

ІНТЕГРАЦІЯ ГЕОПРОСТОРОВИХ ДАНИХ НА ОСНОВІ ЗАСТОСУВАННЯ ОПЕРАЦІЇ З'ЄДНАННЯ (JOIN) РЕЛЯЦІЙНОЇ АЛГЕБРИ

Мета цієї роботи – дослідження інтеграції наборів базових і тематичних геопросторових даних на основі операції з'єднання (Join) реляційної алгебри та її взаємодії із геокодуванням геопросторових об'єктів, яку реалізовано в сучасних геоінформаційних системах (далі – ГІС) та системах керування базами даних (далі – СКБД) для розвитку національної інфраструктури геопросторових даних (далі – НІГД). *Методика*. Основою дослідження є аналіз можливостей застосування теорії баз геопросторових даних і баз знань, міжнародних і національних гармонізованих стандартів у сфері “Географічна інформація/Геоматика” для вирішення питання інтеграції геопросторових даних за допомогою операції з'єднання JOIN реляційної алгебри в об'єктно-реляційних системах керування базами даних (ОР СКБД). *Результати*. В статті досліджено моделі операції з'єднання Join реляційної алгебри, покладені в основу геокодування об'єктів і створення електронних газетирів, і доведено її ефективність: операція з'єднання Join забезпечує інтеграцію наборів базових і тематичних геопросторових даних. Для її виконання необхідно визначити обов'язкові географічні ідентифікатори, які мають бути наявні серед атрибутів наборів базових та тематичних геопросторових даних та за якими виконується з'єднання. Різноманітність видів застосування операції з'єднання Join охоплює всі можливі випадки, які виникають під час їх практичного використання. Отже, для використання операції з'єднання Join на етапі проєктування баз геопросторових даних необхідно визначити ці обов'язкові географічні ідентифікатори. Зокрема, доцільно визначити обов'язкові географічні ідентифікатори (коди) об'єктів за офіційними загальнодержавними системами класифікації (кодифікації) об'єктів у галузевих тематичних реєстрах, за які відповідають визначені тримачі тематичних даних згідно із додатком 2 до Постанови Кабінету Міністрів України “Про затвердження Порядку функціонування національної інфраструктури геопросторових даних” від 26 травня 2021 р. № 532. Досліджено інтеграцію наборів базових і тематичних геопросторових даних на основі моделей операції з'єднання (Join) реляційної алгебри та їх взаємодії із геокодуванням геопросторових об'єктів, яку реалізовано в сучасних ГІС та СКБД для розвитку національної інфраструктури геопросторових даних. Для дослідження використано набір базових геопросторових даних, а саме: відомості про адміністративно-територіальні одиниці Черкаської області, зокрема їх межі; тематичними вибрано дані зі статистичного бюлетеня соціально-економічного становища Черкаської області за січень 2021 р. Головного управління статистики у Черкаській області Державної служби статистики України. Доведено, що операцію з'єднання (JOIN) реляційної алгебри можна використовувати для інтеграції інших тематичних геопросторових даних із базовими геопросторовими даними за допомогою географічних ідентифікаторів, які містять ці набори даних.

Ключові слова: інтеграція геопросторових даних; інтероперабельність; операція з'єднання (Join); національна інфраструктура геопросторових даних; базові геопросторові дані; тематичні геопросторові дані.

Вступ

У 2020 р. ухвалено Закон України “Про національну інфраструктуру геопросторових даних” № 554 від 13.04.2020, що сприяло розвитку національної інфраструктури геопросторових даних в Україні як взаємопов'язаної сукупності організаційної структури, технічних і програмних засобів, базових та тематичних наборів геопросторових даних, метаданих, сервісів, технічних регламентів, стандартів, технічних специфікацій,

необхідних для виробництва, оновлення, оброблення, зберігання, оприлюднення, використання геопросторових даних та метаданих, іншої діяльності із такими даними [Закон України, 2020]. У 2021 р. прийнято Постанову Кабінету Міністрів України “Про затвердження Порядку функціонування національної інфраструктури геопросторових даних” № 532 від 26.05.2021 р. Базові геопросторові дані є ядром НІГД, оскільки вони “забезпечують виробництво та використання тематичних геопросторових даних,

