



Руслан ПОЛІТАНСЬКИЙ, д-р техн. наук, проф.
ORCID ID: 0000-0003-0015-7123
e-mail: r.politansky@chnu.edu.ua

Чернівецький національний університет імені Юрія Федьковича, Чернівці, Україна

Валентин ЛЕСІНСЬКИЙ, канд. техн. наук, доц.
ORCID ID: 0000-0002-1259-1974
e-mail: v.lesinsky@chnu.edu.ua

Чернівецький національний університет імені Юрія Федьковича, Чернівці, Україна

Сергій ГАЛЮК, канд. техн. наук
ORCID ID: 0000-0003-3836-2675
e-mail: s.haliuk@chnu.edu.ua

Чернівецький національний університет імені Юрія Федьковича, Чернівці, Україна

Андрій КУЛЬКО, асп.
ORCID ID: 0009-0006-1185-0774
e-mail: kulko452@gmail.com

Київський національний університет імені Тараса Шевченка, Київ, Україна

МЕТОДИ ВИЗНАЧЕННЯ РІВНОВАЖНОГО СТАНУ МЕРЕЖІ ІЗ ЗАДАНИМИ ПОТОКАМИ У ВУЗЛАХ ІНФОРМАЦІЙНИХ МЕРЕЖ КІБЕРФІЗИЧНИХ СИСТЕМ

Вступ. Багатовимірність кіберпростору та високий рівень інформатизації об'єктів критичної інфраструктури як в Україні, так і у світі, становлять загрозу для людства на глобальному рівні на досягну перспективу. Ситуація ускладнюється тим, що об'єкти критичної інфраструктури, які функціонують в єдиному інформаційному просторі й підтримують широкий спектр сучасних інформаційних технологій, всупереч колосальним зусиллям для протидії стороннім втручанням із кіберпростору та через кіберпростір, й надалі залишаються вразливими до загроз нового типу.

Методи. Виявлено, що дослідження складних систем використовують такі сучасні методи математичної науки, як топологія, теорія графів, лінійна алгебра.

Результати. Ентропійний метод аналізу складних систем широко застосовують у прикладних задачах дослідження природних, технічних і кіберфізичних систем. Ентропійний метод давно використовують для побудови ефективних систем кодування, шифрування та захисту інформації. У зв'язку із зростанням обчислювальних потужностей збільшилася практична цінність дослідження динамічних систем, зокрема й основаних на ентропії, оскільки ентропія є хорошим показником вибірки деяких станів системи із множини всіх її можливих станів.

Висновки. Оскільки дослідження структур і загальних закономірностей зустрічається в багатьох галузях науки і техніки, актуальними є задачі, що мають загальний характер, коли конкретна природа системи не береться до уваги. У цій роботі наведено алгоритм визначення стану багатовузлової мережі інформаційних потоків, у якому її ентропія набуває максимальне значення. Алгоритм оснований на виведених авторами роботи аналітичних виразах ентропії, та її перших і других похідних, визначених на множині ядрових розв'язків системи лінійних рівнянь, до якої зводиться вся початкова задача. Наведено експериментальні результати визначення множини розв'язків і розв'язання оптимізаційної задачі пошуку стану з максимальним значенням ентропії для мереж із 3-, 4-, 5- та 6-ма вузлами. Проаналізовано можливості подальшого вдосконалення методу для розрахунку складніших мереж.

Ключові слова: програмний комплекс, мережний трафік, матриці, графи, ентропія, матриці Гессе, підматриця.

Вступ

В умовах збройної агресії РФ проти України безпека людини, суспільства і держави суттєво залежить від надійного функціонування об'єктів критичної інфраструктури. Окрім фізичних впливів на такі об'єкти летальною зброєю РФ не полишає спроб разом зі своїми сателітами впливати нелетальними засобами – кіберзброєю – на системи управління об'єктів критичної інфраструктури через кіберпростір і з кіберпростору. З огляду на транскордонність кіберпростору та високий рівень інформатизації об'єктів критичної інфраструктури як в Україні, так і у світі, наприклад, об'єктів ядерної енергетики й об'єктів водопостачання, систем управління військами та зброєю Збройних сил України та її складових тощо, ризики масштабуються не тільки на національний безпековий вимір, а і становлять загрозу для людства на глобальному рівні на досягну перспективу (Beshley et al., 2018).

Ситуація ускладнюється тим, що об'єкти критичної інфраструктури, які функціонують в єдиному інформаційному просторі й підтримують широкий спектр сучасних інформаційних технологій, всупереч колосальним зусиллям для протидії стороннім втручанням із кіберпростору та через кіберпростір, і

надалі залишаються вразливими до загроз нового типу. Дослідження складних систем використовують такі сучасні галузі математичної науки, як топологія, теорія графів, лінійна алгебра. Оскільки дослідження структур і загальних закономірностей відбувається у багатьох галузях науки і техніки, то актуальними є задачі, що мають загальний характер, коли конкретну природу системи не беруть до уваги.

Огляд літератури. Обчислення глобальної ентропії, яка характеризує багатовузлову мережу, має практичне застосування у системах захисту інформації (Garofalakis, Gehrke, & Rastogi, 2016). Коли йдеться про раптову зміну значення ентропії, це може бути ознакою хакерської атаки типу DDoS (Distributed Denial of Service) на мережу (Li, & Sun, 2021). У роботі (Madarro-Capó et al., 2021) наведено алгоритми, що основані на ентропії системи, які дають можливість розв'язати "проблему зменшення комунікаційних витрат у розподілених системах" за рахунок зменшення надмірного використання деяких обчислювальних вузлів системи порівняно з іншими її вузлами. Ентропійний метод є також класичним методом аналізу ефективності систем шифрування інформації (Miletic et al., 2022).

Методи

В роботі (Niven et al., 2019) представлено розроблення програмного комплексу аналізу мережного трафіка та виявлення вторгнень. Дослідженню систем з інтенсивними інформаційними потоками присвячено роботу (Pedrosa, & Costa, 2022), де виведено деякі важливі характеристики системи, що характеризують її загалом: основні методи вибірки та логічного висновку для потоків даних; часті схеми руху (найбільш вживані IP-адреси джерела й адресата); найбільш часті адресати для заданого джерела, і навпаки, тощо. Розглянуто також різні методи дискретизації потоків даних для їх подальшого аналізу (пуассонівські потоки, алгоритми резервуарних потоків Watterman і McLeod), кластеризація потоків даних.

Важливу роль дослідження ентропії набувають у зв'язку з виникненням і поширенням застосуванням динамічних мереж із змінною топологією з'єднання вузлів. Такі мережі виникають у задачах використання безпілотних літальних апаратів для забезпечення функціонування мережі, у технологіях ройового використання безпілотних літальних апаратів, технологіях бездротового зв'язку 5G у безлюдних місцевостях, де економічно необґрунтовано використовувати стаціонарні базові станції, а також у інших ситуаціях, де в інформаційному обміні беруть участь машини без участі людини і де це обумовлено правилами безпеки: місця природних і техногенних

катастроф, космічні місії тощо. Дослідження динамічних станів мережі наведено в роботах (Politansky et al., 2022) на основі теорії мультиграфів, у (Shahar, Alfassi, & Keren, 2022) – на підставі прогнозування зміни станів мережі в бік збільшення її ентропії.

