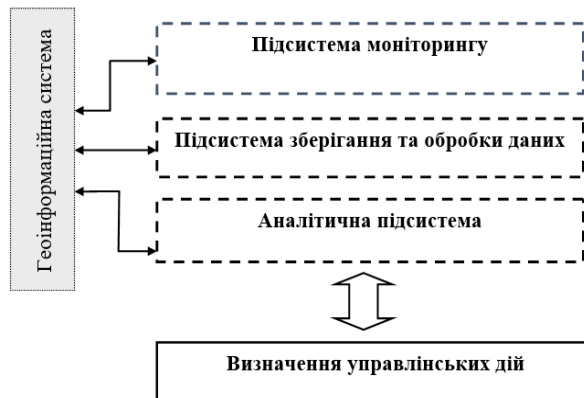


Рисунок 1 – Функціональна модель інформаційної аналітичної системи

У підсистемі моніторингу виконується збирання та первинна обробка інформації про параметри процесів реального світу, що генерують значення просторово-часового ряду.

4 ЕКСПЕРИМЕНТИ

Інтерпретація задачі просторового оцінювання здійснювалась на вихідному просторово-часовому ряді значень локацій, часу і складності аварій, що формувався за даними офіційного сайту комунального підприємства «Харківводоканал».

Моделювання проводилось за районами м. Харкова, зокрема, розглядався Шевченківський район. Загалом проаналізовано та упорядковано більше 4400 спостережень за 2016 – 2019 роки.

На рис. 2 наведено результат побудови регулярної прямокутної сітки у якості розбиття контрольованої території D (території району) та відповідно, мережі G .

Ілюстрацією застосування інтегрального підходу є табл. 1, де наведено приклад обчислення потужностей ϑ_i точок-генераторів v_i мережі водопостачання території Шевченківського району м. Харкова.

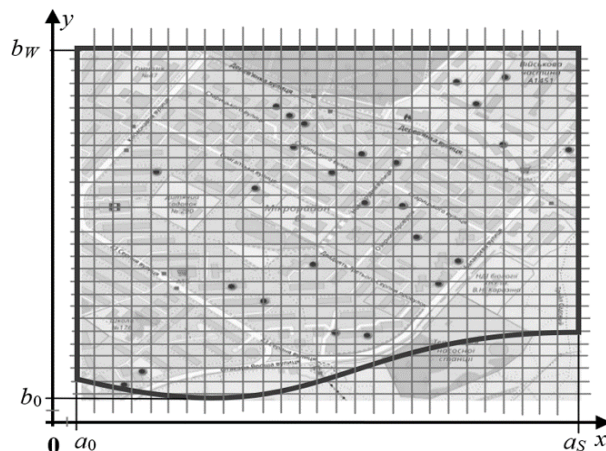


Рисунок 2 – Карта Шевченківського району, м. Харків, з побудованою сіткою. Період 1–10 березня 2018 р. Темні точки – місця аварій

Таблиця 1 – Потужності ϑ , точок-генераторів v , мереж водопостачання території Шевченківського району м. Харкова за березень 1918 р.

		Координата у								
		2,5	7,5	12,5	17,5	22,5	27,5	32,5	37,5	42,5
Координата х	-2,5	3	2	2	1	1	1	3		
	-7,5	4	2	2	2					
	-12,5				3	2	1	2		
	-17,5		3		2	2	3	2	1	
	-22,5	3		1	2	1				
	-27,5	3	1			2	2	1		
	-32,5			1	1	1	2	2	1	
	-37,5	1	4	1		2	1	2		
	-42,5	1								

В комірках табл. 1 наведено кількість аварій на ланках мереж водопостачання, що належать визначеним коміркам регулярної сітки розбиття.

Вибір напрямку координатних осей інформації табл. 1 узгоджений із розташуванням Шевченківського району відносно географічного центру м. Харкова.

За даними табл. 1 побудовано поверхню $\vartheta(x,y)$, яку подано на рис. 3. Більш насиченим кольором на рис. 3 забарвлені комірки з вищою кількістю аварій.

