

В. Б. Мокін¹
Є. М. Крижановський¹
А. М. Лучко¹
Б. С. Білецький¹
С. О. Жуков¹

МЕТОД ОПТИМІЗАЦІЇ ІНФОРМАЦІЙНИХ МОДЕЛЕЙ МАСШТАБОВАНИХ У ПРОСТОРІ АНАЛІТИЧНИХ ВЕБ-СИСТЕМ ЗА КРИТЕРІЄМ ПОВНОТИ ЇХНЬОЇ ТОПОЛОГІЧНОЇ СПОСТЕРЕЖУВАНОСТІ

¹Вінницький національний технічний університет

Розглянуто питання оптимізації інформаційних моделей масштабованих у просторі аналітичних веб-систем за критерієм повноти їхньої топологічної спостережуваності.

Охарактеризовано створену раніше, за участі частини співавторів, інформаційну технологію аналізу та оптимізації топологічної спостережуваності багатозв'язних геоінформаційних систем, тобто — оптимізації моделі системи у вигляді біхроматичного графа, яка дозволяє за довірливих вхідних змінних отримати значення змінних стану чи поточних параметрів режиму системи або забезпечує достатність інформації про поточний стан для синтезу закону керування ним. У такому біхроматичному графі одні вершини є змінними, а інші — залежностями між ними. У цій технології введено критерій спостережуваності, який може приймати значення від 0 до 1 і порівнювати різні моделі за рівнем їхньої спостережуваності та оптимізувати їх, вибираючи найспостережуваніший варіант. Запропоновано яким чином таку формалізацію моделі і цю технологію можна поширити на формалізацію масштабованих у просторі аналітичних веб-систем. А тоді запропоновано новий метод оптимізації інформаційних моделей таких систем за критерієм повноти їхньої топологічної спостережуваності, який дозволяє здійснити аналіз повноти цієї спостережуваності як на кожному рівні масштабування у просторі, на яких планується виконувати аналітичні операції у проєктованій веб-системі, так і для всієї системи загалом. Запропоновано оптимізувати модель, за рахунок введення додаткових вершин-змінних і вершин-залежностей між різними рівнями. Тобто, за відсутності даних чи залежностей на певному рівні системи, іноді можна їх отримати з даних чи залежностей іншого рівня, і — навпаки, що підвищить спостережуваність системи в цілому. Метод дозволяє аналізувати рівень топологічної спостережуваності і аналізувати його вже після оптимізації. За результатами такого аналізу з'являється можливість введення програмних обмежень у функціональності веб-системи, щоб уникнути появи запитів чи залежностей, не забезпечених даними чи залежностями, з яких їх можна було б отримати, через неповну топологічну спостережуваність окремих (чи всіх) рівнів системи.

У роботі продемонстровано роботу методу на прикладі розроблення автоматизованої веб-системи розрахунку і прогнозування водогосподарського балансу басейну Дністра. Застосування розробленого підходу на етапі проєктування цієї системи забезпечило уникнення появи запитів чи залежностей, не забезпечених даними. Систему розроблено та успішно впроваджено.

Розроблений метод оптимізації інформаційних моделей масштабованих у просторі аналітичних веб-систем за критерієм повноти їхньої топологічної спостережуваності може бути застосований для аналізу та проєктування систем, які відображають дані навколишнього середовища чи інфраструктури, розподілені у просторі та мають кілька рівнів деталізації. Це можуть бути кадастрові системи, системи моніторингу стану довкілля тощо.

Ключові слова: аналітична веб-система, топологічна спостережуваність, інформаційна модель, масштабованість просторових даних, проєктування веб-систем.

Постановка задачі та вихідні передумови

До сучасних аналітичних веб-систем ставиться все більше вимог [1]—[3], одними з яких є необхідність масштабування у просторі. Наприклад, усім зручно користуватись сервісом Google Maps і

отримувати інформацію, зокрема аналітичну, спочатку на рівні континентів, потім — країн, регіонів, міст, вулиць (і навіть перегляд вулиць), організацій. Аналогічні вимоги ставляться й до інших веб-систем, які забезпечують аналітику про просторові об'єкти [4], [5].

Для відносно простих систем з малою кількістю даних, теоретично, можна передбачити усі можливі запити на них та запрограмувати елементи інтерфейсу для їхнього обслуговування. Але для складних систем, де є чимало проміжних значень, що визначаються багатьма способами, забезпечення обслуговування довільного запиту перетворюється на нетривіальну задачу. Задача ускладнюється значно, коли система проєктується декількома командами і постійно розвивається. За таких умов, високою є ймовірність того, що якийсь оновлення чи розширення системи не буде узгоджуватись з наявними даними та моделями.

Одними з таких прикладів є «дашборди» (аналітичні панелі) з розвиненим інтерфейсом, коли користувач може сформулювати майже довільний запит і отримати результат на карті, у вигляді графіка або файлу на експорт [5]—[7]. Дуже неприємною є ситуація, коли замість результату складного запиту виводиться повідомлення про відсутність даних. Бажано, щоб вона виводилась ще на етапі формування запиту, не даючи вибрати ті чи інші параметри.

Головним у цій задачі є проєктування інформаційної моделі такої системи, тобто даних та їхніх зв'язків з системою запитів та функцій для обчислення аналітичних даних, які б дозволили отримати інформацію з різним рівнем агрегації у часі, просторі чи за семантикою [8]. Але, якщо для агрегування даних у часі чи за семантикою з подальшим їхнім обробленням існують готові рішення (спеціальні функції усереднення даних з заданою частотою: хвилина, година, доба, тиждень, рік тощо, агрегування семантичних даних з використанням онтологічної моделі та спеціальних словників [8]), то агрегування та оброблення даних на різних просторових рівнях потребує використання моделей різної деталізації, тобто з різними складовими, потребує різних даних та методів їхнього отримання. Зазвичай, для таких задач проєктують окремі математичні та інформаційні моделі і забезпечують їх даними. Але тут вступають у протиріччя дві вимоги: повнота забезпечення кожної такої моделі всією необхідною інформацією із заданим агрегуванням та просторовою прив'язкою, а з іншого боку — відсутність надлишковості інформації та намагання зберегти усе в єдиній базі даних. Деякі дані (наприклад, метеорологічні) можуть одночасно використовуватись у моделях різного масштабу, а деякі (приміром, площа чи кількість населення) для кожного рівня масштабу у просторі — відрізнятимуться. Важливо, при цьому, точно знати, що кожна модель достатньо повно забезпечена даними або розрахунковими співвідношеннями для їхнього обчислення з інших даних у різних комбінаціях запитів.