Розглянемо мережу, яку можна задати плоским повнозв'язним графом, із двонаправленими ребрами, без петльових дуг, що зображений на рис. 1 (для мережі з $n = 5$ вузлів). Для цієї мережі задано сумарні значення потоків, які надходять і виходять із кожного вузла (на рис. 1 вони позначені як f_{ij}):

$$\sum_{j=1}^{n-1} f_{ij} = s_i, i \in \{1, 2, \dots, n\},$$

$$\sum_{i=1}^{n-1} f_{ji} = d_j, j \in \{1, 2, \dots, n\},$$
(1)

де n – порядок системи, що збігається з кількістю вузлів.

Вважатимемо, що петльових (самозамкнених) потоків у системі немає, тоді систему можна відобразити матрицею потоків:

$$\|F\|_{n \times n} = \|f_{ij}\|_n = \begin{bmatrix} 0 & \dots & f_{1n-1} \\ \dots & 0 & \dots \\ f_{n,1} & \dots & 0 \end{bmatrix},$$
(2)

де $\|F\|_{n \times n}$ – матриця розмірності $n \times n$.

Отже, усього є $n \cdot (n-1)$ змінних системи. Вважаємо також, що суми всіх вхідних і всіх вихідних потоків збігаються:

$$\sum_{i=1}^n s_i = \sum_{i=1}^n d_i = S_0.$$
(3)

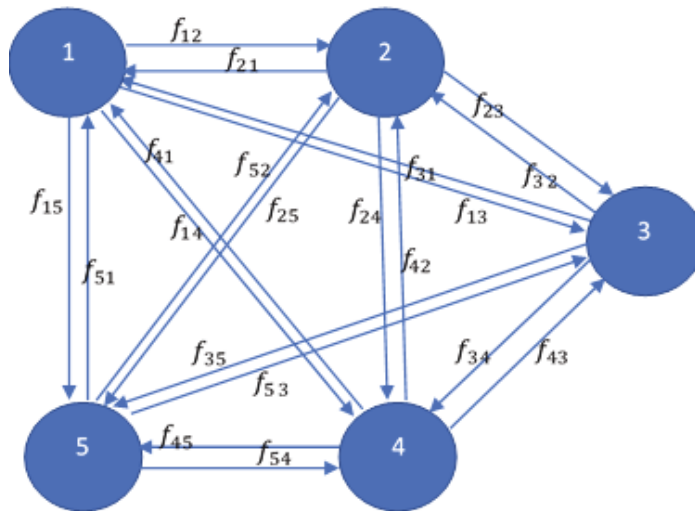


Рис.1. Графічне відображення мережі з п'ятьма вузлами

Тоді рівняння системи, яке враховує умови (1), можна записати у матричному вигляді

$$\|A\|_{(2 \cdot n - 1) \times (n^2 - n)} \times \vec{f}_{n^2 - n} = \vec{v}_{2 \cdot n - 1},$$
(4)

де $\vec{v}$ – стовпець, складений із значень суми вхідних і вихідних потоків для кожного вузла системи, які впорядковані з метою реалізації алгоритму визначення множини розв'язків системи:

$$\vec{v} = (d_2, d_3, \dots, d_n, d_1, s_2, \dots, s_{n-1}, s_n).$$
(5)

$$\vec{f} = (f_{21}, f_{31}, \dots, f_{n1}, f_{11}, f_{32}, \dots, f_{n2}, \dots, f_{n,n-1}, f_{1,n-1}, f_{2,n-1}, \dots, f_{n-1,n-1}).$$
(6)

Таке розміщення змінних і початкових умов задачі обумовлено подальшим спрощенням виконання алгоритму пошуку розв'язків задачі.

Кількість лінійно незалежних рівнянь у системі дорівнює $r = 2 \cdot n - 1$, і є меншою за загальну кількість

Розмірність вектора $\vec{v}$ дорівнює $2 \cdot n - 1$ і збігається з кількістю рівнянь, які забезпечують виконання умов (1), одне рівняння є лінійно залежним від інших, оскільки виконується ще й умова (3). Стовпець змінних системи $\vec{f}$ складений з усіх значень потоків f_{ij} матриці, яка описує систему, які розміщені у порядку проходження стовпців матриці згори донизу, і зліва праворуч:

змінних $n^2 - n$. Тому серед змінних можна виділити $\{\lambda_j\}$, $j \in \{1, \dots, s\}$, ядрових (незалежних змінних):

$$s = n^2 - 3 \cdot n + 1.$$
(7)

Усі інші $\{x_i\}$, $i \in \{1, \dots, r\}$, є лінійними функціями розв'язків $\{\lambda_j\}$.



Визначення множини розв'язків системи для мережі з п'яти вузлів. Визначення множини розв'язків утвореної системи здійснюємо, виконуючи операції віднімання рядків матриці та перестановкою стовпців матриці, внаслідок якої у лівій частині матриці $\|A\|$ буде виділено одиничну діагональну матрицю, розмірність якої $(2 \cdot n - 1) \times (2 \cdot n - 1)$. Зауважимо, що варто дотримуватись такої логіки: після виконання операцій віднімання у деякому стовпці майбутньої одиничної матриці залишиться лише одна одиниця. Перестановку слід виконувати тоді, коли наступний стовпець, який має бути стовпцем одиничної матриці, не містить одиницю в потрібному рядку. Операція віднімання стовпців матриці

не змінює розв'язки усієї системи, а операція перестановки стовпів матриці міняє місцями розв'язки, індекси яких у стовпці $\vec{f}$ збігаються з номерами стовпців матриці, що міняються місцями.

Отже, зручне розміщення змінних системи дає можливість одразу виділити у матриці $\|A\|$ одиничну діагональну підматрицю розмірності $n - 1$ (рис. 1а, де наведено матрицю $\|A\|$ для мережі з $n = 5$ вузлами).

Далі можемо збільшити ранг одиничної матриці без перестановок стовпців матриці. Для цього потрібно послідовно виконати дві операції: відняти від рядка 6 рядок 5 (рис. 1б); відняти від рядка 2 рядок 6 (рис. 1в).

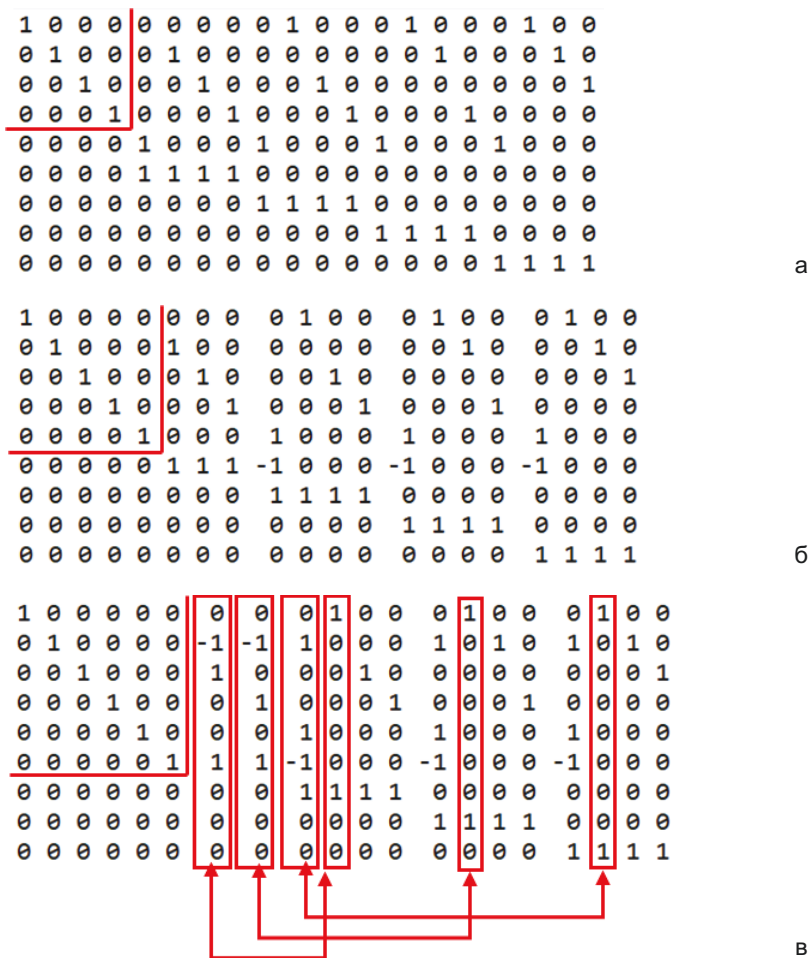


Рис. 1. Початковий етап виділення незалежних розв'язків системи (а та б), позначення перестановок стовпців (в)