становлять уніфіковану єдину топографічну основу для забезпечення інтегрованості геопросторових даних, їх інтеграції та міжвідомчої інформаційної взаємодії. Основою створення наборів базових геопросторових даних є бази топографічних даних та відомості Державного земельного кадастру, Державного реєстру географічних назв, Державного адресного реєстру, містобудівного кадастру та кадастрів природних ресурсів, а також інших геоінформаційних ресурсів” [Постанова..., 2021]. Зазвичай атрибутивні характеристики топографічних об’єктів, які належать до базових геопросторових даних, збирають у результаті виконання топографічних знімів [Карпінський та Лазоренко-Гевель, 2018; Karpinskyi & Lazorenko-Nevel, 2020], тому вони, як правило, не містять повноцінної тематичної інформації про ці об’єкти. Набори тематичних геопросторових даних формують органи виконавчої влади та органи місцевого самоврядування та інші тримачі даних, відповідальні за ці набори даних, згідно із додатком 2 до Постанови Кабінету Міністрів України “Про затвердження Порядку функціонування національної інфраструктури геопросторових даних” від 26.05.2021 р. № 532 [Постанова..., 2021]. Міністерство аграрної політики та продовольства України видало Наказ “Про затвердження технічних вимог до геопросторових даних, метаданих і геоінформаційних сервісів національної інфраструктури геопросторових даних” № 347 від 10.11.2021 р., у якому зазначено, що саме уніфікація та інтеграція базових і тематичних геопросторових даних забезпечують роботу НІГД [Наказ, 2021]. Одним із засобів інтегрування даних є з’єднання різноманітних баз геопросторових даних із даними, які не містять геопросторової складової, за допомогою системного ідентифікатора для їх зведення з різних джерел у НІГД. Є два способи інтегрування даних: пряме геокодування, яке здійснюється в результаті топографо-геодезичних робіт, й опосередковане (непряме), в якому ключову роль відіграють географічні ідентифікатори. Саме завдяки таким географічним ідентифікаторам можливе з’єднання геопросторових координатних описів об’єктів із масивом атрибутивних даних, які не мають прямого позиціонування [Географічна інформація... ДСТУ ISO 19112:2017, 2017; Карпінський та Лазоренко-Гевель, 2020;

Karpinskyi et al., 2020; Karpinskyi & Lazorenko-Nevel, 2020; Лазоренко-Гевель, 2021].

Операція геокодування є одним з найпоширеніших геопросторових інструментів і сервісів у всіх популярних географічних інформаційних системах та системах керування базами даних. За допомогою цієї інформації визначають геопросторову інформацію (наприклад, координати x, y) для описів місцеположення об’єктів, порівнюючи елементи описового місцеположення з тими, що містяться у довідкових даних реєстрів: адресами, географічними назвами тощо. Інакше кажучи, геокодування – це забезпечення відповідності даних реєстрів об’єктам на карті, для роботи з якими необхідно мати базу даних місцеположень об’єктів, так званих газетирів – каталогів географічних ідентифікаторів, в яких описано екземпляри місцеположення. Ці каталоги містять додаткову інформацію про розташування кожного екземпляра місцеположення, можуть охоплювати координатну прив’язку, але можуть бути і суто описовими. Якщо газетир містить координатну прив’язку, то це забезпечує трансформування із просторової референцної системи з використанням географічних ідентифікаторів у іншу референцну систему координат. Якщо ж газетир містить описові посилання, то це буде просторова прив’язка із використанням іншої просторової референцної системи з географічними ідентифікаторами, наприклад, поштовий індекс об’єкта нерухомості. Для будь-якого типу місцеположення може існувати більше ніж один довідник. Описову прив’язку може бути визначено в атрибуті просторового поширення (geographicExtent) для географічного ідентифікатора в газетирі [Національний стандарт України ДСТУ 8774, 2018; Лазоренко-Гевель, 2021].

Нині існує багато настільних та вебінструментів геокодування у ГІС. Наприклад, в геоінформаційній системі із відкритою ліцензією Quantum GIS (далі – QGIS) для геокодування використано спеціальний плагін *MMQGIS Plugin*, який є набором плагінів Python для маніпулювання шарами векторної карти в QGIS: введення/виведення/об’єднання даних у форматі CSV, геокодування з шару вулиць, що дає змогу створювати адресні реєстри із джерел даних у форматі CSV, геокодування у форматі CSV із вебсервісом, який дає

можливість геокодувати таблиці адрес через сервіси геокодування, перетворення геометрії тощо. До таких сервісів належать:

1) Nominatim – всесвітній сервіс геокодування із відкритим вихідним кодом, створений на основі даних OpenStreetMap;