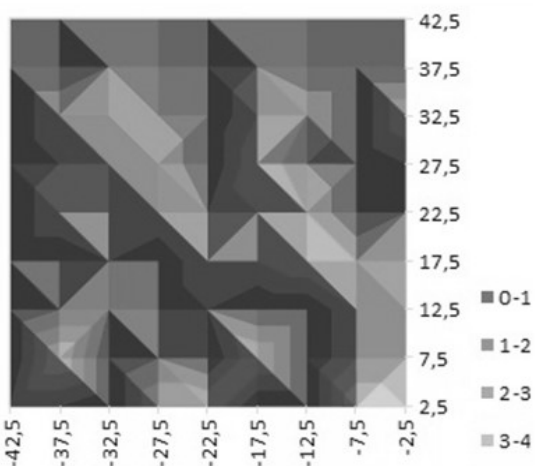


Рисунок 3 – Скалярне поле небезпеки за спостереженнями 1–10 березня 2018 р.

Розглянемо етапи алгоритму реалізації релаксаційного підходу.

Етап 1 – підготовчий.

Крок 1.1. Підготовка вхідних даних. Визначення елементів кортежу (4) стосовно території Шевченківського району м. Харкова, що розглядається.

Крок 1.2. Визначення довжини статистичних рядів даних щодо характеристик ланок мережі за формулою

$$T_{F0} = (T_F - T_0).$$

Крок 1.3. Визначення кроку розбиття H довжини T_{F0} , що застосовується у зворотному порядку. При цьому покладемо, що початковий рік експлуатації ланок мережі, що розглядається – це 1940 р., останній рік початку експлуатації – 1989 рік.

Наприклад, при кроці $H=7$, відповідно визначено 7 кластерів: $[\leq 1940-1950]$, $[1951-1957]$, $[1958-1965]$, $[1966-1973]$, $[1974-1981]$, $[1982-1989]$, $[1990 - \text{по цей час}]$.

Крок 1.4. Кластеризація множини ланок мережі G за терміном введення в експлуатацію. Як зазначалось, кластери K_m , $m=1,2, \dots, M$, в загальному випадку є геометрично незв'язними множинами.

На рис. 4 наведений приклад кластеризації мережі водопостачання району м. Харкова, де елементи кластерів мережі водопостачання виділені різними градаціями сірого.



Рисунок 4 – Кластеризація передавальних мереж

Крок 1.5. Перевірка гіпотези щодо розподілу Пуассона числа k випадкових відмов протягом періоду t на кластерах K_m вигляду (6).

Критерієм справедливості гіпотези про розподіл Пуассона моментів настання ТНС на мережах водопостачання на кластерах g_j , $j=1,2,\dots,J$, є виконання умов

$$E\{X\}=\lambda_j, D\{X\}=\lambda_j. \quad (12)$$

Елементи k_m^n певного кластеру $K_m = \{k_m^1, \dots, k_m^n, \dots, k_m^{N_m}\}$ у подальшому вважаються джерелами (генераторами) аварій, що характеризуються однаковими значеннями величини $Z(v, \tau): Z_m$.

Згідно Наказу [17] подія, яка може трапитись на мережах водопостачання в разі дій несприятливих чинників техногенного походження, буде кваліфікуватись як надзвичайна ситуація, якщо припинено водопостачання більше 30% розрахункового обсягу водопостачання населеного пункту (міста) через аварії

на об'єктах водопостачання або пошкодження мереж водопостачання.

Містом в Україні вважається населений пункт з населенням більше 10000 осіб.

Пропозиція 2. Виходячи з того, що населення Шевченківського району складає 15 відсотків населення м. Харкова, тобто приблизно 230000 людей, будемо вважати що аварії на ділянках мережі водопостачання з діаметром труби I-го класу – від 1000 до 1200 мм, тобто на магістральних водогонях, що тривають 2 або більше діб, задовольняють умови визначення надзвичайної ситуації.

Узагальнена інформація щодо характеристик розподілів та емпіричних значень параметрів λ (інтенсивність настання аварій у місяць) наведена в табл. 3.

Зауваження 3. Відмітимо, що в Табл. 3 враховані всі аварії на мережах водопостачання, в тому числі ті, що не можуть бути класифіковані як надзвичайні ситуації, наприклад, коли тривалість аварії менше 18 годин.