Подальше виділення діагональної матриці можна виконувати тільки після перестановок стовпців. Для матриці, зображеної на рис. 1, потрібно виконати перестановку трьох наступних стовпців (7-й, 8-й і 9-й), для цього можна вибрати будь-які три інші стовпці, що містяться ліворуч від них і які мають одиницю в 7-му, 8-му і 9-му рядках відповідно. З погляду збільшення швидкодії алгоритму доцільним є вибір тих стовпців, які мають мінімальну кількість одиниць, щоб потім зменшити кількість операцій віднімання. Такими є стовпці 10-й, 14-й і 18-й (рис. 1в).

На рис. 2а показано вигляд матриці $\|A\|$ після перестановок стовпців. закінчивши перестановки потрібно відняти від першого рядка 7-й, 8-й і 9-й рядки. На рис. 2б бачимо кінцевий результат перетворення.

Отже, матриця, яку ми шукали має вигляд

$$\|R\|_5 = \begin{bmatrix} 0 & -1 & -1 & -1 & 0 & -1 & -1 & -1 & -1 & -1 \\ -1 & 0 & 0 & 1 & -1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \quad (8)$$

Матриця $\|R\|_5$ визначає лінійний зв'язок між 11 ядровими змінними та іншими 9 розв'язками системи.

Позначимо операцію віднімання i -го і j -го рядків як $Su(i, j)$, а їхню перестановку як $P(i, j)$. Тоді



послідовність перетворень для матриці, яка описує мережу з 5 вузлами можна записати у вигляді

$$O_5 = Su(6,5) \rightarrow Su(2,6) \rightarrow P(7,10) \rightarrow P(8,14) \rightarrow P(9,18) \rightarrow Su(1,7) \rightarrow Su(1,8) \rightarrow Su(1,9),$$

або компактніше

$$O_5 = Su[(6,5), (2,6)] \rightarrow P[(7,10), (8,14), (9,18)] \rightarrow Su[(1,7), (1,8), (1,9)]. \quad (9)$$

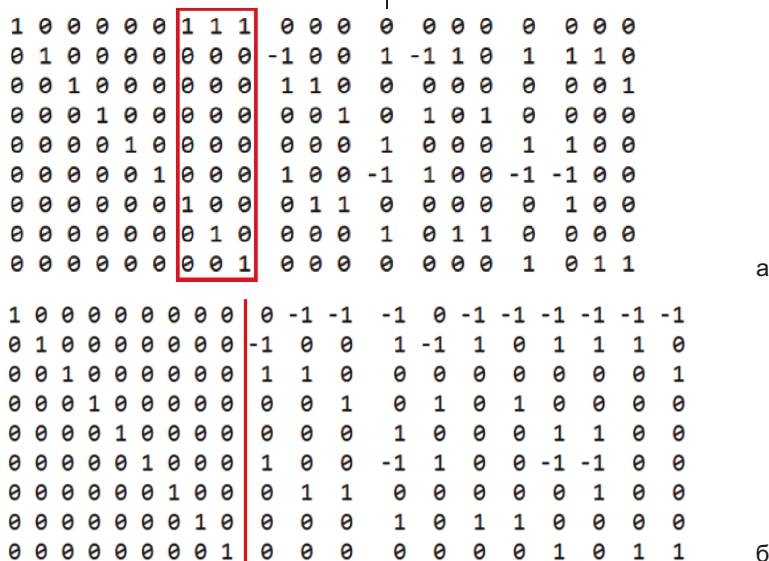


Рис. 2. Перестановка стовпців для виділення діагональної одиничної матриці: матриця після перестановок (а), кінцевий результат перестановок (б)

Разом із відніманням рядків потрібно так само віднімати елементи стовпця $\vec{v}$ у формулі (4), тоді як операції перестановок стовпців матриці $\|A\|$ не змінюють його. Відповідно до розташування елементів стовпця $\vec{v}_5$, операції віднімання змінять його, утворюючи у цьому разі стовпець

$$\vec{b}_5: \vec{v}_5 = \begin{bmatrix} d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{bmatrix} \rightarrow \vec{b}_5 = \begin{bmatrix} d_2 - (s_3 + s_4 + s_5) \\ d_1 + d_3 - s_2 \\ d_4 \\ d_5 \\ d_1 \\ s_2 - d_1 \\ s_3 \\ s_4 \\ s_5 \end{bmatrix} \quad (10)$$

Операція перестановок стовпців змінить вектор змінних системи: $\vec{f} \rightarrow \vec{f}'$, але значення ентропії при цьому не змінюється.

Відтак розв'язки системи утворюють деякий лінійний підпростір 11-вимірного простору незалежних змінних системи $\vec{\lambda}_5^T = (\lambda_1, \lambda_2, \dots, \lambda_{11})$:

$$\vec{f}_5 = - \left\| \frac{R}{E_{11 \times 11}} \right\|_5 \times \vec{\lambda}_5 + \left\| \frac{\vec{b}_5}{\vec{O}_{11}} \right\|, \quad (11)$$

де $E_{11 \times 11}$ – одинична діагональна матриця розмірності 11, $\vec{O}_{11}^T = (0,0,0,0,0,0,0,0,0,0,0)$ – вектор, утворений із 11 елементів, що дорівнюють нулю.

Очевидним обмеженням на розв'язки задачі є те, що потоки повинні мати невід'ємні значення: $f_{ji} \geq 0$. Це накладає також обмеження на значення незалежних змінних системи: $\lambda_j \geq 0$. Виникають також інші, нетривіальні обмеження значень $\{\lambda_j\}$, які забезпечують невід'ємність інших розв'язків системи $\{x_i\}$:

$$\begin{cases} \sum_{j=1}^{11} \|r_{ij}\| \cdot \lambda_j \leq \vec{b}_5(i), i \in \{1, \dots, 9\} \\ \lambda_j \geq 0, j \in \{1, \dots, 11\} \end{cases} \quad (12)$$

Алгоритм визначення розв'язків системи для довільної кількості вузлів. Узагальнимо розглянутий метод для системи довільного порядку.