2) сервіс геокодування OSMNames, що є похідним від даних OSM, вся база даних доступна для завантаження і є можливість (принаймні деяка) подавати запити через їхній API;

3) Google world geocoding service або Google gazetteer – всесвітній сервіс геокодування від компанії Google, який потребує попереднього ключа API геокодування та має щоденний ліміт (нині 2500 адрес);

4) компанії Mapbox та HERE мають власні сервіси геокодування. Геору є клієнтом геокодування для кількох популярних вебсервісів геокодування, урахувавши Nominatim і Google;

5) ESRI World geocoding Service – всесвітній сервіс геокодування від компанії ESRI. Мова програмування Python, а саме ArcPy, має доступ до інструментів геокодування в ArcGIS [How to Geocode, 2022].

У комерційній ГІС ArcGIS використовують кілька інструментів геокодування, з яких три є основними: *Add X,Y Coordinate Data as a Map Layer* (дати координатні дані X, Y як шар карти), *Create an Address Locator* (створити геометричне представлення місцеположень об'єктів) та *ArcGIS World Geocoding Service* (світовий сервіс геокодування ArcGIS), який дає змогу знаходити адреси або місцеположення по всьому світу, геокодувати таблиці адрес або здійснювати зворотне геокодування тощо, використовуючи *ArcGIS Rest API*. Для роботи з конфіденційними даними компанії створюють свої адресні реєстри.

Також в QGIS є можливість скористатися *Pelias Geocoding plugin*, який дає змогу виконувати геокодування за допомогою віддаленого сервісу *Pelias open-source world geocoder*, доступного на *openrouteservice* та *geocode.earth* [Geocoding, 2022].

Аналіз останніх досліджень

Питання інтеграції геопросторових даних та його тенденції розвитку досліджено у таких роботах [Gao et al., 2005; Silberschatz et al., 2011; Franci et al., 2014; Bhattacharya & Painho, 2017; Mardani et al., 2019; Sun et al., 2019].

Особливості інтеграції геопросторових даних для розвитку НІГД, містобудівного кадастру та інших галузях економіки розглянуто у роботах [Hansen, 1999; Максимова, 2016; Пілічева, та ін., 2018; Lemenkova, 2020; Лященко & Черін, 2019, Лященко та ін., 2020; Лященко та ін., 2021; Станкевич та ін., 2021; Шипулін, 2021].

Обґрунтування та застосування моделей операції з'єднання (Join) реляційної алгебри подано в працях [Codd, 1990; Райордан, 2001; Коннолли, Бегг, 2003; Gao et al., 2005; Silberschatz et al., 2011; Глушко, 2013; Буй та Глушко, 2015]. Загальновідомо, що реляційну алгебру і модель запропонував Е. Кодд (E. Codd) у 70-ті роки ХХ ст. і її покладено в основу багатьох сучасних СКБД і мов запитів, зокрема SQL (англ. Structured query language – мови структурованих запитів), які підтримує реляційна модель. Е. Кодд запропонував дев'ять операцій реляційної алгебри: традиційні операції над множинами, такі як об'єднання, перетин, різниця, та спеціальні операції над таблицями: проєкція, декартове з'єднання, тета- (theta-) та екви- (equi-) з'єднання, ділення, вибірка [Codd, 1990]. Проблему інтегрування геопросторових даних вирішено у міжнародному і гармонізованому національному стандарті ДСТУ ISO 19112:2017 (ISO 19112:2003, IDT). Географічна інформація. Просторова прив'язка за географічними ідентифікаторами [Географічна інформація... ДСТУ ISO 19112:2017, 2017].

У статті [Лазоренко-Гевель, 2021] досліджено предмет, ідею, роль та значення географічних ідентифікаторів для забезпечення інтеграції геопросторових даних у цілісних базах топографічних даних та національній інфраструктурі геопросторових даних відповідно до національного стандарту ДСТУ ISO 19112:2017 (ISO 19112:2003, IDT). Географічна інформація. Просторова прив'язка за географічними ідентифікаторами. Метою національного стандарту ДСТУ ISO 19112:2017 (ISO 19112:2003, IDT) є способи визначення та описання систем просторової прив'язки концептуального рівня із використанням географічних ідентифікаторів. Місцеположення об'єкта ідентифікується за допомогою просторової прив'язки. Коли географічний ідентифікатор застосовується для просторової прив'язки, він забезпечує однозначну унікальну ідентифікацію місцеположення

об'єкта. Це місцеположення є об'єктом, використовуваним для прив'язки інших об'єктів [Лященко та ін., 2020; Лазоренко-Гевель, 2021].