Аналіз наявної статистичної інформації надає можливість зробити висновок, що аварії на мережах водопостачання з тривалістю локалізації та ліквідації аварії більше за 18 годин складають приблизно 11–15% (10–21 аварій в місяць, 0–4 аварій в день) від загальної кількості аварій за всіма кластерами.

Таблиця 3 – Характеристики розподілів та емпіричних значень параметрів λ

Характеристики статистичного ряду	Кластери					
	I	II	III	...	VI	VII
Усього значень	988	950	862	293	249
Середнє (за місяць)	27,44	26,39	23,94	...	8,14	6,91
Дисперсія (місяць)	25,11	29,21	21,43	...	9,55	7,22
λ_m	27,44	26,39	23,94	...	8,14	6,91

Однак статистичний ряд моментів настання таких аварій теж добре апроксимується розподілом Пуассона.

Етап 2 – основний. Побудова та запуск програмного симулятора, тобто генерація певної кількості випробувань. Реалізація складових процесу імітаційного моделювання просторово-часового ряду проводиться в рамках парадигми дискретно-подієвого моделювання протягом деякого періоду модельного часу як ітераційний процес, на кожній ітерації якого визначається функція Z .

Крок 2.1. Визначення ймовірностей настання події (аварії, надзвичайної ситуації) A_m на кластері K_m у одиницю часу та, відповідно, ненастання події \bar{A}_m на кластері K_m із урахуванням Зауваження 1.

Крок 2.2. Завдання повної дискретної множини $Q = \{q_\theta\}$, $\theta=1,2,\dots, \Theta$ несумісних подій вигляду (10), та для кожного елемента множини Q визначення ймо-

вірності p_θ реалізації відповідної складної події q_θ , $q_\theta \in Q$, $\theta=1,2,\dots, \Theta$.

Фрагмент визначення необхідних ймовірностей на просторово-часовому ряді параметрів аварій, що розглядається, з тривалістю локалізації та ліквідації більше однієї доби наведений в табл. 4.

Таблиця 4 – Розподіл ймовірностей настання аварій з тривалістю локалізації та ліквідації більше однієї доби

Кластер	λ (доба)	Максимально можлива кількість подій	Ймовірність настання події			
			Жодної події – A_{m0}	Одна подія – A_{m1}	Дві події – A_{m2}	Три події – A_{m3}
I	0,68	3	0,512	0,343	0,115	0,026
II	0,51	3	0,601	0,306	0,078	0,013
III	0,47	3	0,625	0,294	0,069	0,011
...

Крок 2.3. Визначення індексу θ події q_θ з множини (9), що реалізується на даній ітерації, за допомогою генерації рівномірно розподіленого на $[0, 1]$ випадкового числа ξ і перевірки умови (11).

Крок 2.4. Визначення часу локалізації та ліквідації аварії, котрий можна інтерпретувати як час обслуговування Ψ_j аварійного елемента g_j .

Дослідження вихідного просторово-часового ряду аварій за даними офіційного сайту комунального підприємства «Харківводоканал» показало, що час Ψ_j задовольняє нормальному розподілу з заданим середнім μ та стандартним відхиленням σ :

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} e^{-\left(\frac{\ln t - \mu}{2\sigma}\right)^2}, t \geq 0, \quad (13)$$

де параметри μ та σ , що визначаються на кластері K_m , є функціями від рівня Z_m збитків.

Крок 2.5. Визначення загального обсягу збитків в даному випробуванні за умови нормування за часом.

Крок 2.6. Заключний. Визначення середніх (сумарних) характеристик експерименту

В якості середньої (сумарної) характеристики експерименту можна вважати інтенсивність λ .

Зауваження 4. Із теорії ймовірності [16] відомо, що якщо число моментів настання подій за певний проміжок часу задовольняє розподіл Пуассона, то розподіл інтервалів часу між послідовними надходженнями подій відповідає експоненціальному розподілу.

Якщо λ_m – інтенсивність появи подій в розподілі Пуассона, то розподіл часу t між послідовними надходженнями заявок визначається щільністю ймовірності

$$f(t) = \lambda_m e^{-\lambda_m t}. \quad (14)$$

Таким чином уможливується моделювання настання моментів ТНС на часовій осі на певних клас-терах.