Відповідь на питання достатності даних може дати критерій повноти топологічної спостережуваності. Як відомо, для складних систем, модель яких можна представити у вигляді графа зв'язків (зв'язків у розумінні залежностей, функцій чи алгоритмів перетворення одних даних на інші) між змінними, існує поняття топологічної спостережуваності, якою є сукупність умов, що забезпечують отримання інформації з урахуванням вхідних параметрів, про значення поточних параметрів режиму системи або забезпечують достатність інформації про поточний стан для синтезу закону керування ним [9]—[11]. В нашому випадку, цей критерій якраз можна використати для перевірки достатності забезпечення усіх інформаційних моделей системи на кожному рівні масштабу у просторі даними, достатніми для визначення поточного стану системи, які саме і потрібні для формування аналітичних даних та даних, необхідних для прийняття рішень управління системою. Важливим аспектом цього критерію є те, що існує математичний та алгоритмічний апарат оптимізації повноти топологічної спостережуваності системи, який дозволяє визначити які саме залежності чи дані ще необхідні для отримання повністю спостережуваної моделі [9], [10].

Отже, метою статті є розроблення методу оптимізації інформаційних моделей масштабованих у просторі аналітичних веб-систем за критерієм повноти їхньої топологічної спостережуваності та з урахуванням типових вимог до веб-систем, зокрема щодо усунення надлишковості інформації в їхніх базах даних.

Основні поняття та формалізація постановки задачі

Запишемо інформаційну модель M системи у вигляді множини чи системи співвідношень

$$M = \{y_j = f_p(x_1, x_2, \dots, x_N, y_1, y_2, \dots, y_K), j = 1, 2, \dots, K, p = 1, 2, \dots, P\}, \quad (1)$$

де $x_i, i = 1, \dots, N$ — вхідні змінні, якими є первинні дані, що надходять ззовні шляхом уведення

користувачем, чи з якихось датчиків, чи з відкритих даних через API тощо, $y_j, j = 1, \dots, K$ — змінні стану, тобто — проміжні змінні, які формуються вже у системі і є або результатом виконання запитів до бази даних, або результатами обчислень у функціях користувача чи ін. — будь-які, що формуються певною залежністю (чи функцією або алгоритмом) $f_p, p = 1, \dots, P$ з певної кількості (від 0 до N) вхідних змінних та з певної кількості змінних стану (від 0 до N , але на вході алгоритму має бути не менше однієї вхідної змінної або змінної стану), і результат яких теж зберігається у базі даних або у пам'яті комп'ютера, наприклад, для виведення у вигляді графіка чи на карті. У загальному випадку, у моделі (1) можна формалізувати й петлі, тобто та сама змінна стану може бути і на вході, і на виході залежності f_p , оскільки на етапі перевірки моделі (1) на спостережуваність особливості обчислення змінних в часі не вивчаються: інформаційна модель (1) враховує тільки види потоків даних та алгоритмів оброблення їхніх різних комбінацій.

У класичних системах керування, зазвичай, виділяються окремо змінні стану та вихідні змінні, оскільки для задач керування потрібні, передусім, вихідні змінні, але в аналітичних веб-системах результатом може бути й просто графік будь-якої змінної стану (або й первинних даних), тому тут така класифікація не потрібна.

Метод перевірки на топологічну спостережуваність полягає в тому, що насамперед вирази (1) перетворюються на біхроматичний граф з двома типами вершин: вершини-змінні (x_i або y_j) (відображаються, зазвичай, у вигляді заштрихованих кіл) і вершини-залежності (f_p) (відображаються у вигляді незаштрихованих кіл), які пов'язуються ребрами, кількість яких дорівнює кількості вхідних і вихідних змінних у кожній залежності f_p моделі (1). Важливе обмеження цього методу полягає в тому, що кожна залежність моделі (1) повинна мати тільки один вихід. Якщо виходів має бути більше, тоді слід залежність розбити на декілька, результатом яких все-таки буде тільки один вихід. Системи, алгоритми перетворення даних в яких неможливо розбити на такі, де буде тільки один вихід, не є предметом цієї статті і є обмеженням розроблюваного методу.

На другому етапі перевірки на топологічну спостережуваність на біхроматичному графі виділяють усі так звані «сильні» ребра, в кожне з яких входить одна вершина-залежність (тільки одна — не більше і не менше). Звичайно, для складних систем комбінацій таких ребер багато. Але головне, визначити так зване максимальне паросполучення, тобто таку комбінацію, де кількість n таких сильних ребер (точніше: кількість вершин-змінних, кожна з яких входить хоча б в одне «сильне» ребро [9]—[11]) буде найбільшою. А тоді, якщо $n = N$, тобто кількість сильних ребер максимального паросполучення дорівнює кількості вершин-залежностей, тоді підсумовується, що інформаційна модель (1) є топологічно спостережуваною.

У роботі [9] запропоновано ввести критерій повноти топологічної спостережуваності J

$$J = \frac{n}{N}, \quad (2)$$

зокрема, випадок $J = 1$ означає повну топологічну спостережуваність, а менші значення — неповну. Цей критерій потрібний для порівняння різних варіантів інформаційних моделей, коли повноти досягти не вдається, але можна спробувати хоча б удосконалити модель, щоб максимізувати цей критерій, а тоді різними програмними прийомами в інтерфейсі системи обмежити появу ситуацій, коли неповна топологічна спостережуваність системи призведе до непрацездатності системи чи до нештатного виконання її основних функцій.

Зазвичай, для виконання умови $J = 1$ необхідно і достатньо, щоб виконувалась умова в позначеннях моделі (1)

$$P \geq N + K, \quad (3)$$

тобто кількість залежностей має бути не меншою сумарної кількості вхідних змінних та змінних стану.

Наведемо *приклад*. Нехай інформаційна модель певної системи у нотації (1) має інформаційну модель такого вигляду ($N = K = 2; P = 4$):

$$\begin{cases} y_1 = f_1(x_1), \\ y_1 = f_2(x_1, x_2), \\ y_2 = f_3(x_1), \\ y_2 = f_4(x_2, y_1). \end{cases} \quad (4)$$

На рис. 1 показано біхроматичний граф (БГ) для моделі (4).