Разом з тим у науковій літературі питання інтеграції геопросторових даних за допомогою операції з'єднання JOIN в об'єктно-реляційних системах керування базами даних (ОР СКБД) розглянуто недостатньо глибоко.

Мета

Метою статті є дослідження інтеграції наборів базових і тематичних геопросторових даних на основі операції з'єднання (Join) реляційної алгебри та її взаємодії із геокодуванням геопросторових об'єктів, яку реалізовано в сучасних ГІС та СКБД для розвитку національної інфраструктури геопросторових даних.

Методика досліджень

Основа дослідження – аналіз можливостей застосування теорії баз геопросторових даних і баз знань, міжнародних і національних гармонізованих стандартів у сфері “Географічна інформація/Геоматика” для вирішення питання інтеграції геопросторових даних за допомогою операції з'єднання Join реляційної алгебри в об'єктно-реляційних системах керування базами даних (ОР СКБД).

Результати досліджень

Моделі операції з'єднання (Join) реляційної алгебри. З'єднання (Join) в нотації SQL – це операція об'єднання атрибутів з однієї або декількох таблиць у нову таблицю із застосуванням операцій декартового добутку й вибірки. В реляційній алгебрі з'єднанням двох відношень за конкретною умовою називається відношення $(A \text{ TIMES } B) \text{ WHERE } C$, де A та B – будь-які відношення, C – логічний вираз, до якого можуть входити атрибути відношень A та B і/або скалярні вирази (рис. 1).

Операція з'єднання є результатом послідовного застосування операцій декартового добутку й вибірки. Якщо у відношеннях A та B є атрибути із однаковими найменуваннями, то перед виконанням з'єднання такі атрибути перейменовують [Коннолли, Бегг, 2003].

Операція з'єднання в нотації SQL на основі реляційної алгебри поділяється на такі види: внутрішнє (*Inner Join*), зовнішнє (*Outer Join*), перехресне (*Cross Join*), самоз'єднання (*Self-Join*) і просторове (*Spatial Join*).

До внутрішнього належить: тета- (θ -*Join*), екві- (*Equi-Join*), напів- (*Semi-Join*) та натуральне (*Natural Join*) з'єднання. До зовнішнього: ліве (*Left Join*), праве (*Right Join*) та повне зовнішнє (*Full Join*). Окремим випадком еквіз'єднання є натуральне, а окремим випадком напівз'єднання є антиз'єднання (*Anti-Join*). Класифікацію видів операції з'єднання подано на UML-діаграмі, між класами операцій встановлено зв'язок успадкування (рис. 2). У СКБД та ГІС є додаткові можливості та функції для роботи з геопросторовими даними: просторові типи даних; просторові оператори; просторові прикладні процедури та просторове індексування.

Ці можливості дають змогу використовувати операцію просторового з'єднання, враховуючи геометрію об'єктів разом з їх ідентифікаторами.

У цій статті досліджено можливість використання операції з'єднання (JOIN) реляційної алгебри для інтеграції наборів базових і тематичних геопросторових даних в ОРСКБД PostgreSQL. Базовими геопросторовими даними (Відношення А) прийнято відомості про адміністративно-територіальні одиниці Черкаської області, зокрема їх межі, на 2019 р. [Геопортал... <http://atu.gki.com.ua/>], які містять такі атрибути: КОАТУУ (*coatuu*), Геометрія (*geom*), Назва_укр (*name_ua*) тощо. Фрагмент відношення А подано в табл. 1, де: КОАТУУ – Класифікатор об'єктів адміністративно-територіального устрою України; Геометрія (*geometry*) – атрибут, який містить координатні описи геометричних елементів просторових об'єктів у форматі WKT (Well-Known Text); Назва_укр – атрибут, який позначає українську назву об'єкта адміністративно-територіального устрою. Зазначимо, що у відношенні А наявні також новоутворені райони, яким присвоєно не КОАТУУ, а КАТОТТГ відповідно до Постанови Верховної Ради України “Про утворення та ліквідацію районів” від 17.07.2020 р. № 807-IX. Тому в атрибуті КОАТУУ наявні псевдозначення NULL у кортежах новоутворених районів. Кількість кортежів у відношенні А – 30 записів.