Отже, вихідними даними для моделювання є такі:

- масив ймовірностей настання події A_m на кластері K_m у одиницю часу $P_1(t) = (\lambda_m t) e^{-\lambda_m t}$, та, відповідно, ймовірностей $(1 - P_1(t))$ події \bar{A}_m ненастання ТНС на кластері K_m ;

- параметри μ та σ нормального розподілу часу обслуговування Ψ_j .

Для запуску симулятора (кроки 2.2. – 2.4 реалізації релаксаційного підходу) задаються:

- рік прогнозування, що впливає на розподіл ланок інженерних мереж території, що розглядається, за кластерами;

- одиниця часу моделювання;

- точка відліку на часовій осі;

- горизонт планування.

Реалізацію запропонованого алгоритму в рамках інформаційної підсистеми, функціональну модель якої наведено на рис. 1, проведено із застосуванням системи Scilab (версія 6.1.0 від 25.02.2020) – кросплатформеної системи комп'ютерної алгебри.

5 РЕЗУЛЬТАТИ

При проведенні комп'ютерних експериментів щодо прогнозування параметрів просторово-часових статистичних рядів за наведеними методиками визначені наступні особливості застосування проекційного, інтегрального та релаксаційного методів. Перші два методи застосовуються при необхідності ідентифікації просторових параметрів елементів ряду на горизонті планування, тоді як релаксаційний метод дозволяє абстрагуватися від безпосереднього обчислення просторових параметрів і розширити при цьому горизонт планування.

Аналіз ряду інтенсивності λ_m аварій тривалістю ліквідації більше за добу по кластерах K_m з проведеним з використанням тесту Дікі-Фулера. Рівняння відповідного авторегресійного процесу має вигляд:

$$\lambda_m = \alpha_{0m} + \alpha_{1m} \lambda_m,$$

де $\alpha_{0m} \in (0,01 \dots 0,019)$, $\alpha_{1m} \in (1 \dots 1,1)$, що підтверджує властивість нестационарності ряду.

Результати моделювання та побудови прогнозу значень λ_m по кластерах K_m наведені на рис. 6.

Базові значення λ_m – нижня лінія графіку. Вищі лінії – значення λ_m на період та на два періоди за кроком розбиття H вперед відповідно.

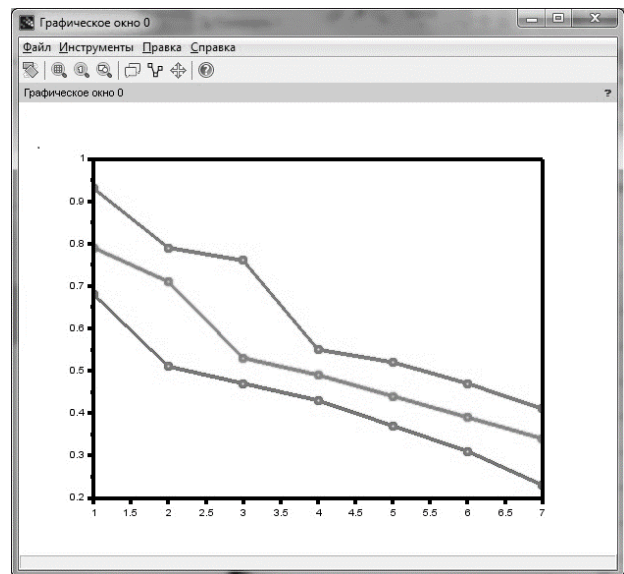


Рисунок 5 – Графічне подання результатів моделювання

Проведені чисельні експерименти за наведеними методиками з історичними рядами даних щодо мереж водопостачання та водовідведення м. Харкова, класифікованими за характеристикою величини оцінки збитків (аварія – надзвичайна ситуація) показали схожість динаміки цих рядів.

6 ОБГОВОРЕННЯ

Проведені етапи комп'ютерного моделювання параметрів просторово-часового ряду за даними щодо мереж водопостачання та водовідведення м. Харкова підтверджують базові припущення моделі щодо характеристик моментів надходження відмов обладнання (настання подій), які є дискретною випадковою величиною; наявність поступових та систематичних відмов обладнання; просторово-часову розподіленість випадкового потоку подій; необхідність урахування кількох джерел настання подій, що в цілому порушує вимогу ординарності процесу. В даній моделі за результатами аналізу наявних історичних даних визначено верхню межу кількості аварій, що відбуваються одночасно – три аварії.