В результаті перших двох операцій віднімання рядків виділяють одиничну діагональну матрицю розмірності $n+1$ (6 для матриці, що зображена на рис. 1). Вказані операції є однотипні для мереж із будь-якою кількістю вузлів. Причому номери рядків, що віднімаються на першому кроці дорівнюють $n+1$ та n , а на другому кроці – 2 та $n+1$. Отже, для мережі довільного порядку ці дві операції можна зобразити у символічному вигляді як $Su[(n+1, n), (2, n+1)]$. Очевидно, що кількість операцій перестановок стовпців становить $2 \cdot n - 1 - (n+1) = n - 2$. Зауважимо, що номери стовпців, які переставляємо ліворуч, послідовно збільшуються, починаючи із стовпця з номером $n+2$. Стовпці, які переставляємо праворуч, можуть бути довільними, але для зручності реалізації алгоритму вибираємо ті, що містять тільки дві одиниці: в першому рядку i в тому рядку, який містить наступну одиницю діагональної матриці. Визначити номери таких стовпців можна, якщо проаналізувати структуру початкової матриці. Це стовпці, номери яких утворюють послідовність $\{2 \cdot n + i \cdot (n-1)\}$, де i – послідовний номер стовпця, який переставляємо, $i \in \{1, 2, \dots, n-2\}$. Відтак для мережі довільного порядку символічне позначення операцій перестановок можна позначити, як $\prod_{i=1}^{n-2} P(n+i+1, 2 \cdot n + i \cdot (n-1))$, де символом $\prod_{i=1}^{n-2}$ позначено послідовне виконання операцій із змінним індексом $i \in \{1, 2, \dots, n-2\}$. Після операцій перестановок потрібно виконувати віднімання рядка з номером стовпця, що переставляється від першого рядка табл. 1 (рис. 2б, відповідні стовпці виділено червоним кольором), у символічному позначенні операції віднімання можна позначити як



$\prod_{i=1}^{n-2} S(1, n + i + 1)$. Отже узагальнена формула виконання операцій алгоритму має вигляд

$$S[(n + 1, n), (2, n + 1)] \rightarrow \prod_{i=1}^{n-2} P[n + i + 1, 2n + i(n - 1)] \rightarrow \prod_{i=1}^{n-2} S(1, n + i + 1). \quad (13)$$

Таким способом початкова матриця $\|A\|$ розділиться на дві частини: ліворуч утворюється одинична діагональна матриця, а праворуч – залишкова матриця

$$\|A\|_{(2 \cdot n - 1) \times n \cdot (n - 1)} \rightarrow \|E\|_{(2 \cdot n - 1) \times (2 \cdot n - 1)} \|R\|_{(2 \cdot n - 1) \times (n^2 - 3 \cdot n + 1)}. \quad (14)$$

Розмірність матриці $\|R\|$ визначається кількістю незалежних рівнянь системи і кількістю незалежних змінних:

$$\dim(\|R\|) = (2 \cdot n - 1) \times s. \quad (15)$$

Елементи матриці $\|R\|$ можуть набувати лише три значення:

$$r_{ij} \in \{-1, 0, 1\}, i \in \{1, 2, \dots, 2 \cdot n - 1\}, j \in \{1, 2, \dots, s\}.$$

Враховуючи тільки операції віднімання, із формули (13) визначимо коефіцієнт вільних доданків $\vec{b}_n$:

$$\begin{cases} b_i = v_i, i \neq \{1, 2, n + 1\} \\ b_1 = v_1 - \sum_{i=1}^{n-2} v_{i+2} \\ b_2 = v_2 - (v_{n+1} - v_n) \\ b_{n+1} = v_{n+1} - v_n \end{cases} \quad (17)$$

$$\vec{f}_n = - \left\| \frac{\|R\|_{(2 \cdot n - 1) \times (n^2 - 3 \cdot n + 1)}}{E_{(n^2 - 3 \cdot n + 1) \times (n^2 - 3 \cdot n + 1)}} \right\|_n \times \vec{\lambda}_{n^2 - 3 \cdot n + 1} + \left\| \frac{\vec{b}_n}{\vec{0}_{n^2 - 3 \cdot n + 1}} \right\|. \quad (19)$$

Значення незалежних змінних $(\lambda_1, \lambda_2, \dots, \lambda_s)$ повинні забезпечувати виконання умови невід'ємності потоків усіх потоків $\vec{f}_n$ у мережі. Із цього слідує обмеження на розв'язки системи $\{\lambda_j\}$, викладені у вигляді нерівностей:

$$\begin{cases} \sum_{j=1}^s r_{ij} \cdot \lambda_j \leq b_i \text{ for } i \in \{1, 2, \dots, 2 \cdot n - 1\} \\ \lambda_j \geq 0 \text{ for } j \in \{1, 2, \dots, n^2 - 3 \cdot n + 1\} \end{cases} \quad (20)$$

Визначення оптимального стану системи ентропійним методом. Маючи аналітичний вигляд розв'язків системи, можемо поставити задачу визначення оптимального значення ентропії. Ентропія мережі з інформаційними потоками може бути визначена на основі нормованого значення потоків, сума яких для всієї матриці потоків рівна 1.

$$f_{ij}^{norm} = f_{ij} / S_0, \quad (21)$$

де f_{ij} – значення потоку між вузлами i та j , f_{ij}^{norm} – нормоване значення потоку між вузлами i та j , S_0 – сума всіх потоків у мережі: $S_0 = \sum_{i=1}^n \sum_{j=1}^{n-1} f_{ij}$, тут n

$$x_i = - \sum_{k=1}^s r_{ik} \cdot \lambda_k + b_i, i \in \{1, 2, \dots, 2 \cdot n - 1\}. \quad (25)$$

Тоді вираз для ентропії матиме вигляд

$$E = - \frac{1}{S_0} \cdot \sum_{i=1}^{2 \cdot n - 1} x_i \cdot \log_2 \left(\frac{x_i}{S_0} \right) - \frac{1}{S_0} \cdot \sum_{i=1}^s \lambda_s \cdot \log_2 \left(\frac{\lambda_s}{S_0} \right) = E_1 + E_2. \quad (26)$$

З огляду на вище наведені формули, можемо зробити висновок, що ентропія є функцією незалежних (ядрових) змінних системи:

$$E = E(\lambda_1, \lambda_2, \dots, \lambda_{s-1}, \lambda_s). \quad (27)$$

Далі шукатимемо стан мережі, в якому ентропія набуває максимальне значення з урахуванням обмежень (20). Оптимізаційна задача пошуку рівноважного стану всієї мережі може бути сформульована так:

$$\frac{\partial E_1}{\partial \lambda_j} = \frac{\partial}{\partial \lambda_j} \left[- \frac{1}{S_0} \cdot \sum_{i=1}^{2 \cdot n - 1} x_i \cdot \log_2 \left(\frac{x_i}{S_0} \right) \right] = - \frac{1}{S_0} \cdot \sum_{i=1}^{2 \cdot n - 1} \frac{\partial}{\partial \lambda_j} \left[x_i \cdot \log_2 \left(\frac{x_i}{S_0} \right) \right], \quad (29)$$

$\|R\|$, яка є основою для подальшого визначення розв'язків системи:

Підставляючи значення $\{v_i\}$, знаходимо значення $\{b_i\}$, виражені через початкові умови задачі:

$$\begin{cases} b_1 = d_2 - \sum_{i=1}^{n-2} s_{i+2} \\ b_2 = d_1 + d_3 - s_2 \\ b_i = d_{i+1}, i \in \{3, 4, \dots, n - 1\} \\ b_n = d_1 \\ b_{n+1} = s_2 - d_1 \\ b_i = s_i, i \in \{3, 4, \dots, n\} \end{cases} \quad (18)$$

Множину розв'язків усієї системи можна записати у вигляді матричного рівняння:

– кількість вузлів у мережі, тоді виконується стандартна умова нормування:

$$\sum_{i=1}^n \sum_{j=1}^{n-1} f_{ij}^{norm} = 1, \quad (22)$$

яка виконується, коли ентропія визначається за ймовірностями станів системи. Тому можливе використання аналогії між інтенсивністю потоків між деякими вузлами та ймовірністю стану системи, в якому реалізується це значення.