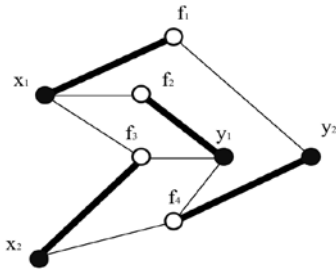


Рис. 1. Біхроматичний граф для інформаційної моделі (4) з виділеними на ньому ребрами з максимальним паросполученням, для якого кількість сильних (товсті лінії) ребер дорівнює $n = 4$

У роботі [9], також, показано, яким чином можна побудувати БГ для багатьох класів інформаційних систем, зокрема для багатозв'язаних із даними з просторовою прив'язкою та формалізацією просторових даних у вигляді геометричних мереж; запропоновано метод оптимізації БГ за критерієм (2) та описана низка правил, які дозволяють виконати зворотню трансформацію, тобто отримати нову інформаційну модель у нотації (1) за оптимізованим БГ.

Повністю топологічно спостережувана інформаційна модель веб-системи дозволяє бути впевненим, що будь-який запит чи функція із множини $F = \{f_p, p = 1, \dots, P\}$ спрацює і дасть на виході результат із множини $Y = \{y_j, j = 1, \dots, K\}$. Одразу зазначимо, що програмні особливості типу ділення на нуль, помилки у роботі програми та інші проблемні результати не є предметом цієї статті — стандартні етапи тестування програмного забезпечення — це подальший етап розроблення чи удосконалення веб-систем. Ця стаття стосується тільки оптимізації інформаційної моделі, тобто уникнення ситуацій, коли для якогось запиту чи алгоритму з множини F просто не вистачатиме первинних даних з множини $X = \{x_i, i = 1, \dots, N\}$ або проміжних змінних з множини Y .

Розроблення методу

Ідея запропонованого методу полягає в тому, що спершу для кожного рівня масштабу будується своя модель

$$M = \{M^r\}, r = 1, 2, \dots, R;$$

$$M^r = \left\{ y_j^r = f_p^r(x_1^r, x_2^r, \dots, x_{N^r}^r, y_1^r, y_2^r, \dots, y_{K^r}^r), j = 1, 2, \dots, K^r, p = 1, 2, \dots, P^r \right\}, \quad (5)$$

де $x_i^r, i = 1, \dots, N^r$ — вхідні змінні для r -го рівня моделі системи, $y_j^r, j = 1, \dots, K^r$ — змінні стану, які формуються у системі за певною залежністю чи алгоритмом $f_p^r, p = 1, \dots, P^r$ з первинних даних та, можливо, й зі змінних стану.

Вважаємо, що 1-й рівень найагрегованіший, а 2-й рівень — менший за розміром, тобто збільшення масштабу (зменшення рівня агрегування) — це перехід від 1-го рівня до 2-го, до 3-го і т. д. Наприклад, 1-й рівень — країна, 2-й — область, 3-й — район, 4-й — місто чи підприємство тощо.

Для кожної моделі (5) r -го рівня будується свій БГ r -го рівня. Для нього визначається кількість сильних ребер (вершин-змінних, що в них входять) n^r і обчислюється критерій J^r повноти топологічної спостережуваності

$$J^r = \frac{n^r}{N^r}, r = 1, 2, \dots, R. \quad (6)$$

Далі, перевіряється виконання умови для інтегрального критерію J повноти топологічної спостережуваності усієї інформаційної моделі системи

$$J = \frac{1}{R} \sum_{r=1}^R J^r = 1. \quad (7)$$

Якщо умова (7) не виконується, тоді проводиться оптимізація кожного БГ r -го рівня шляхом утворення нових залежностей на тому ж рівні. Якщо це не допомогло, тоді залежності утворюються між змінними моделей різних рівнів (назвемо їхні «міжрівневі»), доки не вдасться максимізувати інтегральний критерій J . Прикладами таких міжрівневих залежностей можуть бути, до прикладу, функції агрегування даних, коли дані по адміністративній області формуються як сума даних по районах або дані басейну річки на основі даних по масивах вод тощо.

Оскільки щодо міжрівневих залежностей виникне двозначність щодо того, до якого рівня їхні відносити, зокрема щодо позначення верхнього індексу, пропонується ввести спеціальне позначення для залежностей між змінними рівнів r_1 та r_2 : $g_w^{r_1, r_2}$, $w = 1, \dots, W$. Тоді модель (5) набуде вигляду

$$M = \{M^r\}, r = 1, 2, \dots, R;$$

$$\begin{aligned}
 M^r &= \{Y = F(X, Y), Y = G(X, Y)\}; \\
 X &= \{x_i^r, i = 1, \dots, N^r\}, Y = \{y_j^r, j = 1, \dots, K^r\}; \\
 F &= \{f_p^r(x_1^r, x_2^r, \dots, x_{N^r}^r, y_1^r, y_2^r, \dots, y_{K^r}^r), p = 1, 2, \dots, P^r\}; \\
 G &= \{g_w^{r1, r2}(x_1^r, x_2^r, \dots, x_{N^r}^r, y_1^r, y_2^r, \dots, y_{K^r}^r), w = 1, \dots, W\}.
 \end{aligned}
 \tag{8}$$

Важливо відмітити, що після додавання міжрівневих залежностей слід шукати максимальне паросполучення та визначати загальну кількість сильних ребер n вже для усєї моделі в цілому і користуватись формулою (2).

Якщо жоден з прийомів оптимізації усіх БГ не дав можливості досягти $J = 1$, тоді для даних та залежностей рівнів, де критерій (6) є меншим за 1, вводяться програмні обмеження і додаткові перевірки на функціональність інтерфейсу, враховуючи потенційну неможливість формування певних запитів.