Розроблене програмне забезпечення методики прогнозування параметрів можливої аварії на прикладі множини даних щодо мереж водопостачання та водовідведення, м. Харкова дає можливість її безпосереднього застосування для інших типів систем життєзабезпечення мегаполісу.

ВИСНОВКИ

1. В роботі вирішена актуальна задача з розробки методик побудови аналітичної ймовірнісної моделі просторово-часового ряду із нестационарною часовою складовою на основі формалізації просторово-часового випадкового поля.

2. Розроблено функціональну модель та програмну реалізацію інформаційної аналітичної системи, що реалізує дані методики.

3. Визначено працездатність, особливості та межі застосування розроблених методик для опрацювання статистичної інформації при дослідженні інженерної інфраструктури міст.

Наукова новизна отриманих результатів полягає в тому, що отримали подальший розвиток методи прогнозування параметрів просторово-часових рядів з урахуванням властивості нестационарності розподілу часової складової.

Практичне значення отриманих результатів полягає в побудові алгоритмічної та програмної реалізації імітаційної моделі, що розглядається, як середовища комп'ютерного моделювання зниження масштабів та пом'якшення наслідків впливу небезпечних факторів аварій та надзвичайних ситуацій та інженерній інфраструктурі міста.

Подібна формалізація закладає теоретичну базу побудови методичного й інформаційного забезпечення прогнозування ризику зниження рівня техногенної безпеки мегаполіса як засобу оптимізації відповідних управлінських рішень.

ПОДЯКИ

Робота виконана в рамках держбюджетної НДР Національного університету цивільного захисту України (м. Харків) «Побудова прогнозу виникнення і розвитку надзвичайних ситуацій техногенного характеру» (№ ДР0110U001004).

ЛІТЕРАТУРА / LITERATURA

- Ye J. Comparing different spatio-temporal modeling methods in dengue fever data analysis in Colombia during 2012–2015 / J. Ye., M. J. Moreno-Madriñán // *Spatial and Spatio-temporal Epidemiology*. – 2020. – Vol. 34. – 100360. DOI: /10.1016/j.sste.2020.100360.
- Medrano R. A spatio-temporal attention-based spot-forecasting framework for urban traffic prediction / R. Medrano, J. L. Aznarte // *Applied Soft Computing*. – 2020. – Vol. 96. – 106615. DOI:10.1016/j.asoc.2020.106615.
- The impact of the spatio-temporal neighborhood effect on urban eco-efficiency in China / [P. Chen, R. Xie, M. Lu et al.] // *Journal of Cleaner Production*. – In Press. Available online 30 October 2020. DOI:10.1016/j.jclepro.2020.124860.
- Murakami D. Geostatistics and Gaussian process models / D. Murakami, Y. Yamagata, T. Hirano // *Spatial Analysis Using Big Data*. – Academic Press, 2020. – P. 57–112. DOI:10.1016/b978-0-12-813127-5.00004-7.
- Kyriakidis P. C. Geostatistical space-time models: a review / P. C. Kyriakidis, A. Journel // *Mathematical Geology*. – 1999. – Vol. 31. – P. 651–684.
- Rashid T. Make Your Own Neural Network/ T. Rashid. – CreateSpace Independent Publishing Platform 2016. – 222 p.
- Субботін С. О. Неітеративні, еволюційні та мультиагентні методи синтезу нечіткологічних і нейромережних моделей : монографія / С. О. Субботін, А. О. Олійник, О. О. Олійник. – Запоріжжя : ЗНТУ, 2009. – 375 с.
- Dudin A. N. Optimal hysteresis control for a BMAP/SM/1/N queue with two operation modes / A. N. Dudin, S. Nishimura // *Mathematical Problems in the Engineering*. – 2000. – № 5. – С. 397–420.
- Methods for determining and processing probabilities «Red Book» / [eds.: Schller J. C. H., Brinkman L., Van Gestel P. J., van Otterloo R. W.]. – The Netherlands: Committee for Prevention of Disasters, 1997. – 604 p.
- Tvoroshenko I. S. Software determination of the optimal route by geoinformation technologies / I. S. Tvoroshenko, O. O. Kramarenko // *Radio Electronics, Computer Science, Control*. – 2019. – № 3. – P. 131–142. DOI:10.15588/1607-3274-2019-3-15
- Ongkowiyo C. S. Risk-based resilience assessment model focusing on urban infrastructure system restoration / C. S. Ongkowiyo, H. Doloi // *Procedia Engineering*. – 2018. – Vol. 212. – P. 1115–1122. DOI:10.1016/j.proeng.2018.01.144.
- Попов В. М. Моделирование характеристик потока отказов основных производственных фондов объектов повышенной опасности / В. М. Попов, И. А. Чуб, М. В. Новожилова // *Проблеми надзвичайних ситуацій*. – 2015. – Вип. 21. – С. 93–98.
- Комарницький І. М. Оцінка техногенних збитків та аналіз підходів до їхнього розрахунку у глобальному та регіональному аспектах. [Електрон. ресурс] / І. М. Комарницький, М. І. Бублик. – Режим доступу: vlp.com.ua/files/21_31.pdf.
- Іщенко Г. Г. Аналіз та прогноз природно-техногенної безпеки великих міст із застосуванням нелінійних методів/ Г. Г. Іщенко // *Економіка і держава*. – 2009. – № 8. – С. 48–53.
- He Y. Classification of urban emergency based on fuzzy analytic hierarchy process/ Y. He, S. Du // *Procedia Engineering*. – 2016. – Vol. 137. – P. 630–638.
- Губаревич О. В. Надійність і діагностика електрообладнання: Підручник / О. В. Губаревич. – Северодонецьк : вид-во СНУ ім. В. Даля, 2016. – 248 с.
- Про затвердження Класифікаційних ознак надзвичайних ситуацій: Наказ Міністерства внутрішніх справ України від 06.08.2018р. № 658. [Електрон. ресурс]. – Режим доступу: <https://zakon.rada.gov.ua/laws/show/z0969-18#Text>.