Далі використовуємо логарифмічну міру для визначення ентропії (за аналогією з ентропією системи, яка визначається за ймовірністю її станів):

$$E = - \sum_{i=1}^n \sum_{j=1}^{n-1} \frac{f_{ij}}{S_0} \cdot \log_2 \left(\frac{f_{ij}}{S_0} \right). \quad (23)$$

Для подальшого аналізу змінимо послідовну нумерацію множини розв'язків системи, виділивши окремо незалежні змінні $(\lambda_1, \lambda_2, \dots, \lambda_s)$:

$$\vec{f}_n = (x_1, x_2, \dots, x_{2 \cdot n - 1}, \lambda_1, \lambda_2, \dots, \lambda_s), \quad (24)$$

де розв'язки $\{x_i\}$ можуть бути визначені з матричного рівняння у вигляді лінійної комбінації незалежних змінних:

$$\begin{aligned} \max_{\lambda_1, \lambda_2, \dots, \lambda_{s-1}, \lambda_s} E &= E(\lambda_1, \lambda_2, \dots, \lambda_{s-1}, \lambda_s) \\ \text{subject to } \sum_{i=1}^s r_{ji} \cdot \lambda_i &\leq b_j \text{ for } j \in \{1, 2, \dots, 2 \cdot n - 1\}. \end{aligned} \quad (28)$$

Відтак для визначення оптимального стану мережі визначимо частинні похідні ентропії $E(\lambda_1, \lambda_2, \dots, \lambda_{s-1}, \lambda_s)$ за кожною незалежною змінною $\lambda_i \in \{\lambda_1, \lambda_2, \dots, \lambda_{s-1}, \lambda_s\}$. Для зручності визначатимемо похідні ентропії від $E_1(\lambda_1, \lambda_2, \dots, \lambda_{s-1}, \lambda_s)$ та $E_2(\lambda_1, \lambda_2, \dots, \lambda_{s-1}, \lambda_s)$ окремо (26):



$$\frac{\partial E_2}{\partial \lambda_j} = \frac{\partial}{\partial \lambda_j} \left[-\frac{1}{s_0} \cdot \sum_{i=1}^s \lambda_i \cdot \log_2 \left(\frac{\lambda_i}{s_0} \right) \right] = -\frac{1}{s_0} \cdot \sum_{i=1}^s \frac{\partial}{\partial \lambda_j} \left[\lambda_i \cdot \log_2 \left(\frac{\lambda_i}{s_0} \right) \right]. \quad (30)$$

Знайдемо похідні від кожного доданка (29) і (30):

$$\frac{\partial}{\partial \lambda_j} \left[x_i \cdot \log_2 \left(\frac{x_i}{s_0} \right) \right] = -r_{ij} \cdot \left[\frac{1}{\ln 2} + \log_2 \left(\frac{x_i}{s_0} \right) \right], \quad (31)$$

$$\frac{\partial}{\partial \lambda_j} \left[\lambda_i \cdot \log_2 \left(\frac{\lambda_i}{s_0} \right) \right] = \delta_{ij} \cdot \left[\frac{1}{\ln 2} + \log_2 \left(\frac{\lambda_i}{s_0} \right) \right], \quad (32)$$

де δ_{ij} символ Кронекера: $\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ for $i \neq j$.

$$\frac{\partial E}{\partial \lambda_j} = -\frac{\left[\frac{1}{\ln 2} + \log_2 \left(\frac{\lambda_j}{s_0} \right) \right]}{s_0} + \frac{1}{s_0} \cdot \sum_{i=1}^{2n-1} r_{ij} \cdot \left[\frac{1}{\ln 2} + \log_2 \left(\frac{-\sum_{k=1}^s r_{ik} \cdot \lambda_k + b_i}{s_0} \right) \right]. \quad (33)$$

Так само знаходимо похідні другого порядку $\frac{\partial^2 E}{\partial \lambda_i \partial \lambda_j}$:

$$\frac{\partial^2 E}{\partial \lambda_i \partial \lambda_j} = -\frac{1}{\ln 2 \cdot s_0} \cdot \left\{ \frac{1}{\lambda_j} \cdot \delta_{ij} + \sum_{i=1}^{2n-1} \frac{r_{ij} \cdot r_{il}}{b_i - \sum_{k=1}^s r_{ik} \cdot \lambda_k} \right\}, \quad (34)$$

або

$$\frac{\partial^2 E}{\partial \lambda_i \partial \lambda_j} = -\frac{1}{\ln 2 \cdot s_0} \cdot \left\{ \frac{1}{\lambda_j} \cdot \delta_{ij} + \sum_{i=1}^{2n-1} \frac{r_{ij} \cdot r_{il}}{x_i} \right\}. \quad (35)$$

Результати

Визначення множини цілочисельних розв'язків системи. Задача визначення всієї множини розв'язків системи є дуже складною з погляду швидкодії та потужності обчислювальних ресурсів, які необхідні для її розв'язання. Алгоритм цієї задачі ґрунтується на

циклічній перевірці умов (1) на значення незалежних змінних, і в разі їх генерування цілочисельних розв'язків задачі. Але у зв'язку з обчислювальною складністю, крок визначення значень незалежних параметрів варіювався від 2 до 5 залежно від порядку системи. У табл. 1 наведено результати обчислень.

Таблиця 1

Результати визначення підмножин розв'язків системи із $n = 3, 4, 5$ і 6 вузлів

Порядок системи, початкові умови задачі	Крок вибірки незалежних параметрів	Кількість розв'язків	Середнє значення ентропії	Максимальне значення ентропії
$n = 3, S_0 = 1000$	1	1499	1,21	2,42
$n = 4, S_0 = 1000$	20	1125	1,7	3,40
	10	33469	1,75	3,41
$n = 5, S_0 = 1130$	50	3517	1,99	3,97
$n = 6, S_0 = 1580$	80	899	2,22	4,44

Очевидно, що пошук розв'язків методом циклічної перевірки умов (1) значно обмежує можливість його застосування для задач високої розмірності ($n > 7$). Тому залишається відкритим питання про можливість ефективного пошуку хоча б одного розв'язку задачі для застосування оптимізаційних алгоритмів для розв'язання задач більш розмірності ($n > 7$).

Визначення оптимального стану з максимальним значенням ентропії на основі аналітичних виразів. Оскільки метод визначення всієї множини розв'язків потребує значних обчислювальних ресурсів, або значних витрат машинного часу, то розв'язки

визначали з деяким кроком дискретизації у просторі $\{\lambda_j\}$. Тому оптимальний розв'язок, для якого реалізується глобальний максимум ентропії, у більшості випадків пропущено, хоча стан системи є наближеним до рівноважного.

Для точного визначення глобального максимуму ентропії використовуємо властивість унімодальності цієї функції. Наведемо без доведення дві лема, які дозволяють зробити висновок про те, що ентропія є унімодальною (Shahar, Alfassi, & Keren, 2022).

Лема 1. Сума унімодальних функцій вигляду

$$f_j(\lambda_j) = -\sum_{i=1}^r \frac{(b_{ij} - r_{ij} \cdot \lambda_j)}{s_0} \cdot \log_2 \left(\frac{b_{ij} - r_{ij} \cdot \lambda_j}{s_0} \right) - \frac{\lambda_j}{s_0} \cdot \log_2 \left(\frac{\lambda_j}{s_0} \right),$$

$$\tilde{b}_{ij} = b_i - \sum_{k=1, k \neq j}^s r_{ik} \cdot \lambda_k, j \in \{1, \dots, s\}$$

є унімодальною функцією змінної λ_j .