Пояснимо алгоритм та ефективність додавання міжрівневих залежностей на прикладі. Нехай модель у нотації (5) має вигляд (на кожному рівні по дві вхідні змінні і по одній вихідній)

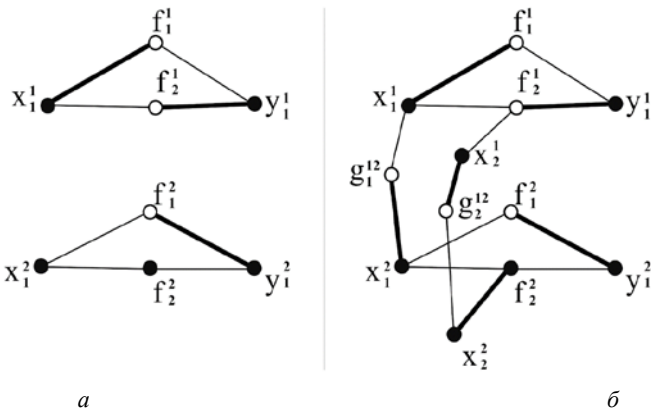
$$\begin{cases}
 y_1^1 = f_1^1(x_1^1), \\
 y_1^1 = f_2^1(x_1^1, x_2^1), \\
 y_1^2 = f_1^2(x_1^2), \\
 y_1^2 = f_2^2(x_1^2, x_2^2).
 \end{cases}
 \tag{9}$$

На рис. 2а показано БГ моделі (9), з якої видно, що максимальне паросполучення кожного рівня має тільки 2 сильних ребра, тобто

$$J^1 = \frac{n^1}{N^1} = \frac{2}{2+1} = 0,67 < 1, J^2 = \frac{n^2}{N^2} = \frac{2}{2+1} = 0,67 < 1.
 \tag{10}$$

Обчислюємо інтегральний критерій моделі, підставляючи (10) в (7),

$$J = \frac{1}{2} \sum_{r=1}^2 J^r = 0,67 < 1.
 \tag{11}$$



Отже, умова (7) не виконується. Переходимо до оптимізації моделі. Нехай, оптимізація на кожному рівні неможлива (чи вичерпано усі спроби, які не дали ефекту), але міжрівневі залежності можливі. Доповнюємо модель (9) до загального вигляду (8)

$$\begin{cases}
 y_1^1 = f_1^1(x_1^1), \\
 y_1^1 = f_2^1(x_1^1, x_2^1), \\
 y_1^2 = f_1^2(x_1^2), \\
 y_1^2 = f_2^2(x_1^2, x_2^2), \\
 x_1^1 = g_1^{1,2}(x_1^2), \\
 x_2^1 = g_2^{1,2}(x_2^2).
 \end{cases}
 \tag{12}$$

На рис. 2б показано БГ моделі (12), з якої випливає, що максимальне паросполучення усіх рівнів разом має $n = 6$ сильних ребер, що дорівнює сумарній кількості усіх вхідних змінних та змінних стану, а тому з (2)

$$J = \frac{n}{N} = \frac{6}{4+2} = 1.
 \tag{13}$$

Отже, система стала повністю топологічно спостережуваною.

Алгоритм застосування методу

Таким чином, алгоритм застосування запропонованого методу оптимізації інформаційних моделей масштабованих у просторі аналітичних веб-систем за критерієм повноти їхньої топологічної спостережуваності полягає у такому:

1. Виділити усі рівні $r = 1, 2, \dots, R$ масштабування у просторі, на яких планується виконувати аналітичні операції у проєктованій веб-системі.

2. Формалізувати задачу, виділивши на кожному r -му рівні елементи множин вхідних змінних X і змінних стану Y та множин залежностей (алгоритмів, функцій тощо) F .

3. Записати інформаційні моделі M^r для кожного рівня в нотації (5).

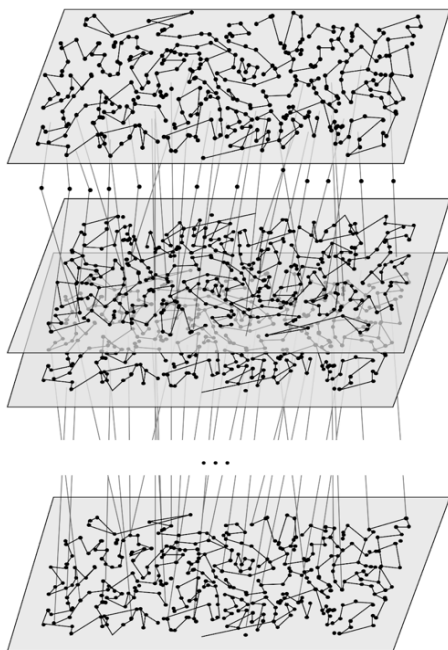


Рис. 3. Узагальнена схема біхроматичного графу багаторівневої масштабованої у просторі аналітичної веб-системи

4. Побудувати для кожної моделі M^r біхроматичний граф (БГ).

5. Для кожного БГ r -го рівня визначити максимальне паросполучення та визначити у ньому кількість сильних ребер n^r .

6. Визначити критерії повноти топологічної спостережуваності на кожному рівні J^r за формулою (6).

7. Обчислити інтегральний критерій повноти топологічної спостережуваності моделі системи в цілому J .

8. Перевірити виконання умови (7) ($J = 1?$).

9. Якщо умова (7) виконується, тоді перейти до завершення, інакше — на наступний етап.

10. Оптимізувати БГ для M^r для кожного r -го рівня окремо, після чого повторити пп. 4—6.

11. Якщо умова (7) виконується, то перейти до завершення, інакше — на наступний етап.

12. Доповнити модель (5) до моделі (8) введенням міжрівневих залежностей G у БГ системи в цілому.

13. Для БГ усієї системи в цілому визначити максимальне паросполучення та порахувати у ньому кількість сильних ребер n .

14. Обчислити інтегральний критерій повноти топологічної спостережуваності моделі системи в цілому J за формулою (2) та перевірити виконання умови « $J = 1?$ ».

15. Якщо умова $J = 1$ виконується, то перейти до завершення, інакше — на наступний етап.

16. Ввести програмні обмеження у функціональність веб-системи, щоб уникнути появи запитів чи залежностей, незабезпечених даних, через неповну топологічну спостережуваність окремих (чи усіх) рівнів системи.

17. Завершення алгоритму.

На рис. 3 показано схематичне представлення результату виконання п.12 цього алгоритму, тобто підсумковий оптимізований БГ моделі масштабованої у просторі аналітичної веб-системи.

Блок-схема цього алгоритму подана на рис. 4.

Наведемо приклад застосування запропонованого методу на вже реально створеній веб-системі.

Приклад застосування методу оптимізації інформаційних моделей масштабованих у просторі аналітичних веб-систем

Одна з проблем розвитку економіки будь-якої країни це — прогноз забезпеченості водою заданого регіону. Таке дослідження проводиться і в довгостроковій перспективі під час вибору місць розташування нових підприємств або інших виробничих потужностей, зокрема місць вирощування певних сільськогосподарських культур та ін., і в короткостроковій перспективі, щоб правильно спланувати роботу кожного виробництва чи водогосподарської споруди протягом кожного року, кварталу, місяця з урахуванням різних, і навіть фінансових та екологічних обмежень.