Стаття надійшла до редакції 18.11.2020.
Після доробки 26.01.2021.

УДК 005:355:[519.8::504.4]

ПРОГНОЗИРОВАНИЕ ПАРАМЕТРОВ МНОГОМЕРНЫХ ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ РЯДОВ С УЧЕТОМ НЕСТАЦИОНАРНОСТИ РАСПРЕДЕЛЕНИЯ ВРЕМЕННОЙ СОСТАВЛЯЮЩЕЙ

Новожилова М. В. – д-р физ.-мат. наук, профессор, заведующий кафедрой компьютерных наук и информационных технологий, Харьковская национальная академия городского хозяйства, г. Харьков, Украина.

Андронов В. А. – д-р техн. наук, профессор, проректор, Национальный университет гражданской защиты Украины, г. Харьков, Украина.

Мележик Р. С. – адъюнкт, Национальный университет гражданской защиты Украины, г. Харьков, Украина.

АННОТАЦИЯ

Актуальность. Актуальность работы заключается в создании методов анализа и обработки пространственно-временной информации, а именно рядов данных, распределенных как в пространстве, так и во времени, и построении на этой основе компьютерной вероятностной модели процесса прогнозирования возникновения и развития чрезвычайных ситуаций техногенного характера на инженерной инфраструктуре города. Пространственно-временной характер рядов данных определяет дополнительные требования к процедурам идентификации математической модели ряда, и в статье предлагается ряд подходов к ее структурной и параметрической идентификации.

Цель. Создание, методическая и программная реализация компьютерной модели пространственно-временного ряда, который будет использоваться для прогнозирования будущих значений локаций и времени возникновения чрезвычайных ситуаций на инженерной инфраструктуре мегаполиса и повышения эффективности принятия управленческих решений.

Метод. Разработан проекционный подход, предусматривающий независимое определение случайных пространственных параметров аварийных звеньев инженерной инфраструктуры как последовательности двух одномерных равномерных распределений и определения временного распределения моментов возникновения аварий как нестационарного распределения Пуассона. Предложен интегральный подход, включающий построение точек генераторов, мощность которых (характеристика сложности аварии) определяется на основе реализации подхода сравнительной статистики с так называемым накопительным эффектом в пределах определенного времени. Построен релаксационный подход к определению параметров модели исходного пространственно-временного ряда, основанный на кластеризации пространственной составляющей в соответствии с временной характеристикой.