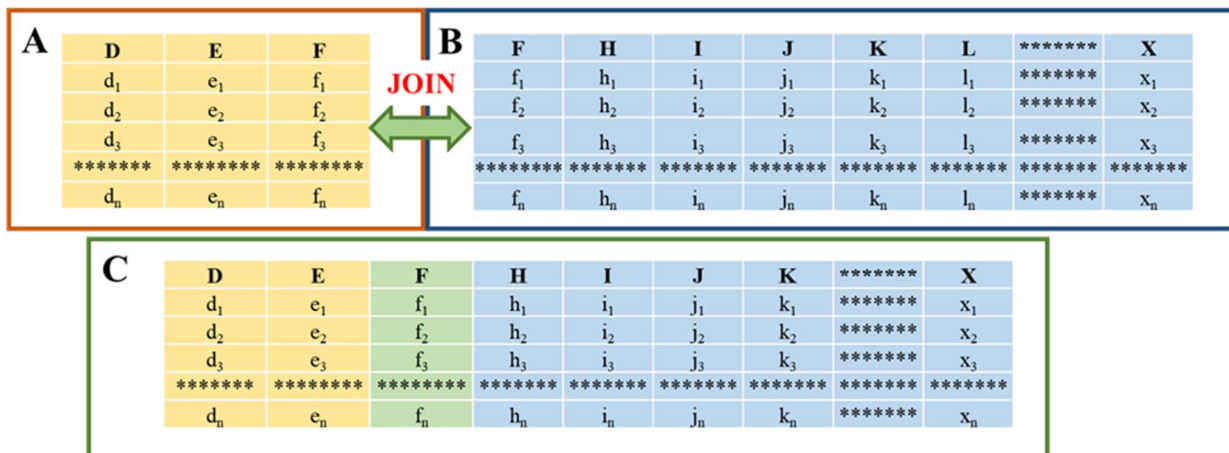


Рис. 1. Інтегрування (з'єднання – Join) наборів базових (відношення А – жовтий колір) і тематичних (відношення В – блакитний колір) геопросторових даних у базі геопросторових даних [Системна модель, 2020]