У трансграничному басейні Дністра протягом 2017—2021 років реалізовувався проєкт ГЕФ/ПРООН/ОБСЄ «Сприяння транскордонному співробітництву та комплексного управління водними ресурсами в басейні річки Дністер», у межах фінансування якого автори статті В. Б. Мокін і Є. М. Крижановський створювали автоматизовану веб-систему розрахунку і прогнозування водо-

господарського балансу басейну Дністра (<http://vb.dniester-commission.com/>), яка з 2018 року передана ОБСЄ для впровадження в систему управління урядами України та Республіки Молдови, а також уряду Республіки Польщі, на території якої розташована порівняно невелика ділянка басейну Дністра. На вимогу ОБСЄ веб-система має чотиримовний інтерфейс (англійський, румунський, український та російський) (рис. 5).

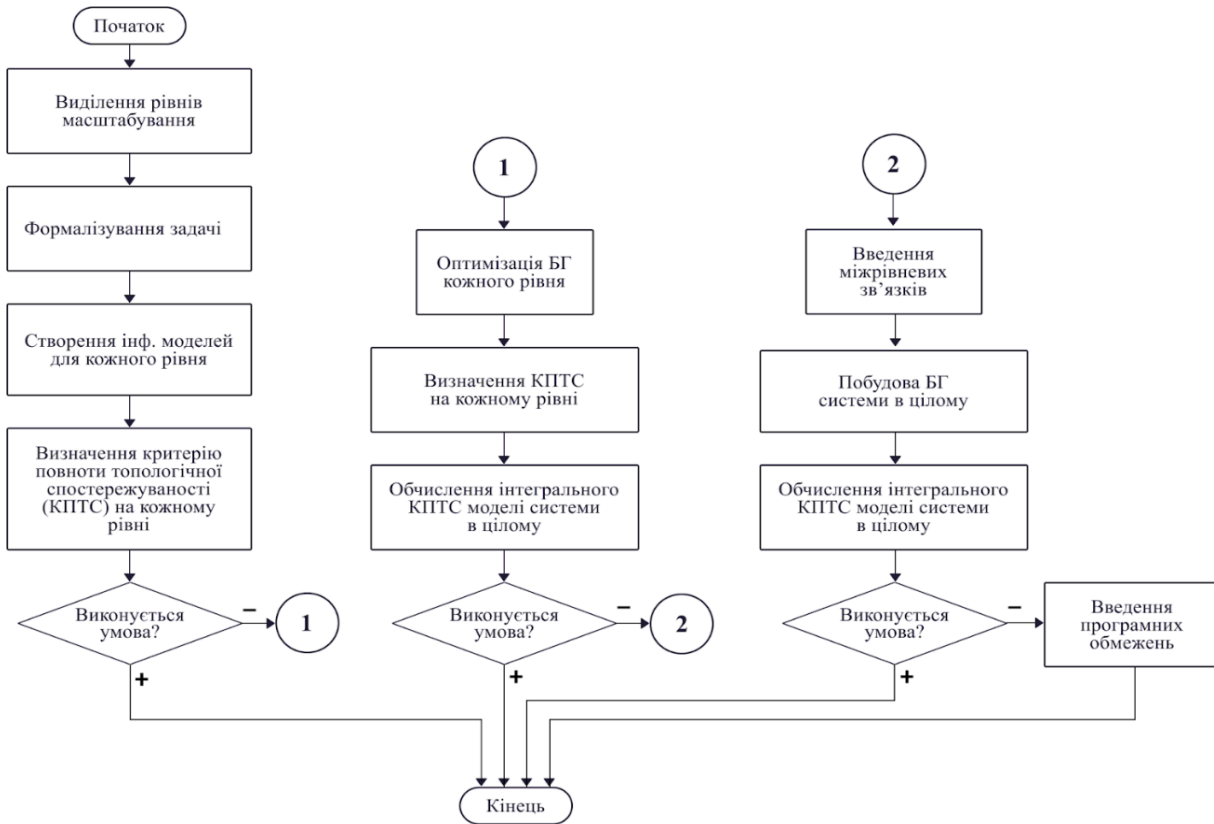


Рис. 4. Блок-схема алгоритму методу оптимізації інформаційних моделей масштабованих у просторі аналітичних веб-систем за критерієм повноти їхньої топологічної спостережуваності

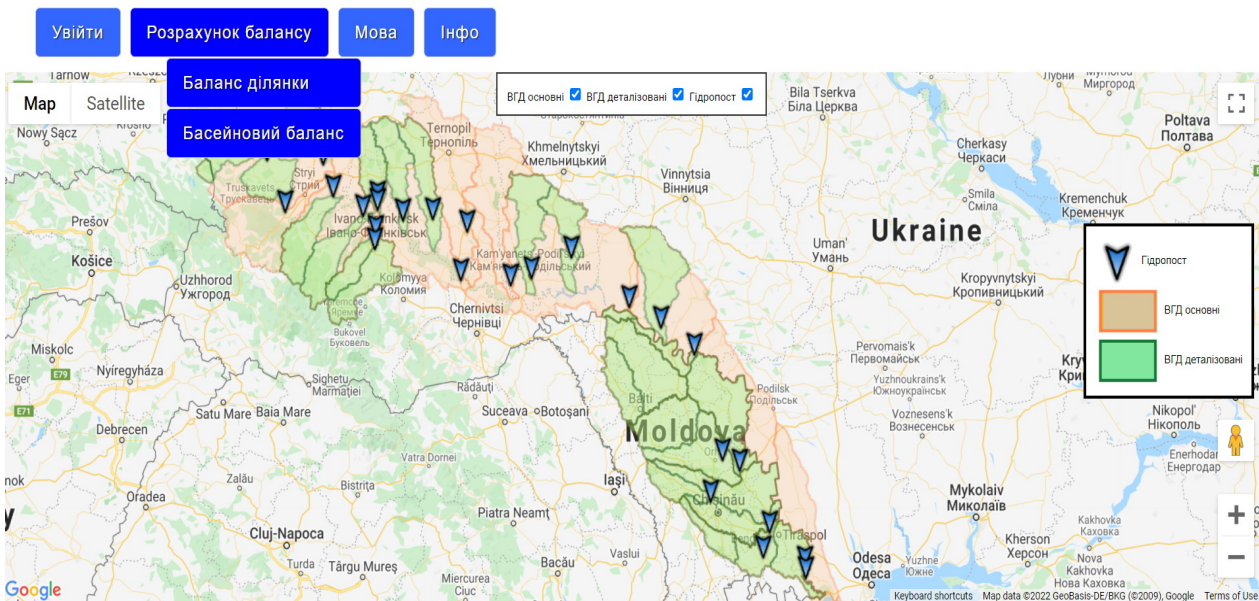


Рис. 5. Інтерфейс аналітичної веб-системи розрахунку і прогнозування водогосподарського балансу, створеної авторами у межах проєкту «Сприяння транскордонному співробітництву та комплексного управління водними ресурсами в басейні річки Дністер» (<http://vb.dniester-commission.com/>)