Лема 2. Багатовимірна функція $f(\lambda_1, \dots, \lambda_s)$ є унімодальною функцією.

Для знаходження оптимального стану мережі можемо далі застосувати чисельні методи, а саме метод другого порядку – метод Ньютона:

$$\tilde{\lambda}^{(k+1)} = \tilde{\lambda}^{(k)} - (\|H\|^{(k)})^{-1} \times \nabla E^{(k)}, \quad (36)$$

де $\|H\|$ – матриця Гессе (гесіан), що складена з похідних другого порядку (33):

$$\|H\| = \left\| \frac{\partial^2 E}{\partial \lambda_i \partial \lambda_j} \right\|. \quad (37)$$

Одним із критеріїв збіжності методу є визначення норми градієнта на кожному кроці ітерації. Для цього використано евклідову міру відстані і довжини вектора у багатовимірному просторі:

$$|\nabla E| = \sqrt{\sum_{j=1}^s \left(\frac{\partial E}{\partial \lambda_j} \right)^2}. \quad (38)$$

Для найменш складної системи $n = 3$ на рис. 3 зображено графіки ентропії, першої і другої похідних, обчислених на множині значень єдиного параметра λ для системи із $S_0 = 100$.

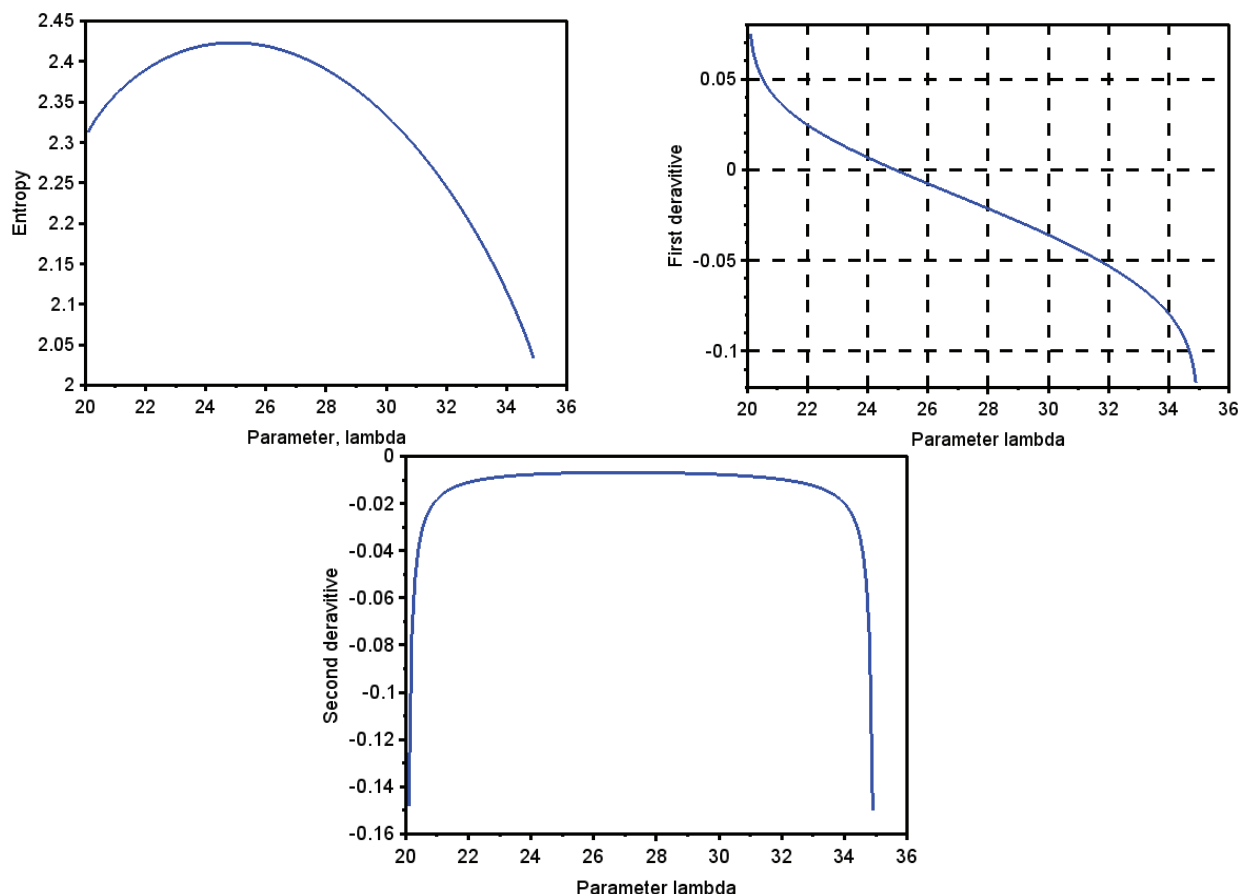


Рис. 3. Ентропія $E(\lambda)$ та похідна $\frac{\partial E}{\partial \lambda}$ для системи $n = 3$

Чисельний метод показує дуже хорошу збіжність: для визначення оптимального значення для всіх досліджених порядків системи ($n = 3, 4, 5, 6$) достатньо всього 5 ітерацій методу. Результати ітераційного

процесу зображено на рис. 4а, де показано обчислену ентропію, і на рис. 4б, де показано норму градієнта на кожному кроці ітерації.

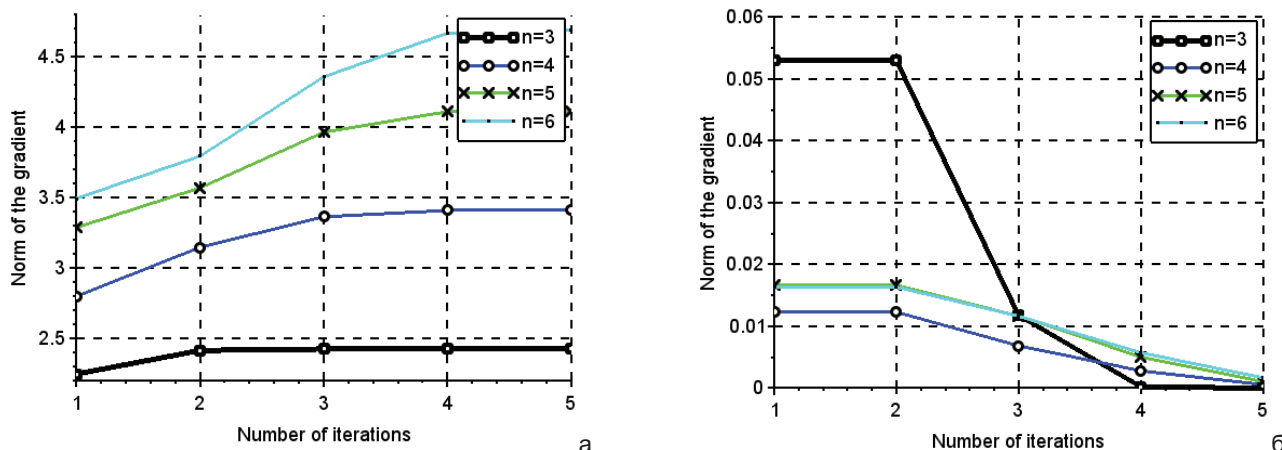


Рис. 4. Збіжність методу для системи дослідженої розмірності ($n = 3, 4, 5, 6$): а – ентропія; б – норма градієнта

Зауважимо, що залежно від вибору початкової точки алгоритм, який оснований на застосуванні методу Ньютона, може визначити кілька рівноправних розв'язків, для яких ентропія мережі має однакові значення. Це можна зрозуміти, якщо зважити на те, що абсолютне значення градієнта має досить малі

значення. Правильність такого результату підтверджують статистичні розподіли ентропії, які одержані нами на основі множини розв'язків системи (табл. 1). Для мереж із 3, 4, 5 і 6 вузлами на рис. 5 показано гістограми значень ентропії, розділені на 10 інтервалів.