Проведена формализация пространственно-временного поля, построены процедуры информационного обеспечения процесса прогнозирования параметров возможной чрезвычайной ситуации, выполнена типизация исходных данных для проведения численных экспериментов на примере сети водоснабжения и водоотведения коммунального предприятия «Харьковводоканпл», г. Харьков.

Результаты. Предложена дуальная методология определения параметров имитационной модели пространственно-временного ряда, содержащая как проекционный так и интегральный подходы, а также комбинированная методика – релаксационный подход. Проведены численные эксперименты на основе построенной модели, которая является теоретической основой для осуществления и построения прогноза с использованием большого объема исторических данных.

Выводы. В работе получил дальнейшее развитие метод прогнозирования параметров пространственно-временных рядов с учетом свойства нестационарности распределения временной составляющей. Применение предложенных инструментальных средств компьютерного моделирования позволяет повысить точность прогноза локации, времени наступления и тяжести возможной аварии на инженерной инфраструктуре мегаполиса.

КЛЮЧЕВЫЕ СЛОВА: пространственно-временные ряды, нестационарное распределение Пуассона, имитационная модель, инженерная инфраструктура.

UDC 005:355:[519.8::504.4]

COMPUTER MODELING PARAMETERS OF TECHNOGENIC EMERGENCY SITUATIONS ON ENGINEERING INFRASTRUCTURE OF THE MEGAPOLIS

Novozhylova M. V. – Dr. Sc., Professor, Head of the Department of Computer Science and Information Technologies, O. M. Beketov National University of Urban Economy in Kharkiv, Kharkiv, Ukraine.

Andronov V. A. – Dr. Sc., Professor, Vice-Rector of the National University of Civil Defense of Ukraine, Kharkiv, Ukraine.

Melezhik R. S. – Adjunct of the National University of Civil Defense of Ukraine, Kharkiv, Ukraine.

ABSTRACT

Context. The urgency of the research is to develop methods for analyzing and processing space-time information, namely the set of data distributed both in space and time and creating on this basis a computer probabilistic model of the process of predicting man-made emergencies on city engineering infrastructure. The spatio-temporal nature of data series causes additional requirements for the identification procedures of the mathematical model of a series, therefore, the number of approaches identifying its structure and construction of a series model has been proposed.

Objective is methodical and software implementation of a computer model of the space-time series being intended to predict the future values of locations and times of man-made emergencies on the engineering infrastructure of the metropolis and increase decision-making efficiency.

Method. A projection approach providing independent determination of random spatial parameters defining location of emergency units on engineering infrastructure as a sequence of two one-dimensional uniform distributions and describing time distribution of moments of accidents as non-stationary Poisson distribution has been developed. Proposed is an integrated approach which includes the construction of generator points, the power of which (characteristic of the accident complexity) based on the implementation of the comparative statics approach with so-called cumulative effect within a certain time. A relaxation approach based on the reduction of a two-dimensional simulation model of determining the city of possible emergency location to a set of independent one-dimensional non-stationary (including stationary) distributions to generate the time of occurrence has been constructed. Formalization of the space-time field, procedures of information support of the process of forecasting the parameters of a possible emergency, typification of initial data for numerical experiments on the implementation of methods for forecasting the parameters of a possible emergency on the example of water supply and sewerage network of utility company Kharkivvodokal, city Kharkiv have been developed.

Results. A dual methodology to determine the simulation model parameters of the space-time series, which contains both projection and integral approaches, as well as a combined method – relaxation approach, have been proposed. Numerical experiments

based on the constructed model were performed. The model being considered is the theoretical basis to construct the forecast using a large amount of historical data.

Conclusions. The method to predict the parameters of space-time series considering the nonstationarity property of the time component distribution has been further developed. Using the proposed computer simulation tools allows to increase the accuracy of the forecast of the location, time of occurrence and severity of a possible accident on the engineering infrastructure of the metropolis.