Автоматизована веб-система розрахунку і прогнозування водогосподарського балансу басейну Дністра забезпечує:

- 1) для неавторизованих користувачів:
 - перегляд верифікованих розрахунків балансу по водогосподарським ділянках з архіву системи;
 - перегляд карти з водогосподарськими ділянками і гідропостів;
 - розрахунок водогосподарського балансу для водогосподарських ділянок басейну Дністра для 50 %, 75 %, 95 % забезпеченості, а також баланс за вибраним роком і прогноз балансу з урахуванням зміни клімату;
- 2) для авторизованих користувачів:
 - перегляд інформації в системі;
 - введення, зміна інформації в системі в тих таблицях, за якими у користувачів є достовірні дані (права на доступ до кожної таблиці видаються адміністратором системи);
 - розрахунок водогосподарського балансу для водогосподарських ділянок басейну Дністра для 50 %, 75 %, 95 % забезпеченості, а також баланс за вибраним роком і прогноз балансу з урахуванням зміни клімату;
 - збереження в архів верифікованих розрахунків водогосподарського балансу;
 - перегляд карти з водогосподарськими ділянками і гідропостами;
 - перегляд верифікованих розрахунків балансу по водогосподарських ділянках з архіву системи;
 - перегляд статистики;
- 3) для адміністратора системи:
 - управління користувачами системи і їхніми правами;
 - управління архівом.

Створена у 2017—2021 роках веб-система є другою версією цієї ж системи, створеної співавторами В. Б. Мокіним та Є. М. Крижановським у межах компоненти «Зміна клімату та безпека в басейні річки Дністер» іншого проекту ОБСЄ «Зміна клімату та безпека у Східній Європі, Центральній Азії та на Північному Кавказі» протягом 2015—2016 років [12]—[14].

Ключова відмінність системи другої версії від першої — це можливість масштабування у просторі основних параметрів та створених прогнозів. На вимогу представників урядів Республіки Молдова та України забезпечено можливість розрахунку водогосподарського балансу не тільки водогосподарських ділянок, згідно із законодавчо затвердженим в Україні районуванням та проектом аналогічного районування у Республіці Молдова, а й — значно менших ділянок території. Адже, як відомо, чим менше просторова ділянка, дані якої аналізуються, тим точнішими і практично кориснішими є зроблений для неї аналіз та прогноз.

Побудовано інформаційну модель для кожного рівня: «загальне районування» та «деталізоване районування». Проаналізовані усі можливі запити та співвідношення між ними. Частина результату такого аналізу опублікована у монографії [9], де подано й біхроматичний граф, точніше — його формалізація у вигляді геоінформаційного простору параметрів (це варіант БГ, адаптований одночасно і до рівнянь у вигляді (5), і до формату шарів геоінформаційних систем, у вигляді яких запропоновано зберігати і редагувати БГ). Аналіз виявив, що для низки ділянок досягнення умови (7) є неможливим, оскільки не вистачає первинних даних про стік річки, через відсутність жодного метеопоста на території. Тоді знову створили робочу групу проекту, яка оптимізувала деталізоване районування, щоб забезпечити в такий спосіб повноту топологічної спостережуваності інформаційної моделі на обох рівнях. Одним з елементів такої оптимізації стало обмеження ділянок деталізованого районування кордонами держав — на території кожної із трьох країн виділено свої ділянки. Саме це дозволило узгодити міжрівневі залежності. Справа в тім, що інформаційні моделі створені для районування двох видів:

- басейнове (основне): басейн, водогосподарська ділянка основного районування, водогосподарська ділянка деталізованого районування;
- адміністративне (для частини даних на кшталт обсягів водокористування та ін.): країна, для України — область країни.

Частина даних була тільки по Польщі, тільки по Молдові, тільки по Україні, частина — зібрана та агрегована за басейновим принципом. Оскільки інтерфейс введення первинних даних передбачає, що кожна країна вводить свої дані окремо, то важливо було передбачити поєднання і басейнової, і адміністративної інформаційної моделі даних. Саме метод оптимізації цих моделей з використання міжрівневих залежностей допоміг розв'язати задачу.

В подальшому виконанні програмні роботи, проведено навчання представників обох основних країн, усім передано паролі доступу. Частина функціональності системи (без редагування первинних даних) є доступною без авторизації.

Отже, запропонований метод довів свою ефективність для розв'язання складної прикладної задачі.

Висновки

Розглянуто деякі питання оптимізації інформаційних моделей масштабованих у просторі аналітичних веб-систем за критерієм повноти їхньої топологічної спостережуваності.

Розроблено новий метод оптимізації інформаційних моделей масштабованих у просторі аналітичних веб-систем за критерієм повноти їхньої топологічної спостережуваності, який дозволяє здійснити аналіз повноти топологічної спостережуваності як на кожному рівні масштабів у просторі, на яких планується виконувати аналітичні операції у проєктованій веб-системі, так і для всієї системи загалом. За результатами такого аналізу з'являється можливість введення програмних обмежень у функціональності веб-системи, щоб уникнути появи запитів чи залежностей, не забезпечених даними, через неповну топологічну спостережуваність окремих (чи усіх) рівнів системи.

У роботі продемонстровано роботу методу на прикладі розробки автоматизованої веб-системи розрахунку і прогнозування водогосподарського балансу басейну Дністра. Застосування розробленого підходу на етапі проєктування цієї системи забезпечило уникнення появи запитів чи залежностей, не забезпечених даними. Систему розроблено та успішно впроваджено.