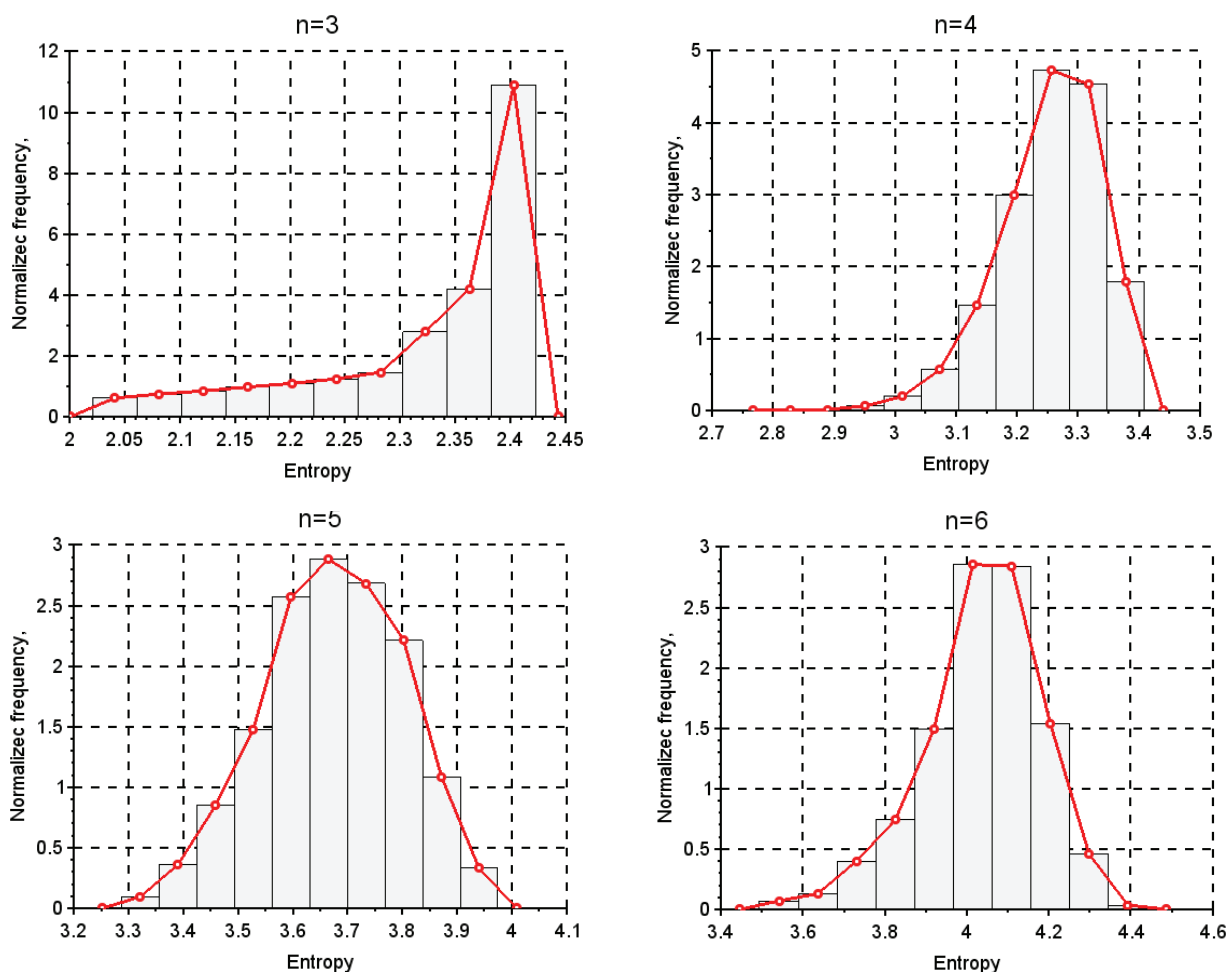


Рис. 5. Гістограми розподілу ентропії для мережі з 3, 4, 5 і 6 вузлами, побудовані на основі множини цілочисельних розв'язків із поділом на 10 інтервалів (табл.1)

Розв'язок оптимізаційної задачі, що оснований на аналітичних виразах (29) і (33) має очевидні обмеження: $\lambda_j \neq 0, j \in \{1, 2, \dots, s\}$ і $x_i \neq 0, i \in \{1, 2, \dots, 2n-1\}$. Але в загальній постановці задачі (28) нульові розв'язки є допустимими. Тому для більш строгого пошуку розв'язку задачі варто окремо проаналізувати значення ентропії у точках, що є вершинами опуклої множини у s -вимірному просторі незалежних змінних $\{\lambda_j\}$, де $\lambda_{t1} = 0$ для деяких $t_1 \in \{1, 2, \dots, s\}$ або $x_{t2} = 0$ для деяких $t_2 \in \{1, 2, \dots, 2n-1\}$.

Дискусія і висновки

Внаслідок проведених досліджень розв'язано такі задачі:

- розроблено алгоритм визначення множини цілочисельних розв'язків задачі (4), які відповідають умовам (1) і (3);
- одержано аналітичні вирази для ентропії, як функції змінних $\{\lambda_j\}$;
- одержано аналітичні вирази для градієнта ентропії, визначеного у просторі ядрових змінних $\{\lambda_j\}$;
- одержано аналітичні вирази для всіх других похідних від ентропії у просторі $\{\lambda_j\}$, які утворюють матрицю Гессе;
- на основі градієнта ентропії і її матриці Гессе розроблено алгоритм визначення стану мережі, в якому її ентропія набуває максимальне значення (Politanskyi et al., 2022).

Перерахуємо переваги розробленого методу. По-перше, це є швидкодія визначення залишкової матриці. Для матриці з $n = 10$ вузлів, матриця $\|R\|$ розмірністю 19×71 обчислюється за допомогою звичайного персонального комп'ютера за час, що не перевищує 0,5 с. По-друге, виконується хороша збіжність ітераційного процесу, яка для досліджених розмірностей ($n = 3, 4, 5, 6$) практично не залежить від порядку системи, і тому цей процес не потребує значних витрат обчислювальних потужностей.

Очевидною проблемою застосування аналітичного методу є вибір початкового розв'язку задачі ($\vec{\lambda}^{(0)}$ у (36)). Для цього використовуємо алгоритм визначення цілочисельних розв'язків системи, описаний вище. Внаслідок значної кількості вкладених циклів перевірки умов задачі (28), це досить складна задача, яка потребує значних витрат машинного часу навіть для пошуку єдиного розв'язку задачі. Тому визначення початкового розв'язку для оптимізаційної задачі має вирішальне значення для застосування алгоритму у системах, розмірність яких перевищує 7 вузлів.

Внесок авторів: Руслан Політанський – концептуалізація; методологія; Валентин Лесінський – аналіз джерел, підготування огляду літератури; Сергій Галюк – аналіз теоретичних засад дослідження; Андрій Кулько – збір емпіричних даних та їх валідація; емпіричне дослідження.