KEYWORDS: space-time series, nonstationary Poisson distribution, simulation model, engineering infrastructure.

REFERENCES

1. Ye J., Moreno-Madriñán M. J. Comparing different spatio-temporal modeling methods in dengue fever data analysis in Colombia during 2012–2015, *Spatial and Spatio-temporal Epidemiology*, 2020, Vol. 34, 100360. DOI: /10.1016/j.sste.2020.100360.
2. Medrano R., Aznarte J. L. A spatio-temporal attention-based spot-forecasting framework for urban traffic prediction, *Applied Soft Computing*, 2020, Vol. 96, 106615. DOI:10.1016/j.asoc.2020.106615.
3. Chen P., Xie R., Lu M., Huang Z. The impact of the spatio-temporal neighborhood effect on urban eco-efficiency in China, *Journal of Cleaner Production*, In Press. Available online 30 October 2020. DOI:10.1016/j.jclepro.2020.124860.
4. Murakami D., Yamagata Y., Hirano T. Geostatistics and Gaussian process models, *Spatial Analysis Using Big Data*, Academic Press, 2020, P.57–112. DOI:10.1016/b978-0-12-813127-5.00004-7.
5. Kyriakidis P., Journel C. A. Geostatistical space-time models: a review, *Mathematical Geology*, 1999, Vol. 31, pp. 651–684.
6. Rashid T. Make Your Own Neural Network, Create Space Independent Publishing Platform, 2016, 222 p.
7. Subbotin S. O., Olijnyk A. O., Olijnyk O. O. Neiteratyvni, evoljucijni ta mul'tyagentni metody syntezy nechtikologichnyh i nejmomeznyh modelej, Monografija, Zaporizhzhja, ZNTU, 2009, 375 p.
8. Dudin A.N., Nishimura S. Optimal hysteresis control for a BMAP/SM/1/N queue with two operation modes. *Mathematical Problems in the Engineering*, 2000, No. 5, pp. 397–420.
9. Schller J. C. H., Brinkman L., Van Gestel P. J., Van Otterloo R. W. Methods for determining and processing probabilities «Red Book». The Netherlands, Committee for Prevention of Disasters, 1997, 604 p.
10. Tvoroshenko I. S., Kramarenko O. O. Software determination of the optimal route by geoinformation technologies, *Radio Electronics, Computer Science, Control*, 2019, No. 5, pp. 131–142. DOI:10.15588/1607-3274-2019-3-15
11. Ongkowitzo C. S., Doloi H. Risk-based resilience assessment model focusing on urban infrastructure system restoration, *Procedia Engineering*, 2018, Vol. 212, pp. 1115–1122. DOI:10.1016/j.proeng.2018.01.144
12. Popov V. M., Chub I. A., Novozhilova M. V. Modelirovanie harakteristik potoka otkazov osnovnyh proizvodstvennyh fondov ob#ektov povyshennoj opasnosti, *Problemi nadzvichajnih situacij*, 2015, Vol. 21, pp. 93–98.
13. Komarnyc'kyj I. M., Bublyk M. I. Ocinka tehnogennyh zbytkiv ta analiz pidhodiv do i'hn'ogo rozrahunku u global'nomu ta regional'nomu aspektah. [Online]. 2008, Available: vlp.com.ua/files/21_31.pdf.
14. Ishhenko G. G. Analiz ta prognos pryrodno-tehnogennoi' bezpeky velykyh mist iz zastosuvannjam nelinejnyh metodiv. *Ekonomika i derzhava*, 2009, № 8, pp. 48–53.
15. He Y., Du S. Classification of urban emergency based on fuzzy analytic hierarchy process, *Procedia Engineering*, 2016, Vol. 137., pp. 630–638.
16. Gubarevych O. V. Nadijnist' i diagnostyka elektroobladnannya: Pidruchnyk. Sjevjerodonec'k: vyd-vo SNU im. V. Dalja, 2016, 248 p.
17. Pro zatverdzhennia Klasyfikatsiinykh oznak nadzvichajnykh sytuatsii: Nakaz Ministerstva vnutrishnykh sprav Ukrainy vid 06.08.2018r. № 658. [Online]. Available: <https://zakon.rada.gov.ua/laws/show/z0969-18#Text>