Наведено приклад у галузі управління водними ресурсами, який продемонстрував працездатність запропонованих рішень. Використовуючи запропонований метод в межах проєкту ГЕФ/ПРООН/ОБСЄ «Сприяння транскордонному співробітництву та комплексного управління водними ресурсами в басейні річки Дністер» розроблено та впроваджено автоматизовану веб-систему розрахунку і прогнозування водогосподарського балансу басейну Дністра.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- [1] M. Bandara, A. Behnaz, F. A. Rabhi, and O. Demirors, "From Requirements to Data Analytics Process: An Ontology-Based Approach," *Business Process Management Workshops*, vol. 342, pp. 543-552, 2019. https://doi.org/10.1007/978-3-030-11641-5_43.
- [2] M. Shu, S. Tan, L. Fu, Y. Zeng, X. Cao, and Y. Zeng, "Application of Design Methodologies to Web System Design: A Case Study of JIDPS Editorial System," *Journal of Integrated Design and Process Science*, vol. 21, no. 4, pp. 79-112, 2017. <https://doi.org/10.3233/jid-2017-0020>.
- [3] H.-M. Chen, R. Kazman, and S. Haziyevev, "Agile Big Data Analytics for Web-Based Systems: An Architecture-Centric Approach," *IEEE Transactions on Big Data*, vol. 2, pp. 234-248, 2016. <https://doi.org/10.1109/tbdata.2016.2564982>.
- [4] L. F. F. G. Assis, et al., "TerraBrasilis: A Spatial Data Analytics Infrastructure for Large-Scale Thematic Mapping," *ISPRS International Journal of Geo-Information*, vol. 8, no. 11, pp. 513, 2019. <https://doi.org/10.3390/ijgi8110513>.
- [5] K. S. Duisebekova, D. K. Kozhamzharova, S. B. Rakhmetulayeva, F. A. Umarov, and M. Zh. Aitimov, "Development of an information-analytical system for the analysis and monitoring of climatic and ecological changes in the environment," *Procedia Computer Science*, vol. 170, pp. 578-583, 2020. <https://doi.org/10.1016/j.procs.2020.03.128>.
- [6] D. Y. Paramartha, A. L. Fitriyani, and S. Pramana, "Development of Automated Environmental Data Collection System and Environment Statistics Dashboard," *Indonesian Journal of Statistics and Its Applications*, vol. 5, no. 2, pp. 314-325, 2021. <https://doi.org/10.29244/ijsa.v5i2p314-325>.
- [7] A. Fruhling, M. Hall, S. Medcalf, and A. Yoder, "Designing a Real-Time Integrated First Responder Health and Environmental Monitoring Dashboard," *Designing for Digital Transformation. Co-Creating Services with Citizens and Industry*, vol. 12388, pp. 28-34, 2020. https://doi.org/10.1007/978-3-030-64823-7_3.
- [8] В. Б. Мокін, І. І. Овчаренко, А. М. Лучко, і О. М. Давидюк, «Побудова масштабованої інформаційно-пошукової системи для управління річковим басейном на основі реєстрів та онтологічних моделей», *Математичне моделювання в економіці*, № 2 (15), с. 45-56, 2019.
- [9] В. Б. Мокін, І. В. Варчук, і Є. М. Крижановський, *Інформаційна технологія аналізу та оптимізації топологічної спостережуваності багатозв'язних геоінформаційних систем*, моногр. Вінниця, Україна: ВНТУ, 2019, 121 с.
- [10] А. З. Гамм, І. І. Голуб, і Г. Н. Ополева, «Некоторые задачи анализа режима электроэнергетических систем по данным измерений», *Электричество*, № 6, с. 1-6, 1984.
- [11] A. N. Montanari, and L. A. Aguirre, "Observability of Network Systems: A Critical Review of Recent Results," *Journal of Control, Automation and Electrical Systems*, № 31, pp. 1348-1374, 2020. [Online]. Available: <https://doi.org/10.1007/s40313-020-00633-5>.
- [12] Б. Либерт и др., *Днестр без границ: результаты проекта «Трансграничное сотрудничество и устойчивое управление в бассейне реки Днестр: фаза III – реализация программы действий» ("Днестр-III")*. Киев, Украина: Вайтэ, 2013, 172 с. ISBN 978-966-2310-13-9.
- [13] В. Б. Мокін, Є. М. Крижановський, Л. М. Скорина, і В. В. Гребінь, «Технологія оптимізації управління водними ресурсами басейну р. Дністер шляхом автоматизації складання його водогосподарського балансу», на *XIV Міжнародна науково-практична конференція: Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях*: Київ, 2015, с. 131-134.

[14] В. Б. Мокін та ін., «Розробка моделі водогосподарського балансу української частини районів басейнів річок Дунай, Західний Буг та суббасейну річки Десна. Верифікація та адаптація водогосподарських балансів української частини районів басейнів річок Дон, Дністер, району басейну річки Південний Буг та суббасейну річки Прип'ять до вимог нормативно-правових актів Мінприроди.» звіт про НДР: № 2853. Вінницький національний технічний університет; Київ, Україна, 2017, 108 с. № ДР 0117U005018. Інв. № 0217U002783.

Рекомендовано до друку кафедрою системного аналізу та інформаційних технологій ВНТУ

Стаття надійшла до редакції 17.12.2021

Мокін Віталій Борисович — д-р техн. наук, професор, завідувач кафедри системного аналізу та інформаційних технологій, e-mail: vbmokin@gmail.com ;

Крижановський Євгеній Миколайович — канд. техн. наук, доцент кафедри системного аналізу та інформаційних технологій, e-mail: kruzhan@gmail.com ;

Лучко Андрій Михайлович — аспірант кафедри системного аналізу та інформаційних технологій, e-mail: andriyluchko@gmail.com ;

Білецький Богдан Сергійович — аспірант кафедри системного аналізу та інформаційних технологій, e-mail: bohdanbeletskyi@gmail.com ;

Жуков Сергій Олександрович — канд. техн. наук, доцент кафедри системного аналізу та інформаційних технологій, e-mail: sazhukov@gmail.com .

Вінницький національний технічний університет, Вінниця

V. B. Mokin¹
Ye. M. Kryzhanovskiy¹
A. M. Luchko¹
B. S. Biletskyi¹
S. O. Zhukov¹

Method of Optimization of Information Models Scaled in Space of Analytical Web-Systems by the Criteria of Completeness of their Topological Observation

¹Vinnitsia National Technical University

The article considers the issue of optimization of information models of space-scaled analytical web systems by the criterion of completeness of their topological observability.