Джерела фінансування. Це дослідження не отримало жодного гранту від фінансової установи в державному, комерційному або некомерційному секторах.

Список використаних джерел

- Beshley, M., Toliupa, S., Pashkevych, V., & Kolodiy, R. (2018). Development of software system for network traffic analysis and intrusion detection. *2018 International Conference on Information and Telecommunication Technologies and Radio Electronics, UkrMiCo 2018 – Proceedings*. Igor Sikorsky Kyiv Polytechnic Institute. <https://doi.org/10.1109/UkrMiCo43733.2018.9047546>
- Garofalakis, M. N., Gehrke, J., & Rastogi, R. (Eds.). (2016). *Data stream management – Processing high-speed data streams. Data-Centric Systems and Applications*. Springer. <https://doi.org/10.1007/978-3-540-28608-0>
- Li, X., & Sun, Q. (2021). Identifying and ranking influential nodes in complex networks based on dynamic node strength. *Algorithms*, 14(3), 82. <https://doi.org/10.3390/a14030082>
- Madarro-Capó, E. J., Legón-Pérez, C. M., Rojas, O., & Sosa-Gómez, G. (2021). Information theory based evaluation of the RC4 stream cipher outputs. *Entropy*, 23(7), 896. <https://doi.org/10.3390/e23070896>
- Miletic, S., Pokrajac, I., Pena-Pena, K., Arce, G. R., & Mladenovic, V. A. (2022). Multigraph-defined distribution function in a simulation model of a communication network. *Entropy*, 24(9), 1294. <https://doi.org/10.3390/e24091294>
- Niven, R. K., Abel, M., Schlegel, M., & Waldrip, S. H. (2019). Maximum entropy analysis of flow networks: Theoretical foundation and applications. *Entropy*, 21(8), 776. <https://doi.org/10.3390/e21080776>
- Pedrosa, V. G., & Costa, M. H. M. (2022). Index coding with multiple interpretations. *Entropy*, 24(8), 1149. <https://doi.org/10.3390/e24081149>
- Politanskyi, R. L., Bobalo, Y. Y., Zarytska, O. L., Kiselychuk, M. D., & Vistak, M. V. (2022). Entropy calculation for networks with determined values of flows in nodes. *Mathematical Modelling and Computing*, 9(4), 936–944. <https://doi.org/10.23939/mmc2022.04.936>
- Shahar, A., Alfassi, Y., & Keren, D. (2022). Communication efficient algorithms for bounding and approximating the empirical entropy in distributed systems. *Entropy*, 24(11), 1611. <https://doi.org/10.3390/e24111611>

References

- Beshley, M., Toliupa, S., Pashkevych, V., & Kolodiy, R. (2018). Development of software system for network traffic analysis and intrusion detection. *2018 International Conference on Information and Telecommunication Technologies and Radio Electronics, UkrMiCo 2018 – Proceedings*. Igor Sikorsky Kyiv Polytechnic Institute. <https://doi.org/10.1109/UkrMiCo43733.2018.9047546>
- Garofalakis, M. N., Gehrke, J., & Rastogi, R. (Eds.). (2016). *Data stream management – Processing high-speed data streams. Data-Centric Systems and Applications*. Springer. <https://doi.org/10.1007/978-3-540-28608-0>
- Li, X., & Sun, Q. (2021). Identifying and ranking influential nodes in complex networks based on dynamic node strength. *Algorithms*, 14(3), 82. <https://doi.org/10.3390/a14030082>
- Madarro-Capó, E. J., Legón-Pérez, C. M., Rojas, O., & Sosa-Gómez, G. (2021). Information theory based evaluation of the RC4 stream cipher outputs. *Entropy*, 23(7), 896. <https://doi.org/10.3390/e23070896>
- Miletic, S., Pokrajac, I., Pena-Pena, K., Arce, G. R., & Mladenovic, V. A. (2022). Multigraph-defined distribution function in a simulation model of a communication network. *Entropy*, 24(9), 1294. <https://doi.org/10.3390/e24091294>
- Niven, R. K., Abel, M., Schlegel, M., & Waldrip, S. H. (2019). Maximum entropy analysis of flow networks: Theoretical foundation and applications. *Entropy*, 21(8), 776. <https://doi.org/10.3390/e21080776>
- Pedrosa, V. G., & Costa, M. H. M. (2022). Index coding with multiple interpretations. *Entropy*, 24(8), 1149. <https://doi.org/10.3390/e24081149>
- Politanskyi, R. L., Bobalo, Y. Y., Zarytska, O. L., Kiselychuk, M. D., & Vistak, M. V. (2022). Entropy calculation for networks with determined values of flows in nodes. *Mathematical Modelling and Computing*, 9(4), 936–944. <https://doi.org/10.23939/mmc2022.04.936>
- Shahar, A., Alfassi, Y., & Keren, D. (2022). Communication efficient algorithms for bounding and approximating the empirical entropy in distributed systems. *Entropy*, 24(11), 1611. <https://doi.org/10.3390/e24111611>

Отримано редакцією журналу / Received: 26.03.25
Прорецензовано / Revised: 28.04.25
Схвалено до друку / Accepted: 22.05.25

Ruslan POLITANSKYI, DSc (Engin.), Prof.
ORCID ID: 0000-0003-0015-7123
e-mail: r.politansky@chnu.edu.ua
Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine

Valentyn LESINSKYI, PhD (Engin.), Assoc. Prof.
ORCID ID: 0000-0002-1259-1974
e-mail: v.lesinsky@chnu.edu.ua
Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine

Serhii HALIUK, PhD (Engin.)
ORCID ID: 0000-0003-3836-2675
e-mail: s.haliuk@chnu.edu.ua
Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine

Andrii KULKO, PhD Student
ORCID ID: 0009-0006-1185-0774
e-mail: kulko452@gmail.com
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

METHODS FOR DETERMINING THE EQUILIBRIUM STATE OF A NETWORK WITH GIVEN FLOWS IN NODES OF INFORMATION NETWORKS OF CYBER-PHYSICAL SYSTEMS

Background. The multidimensionality of cyberspace and the high level of informatization of critical infrastructure objects, both in Ukraine and globally, pose a threat to humanity at the global level in the foreseeable future. The situation is complicated by the fact that critical infrastructure objects, which operate within a unified information space and support a wide range of modern information technologies, remain vulnerable to new types of threats, despite tremendous efforts to counteract external cyber intrusions.

Methods. It has been identified that the study of complex systems employs modern mathematical methods such as topology, graph theory, and linear algebra.

Results. The entropic method for analyzing complex systems has found wide application in practical problems related to the study of natural, technical, and cyber-physical systems. This method has long been used to develop efficient systems for coding, encryption, and information protection. With the increase in computational power, the practical value of researching dynamic systems – including those based on entropy – has grown, as entropy serves as a good indicator for selecting certain states of a system from the set of all its possible states.

Conclusions. Since the study of structures and general patterns is common across many branches of science and technology, tasks of a general nature – where the specific nature of the system is not taken into account – remain relevant. This study presents an algorithm for determining the state of a multi-node information flow network in which its entropy reaches a maximum value. The algorithm is based on analytical expressions for entropy and its first and second derivatives, derived by the authors and defined on the set of kernel solutions of a system of linear equations to which the original problem is reduced. Experimental results are presented for determining the solution set and solving the optimization problem of finding the state with the maximum entropy value for networks with 3, 4, 5, and 6 nodes. Possibilities for further improvement of the method for calculating more complex networks are analyzed.

Keywords: software complex, network traffic, matrices, graphs, entropy, Hessian matrices, submatrix.

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.