The previously created, with the participation of some co-authors, information technology analysis and optimization of topological observability of multilinked geographic information systems, i.e. — optimization of the system model in the form of a dichromatic graph, which allows arbitrary input variables to obtain values of state variables or current system parameters. information about the current state to synthesize the law of its management. In such a bichromatic graph, some vertices are variable and others are dependencies between them. This technology introduces the criterion of observability, which can take values from 0 to 1 and compare different models according to the level of their observability and optimize them by choosing the most observable option. This paper proposes how such a model formalization and this technology can be extended to the formalization of space-scaled analytical web systems. Then a new method of optimizing information models of such systems by the criterion of completeness of their topological observability was proposed, which allows to analyze the completeness of this observability both at each level of scaling in space where analytical operations are planned in the projected web system and for the whole systems in general. It is proposed to optimize the model by introducing additional vertices-variables and vertices-dependencies between different levels. That is, in the absence of data or dependencies at one level of the system, they can sometimes be obtained from data or dependencies at another level, and vice versa, which will increase the observability of the system as a whole. The method allows both to analyze the level of topological observability and to analyze it after optimization. As a result of such analysis, it is possible to impose software constraints on the functionality of the web system to avoid queries or dependencies that are not provided with data or dependencies from which they could be obtained due to incomplete topological observability of individual (or all) levels.

The paper demonstrates the work of the method on the example of developing an automated web system for calculating and forecasting the water balance of the Dniester basin. The application of the developed approach at the design stage of this system ensured the avoidance of queries or dependencies not provided with data. The system was developed and successfully implemented.

The developed method of optimizing information models of space-scaled analytical web systems by the criterion of completeness of their topological observability can be used in the analysis and design of systems that reflect environmental or infrastructure data distributed in space and have several levels of detail. These can be cadastral systems, environmental monitoring systems, etc.

Keywords: analytical web system, topological observability, information model, scalability of spatial data, web systems design.

Mokin Vitalii B. — Dr. Sc. (Eng.), Professor, Head of the Chair of System Analysis and Information Technology, e-mail: vbmokin@gmail.com ;

Kryzhanovskiy Yevhenii M. — Cand. Sc. (Eng.), Associate Professor of the Chair of System Analysis and Information Technology, e-mail: kruzhan@gmail.com ;

Luchko Andrii M. — Post-Graduate Student of the Chair of System Analysis and Information Technology, e-mail: andriyluchko@gmail.com ;

Biletskyi Bohdan S. — Post-Graduate Student of the Chair of System Analysis and Information Technology, e-mail: bohdanbeletskyi@gmail.com ;

Zhukov Serhii O. — Cand. Sc. (Eng.), Associate Professor of the Chair of System Analysis and Information Technologies, e-mail: sazhukov@gmail.com

В. Б. Мокин¹
Е. Н. Крыжановский¹
А. М. Лучко¹
Б. С. Белецкий¹
С. А. Жуков¹

Метод оптимизации информационных моделей масштабированных в пространстве аналитических веб-систем по критерию полноты их топологической наблюдаемости

¹Вінницький національний технічний університет

Рассмотрены вопросы оптимизации информационных моделей масштабируемых в пространстве аналитических веб-систем по критерию полноты их топологической наблюдаемости. Охарактеризовано созданную ранее, при участии части соавторов, информационную технологию анализа и оптимизации топологической наблюдаемости многосвязных геоинформационных систем, т.е. сведения о текущем состоянии для синтеза закона управления ним. В таком бихроматическом графе одни вершины переменные, а другие — зависимости между ними. В этой технологии введен критерий наблюдаемости, который может принимать значения от 0 до 1 и сравнивать разные модели по уровню их наблюдаемости и оптимизировать их, выбирая наиболее наблюдаемый вариант. Предложено каким образом такая формализация модели и эта технология может быть распространена на формализацию масштабируемых в пространстве аналитических веб-систем. А также предложен новый метод оптимизации информационных моделей таких систем по критерию полноты их топологической наблюдаемости, позволяющий осуществить анализ полноты этой наблюдаемости как на каждом уровне масштабирования в пространстве, на которых планируется выполнять аналитические операции в проектируемой веб-системе, так и для всей системы в целом. Предложено оптимизировать модель, за счет ввода дополнительных вершин-переменных и вершин-зависимостей между разными уровнями. То есть, при отсутствии данных или зависимостей на определенном уровне системы, иногда можно их получить из данных или зависимостей другого уровня, и наоборот, что повысит наблюдаемость системы в целом. Метод позволяет как анализировать уровень топологической наблюдаемости, так и анализировать его после оптимизации. По результатам такого анализа появляется возможность ввода программных ограничений в функциональности веб-системы во избежание появления запросов или зависимостей, не обеспеченных данными или зависимостями, из которых их можно было бы получить, из-за неполной топологической наблюдаемости отдельных (или всех) уровней системы.

Продемонстрирована работа метода на примере разработки автоматизированной веб-системы расчета и прогнозирования водохозяйственного баланса бассейна реки Днестр. Применение разработанного подхода на этапе проектирования этой системы обеспечило избежание появления запросов или зависимостей, не обеспеченных данными. Система была разработана и успешно внедрена.

Разработанный метод оптимизации информационных моделей масштабируемых в пространстве аналитических веб-систем по критерию полноты их топологической наблюдаемости может применяться при анализе и проектировании систем, отражающих данные окружающей среды или инфраструктуры, распределенные в пространстве и имеющие несколько уровней детализации. Это могут быть кадастровые системы, системы мониторинга состояния окружающей среды и т. п.

Ключевые слова: аналитическая веб-система, топологическая наблюдаемость, информационная модель, масштабированность пространственных данных, проектирование веб-систем.

Мокін Віталій Борисович — д-р техн. наук, професор, завідувач кафедри системного аналізу та інформаційних технологій, e-mail: vbmokin@gmail.com ;

Крыжановський Євгеній Николаєвич — канд. техн. наук, доцент кафедри системного аналізу та інформаційних технологій, e-mail: kruzhan@gmail.com ;

Лучко Андрій Михайлович — аспірант кафедри системного аналізу та інформаційних технологій, e-mail: andriyluchko@gmail.com ;

Белецький Богдан Сергєєвич — аспірант кафедри системного аналізу та інформаційних технологій, e-mail: bohdanbeletskyi@gmail.com ;

Жуков Сергєй Александрович — канд. техн. наук, доцент кафедри системного аналізу та інформаційних технологій, e-mail: sazhukov@gmail.com