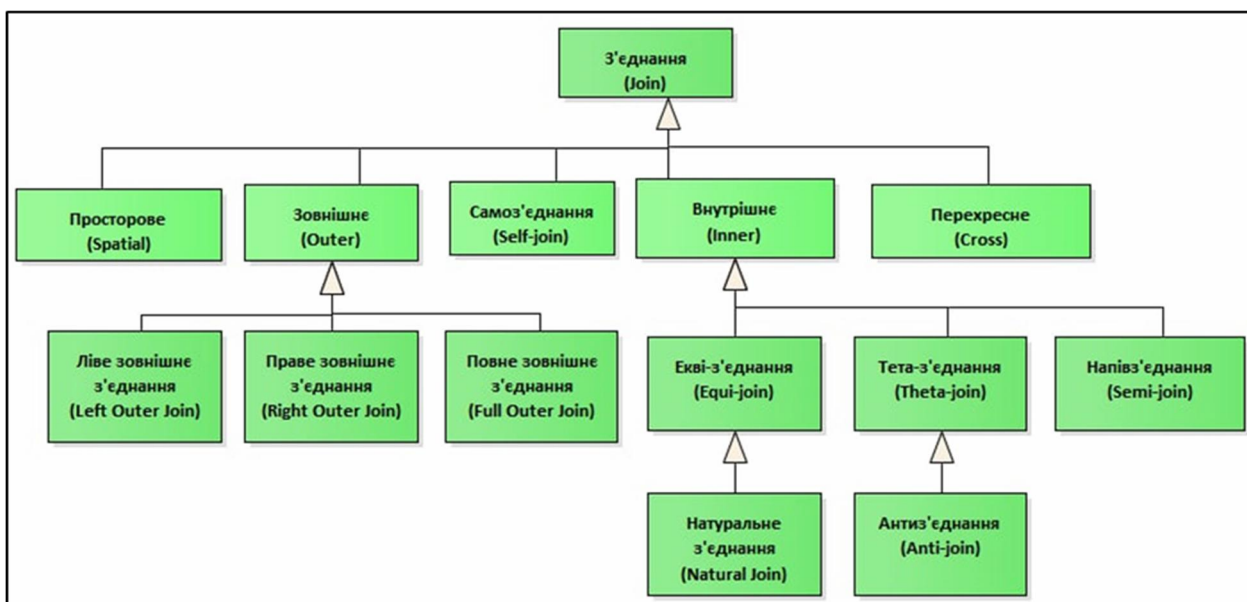


Рис. 2. UML-діаграма класифікації операції з'єднання

Таблиця 1

Фрагмент відношення А. Дані щодо районів Черкаської області

КОАТУУ coatuu	Геометрія geom	Назва району name_ua
7121500000	MULTIPOLYGON(((6392148.42197378 5478804.02538396...)))	Золотоніський
7121200000	MULTIPOLYGON(((6380005.8326648 5402075.15493481...)))	Звенигородський
7120600000	MULTIPOLYGON(((63356005.8326648 5412565.15493481...)))	Драбівський
7120900000	MULTIPOLYGON(((63569775.8326648 5434895.15493481...)))	Жашківський
NULL	MULTIPOLYGON(((6386787.50122662 5493188.28142229...)))	Черкаський
NULL	MULTIPOLYGON(((6450505.70254616 5545812.54783928...)))	Уманський

За тематичні дані (Відношення В) прийнято дані зі статистичного бюлетеня соціально-економічного становища Черкаської області за січень 2021 р. Головного управління статистики у Черкаській області Державної служби статистики України, а саме показники демографічної ситуації. Фрагмент відношення В подано в табл. 2 [Статистичний бюлетень, 2021]. Державна служба статистики України ще використовує Класифікатор об'єктів адміністративно-територіального устрою України (КОАТУУ) для збирання і зберігання статистичних даних і поступово переходить на Кодифікатор адміністративно-територіальних одиниць та територій територіальних громад (КАТОТТГ) у зв'язку з прийняттям Постанови Верховної Ради України від 17.07.2020 № 807-ІХ “Про утворення та ліквідацію районів”, тому КОАТУУ втратив актуальність. Міністерство розвитку громад та територій розробило та затвердило Кодифікатор адміністративно-територіальних одиниць та територій територіальних громад (наказ від 26.11.2020 р. № 290 зі змінами, внесеними наказом від 12.01.2021 р. № 3), який відповідає чинному законодавству і запроваджується на заміну КОАТУУ [Класифікатор, 2022].

Кількість кортежів у відношенні В становить 26 записів.

КОАТУУ та КАТОТТГ є непрямими або географічними ідентифікаторами, які наявні у відношеннях А і В та за допомогою яких у статті досліджено з'єднання геопросторової складової базових геопросторових даних з атрибутивною тематичних геопросторових даних, які не містять координат.

Якщо потрібно виконати звичайне з'єднання за однаковим кортежем “КОАТУУ”, а значення його атрибутів однакові, тоді використовують операції натурального та зовнішніх з'єднань.

Натуральне з'єднання. Нехай дано відношення $A(A_1, A_2, \dots, A_n, X_1, X_2, \dots, X_p)$ та $B(X_1, X_2, \dots, X_p, B_1, B_2, \dots, B_m)$, які мають атрибути X_1, X_2, \dots, X_p з однаковими іменами та визначені на однакових доменах. Тоді *натуральним з'єднанням* відношень А та В називають відношення із заголовком $(A_1, A_2, \dots, A_n, X_1, X_2, \dots, X_p, B_1, B_2, \dots, B_m)$ і тілом, що містить множину кортежів $(a_1, a_2, \dots, a_n, x_1, x_2, \dots, x_p, b_1, b_2, \dots, b_m)$, таких, що $(a_1, a_2, \dots, a_n, x_1, x_2, \dots, x_p) \in A$ та $(x_1, x_2, \dots, x_p, b_1, b_2, \dots, b_m) \in B$.

Таблиця 2

Фрагмент відношення В. Дані щодо соціально-економічного становища Черкаської області за січень 2021 р.

КОАТУУ coatuu	Назва району name_ua	Наявне населення existing population	Постійне населення permanent population	Кількість живона- роджених number of births	Кількість померлих number of deaths	Кількість прибулих number of arrivals
7121500000	Золотоніський	39031	39339	218	824	441
7121200000	Звенигородський	41845	41887	282	804	503
7120600000	Драбівський	32602	32806	154	717	269
7120900000	Жашківський	34831	34726	235	756	430

Для натурального з'єднання використовують синтаксис: $A \text{ JOIN } B$. *Натуральне з'єднання* має властивість *асоціативності*, тобто $(A \text{ JOIN } B) \text{ JOIN } C = A \text{ JOIN } (B \text{ JOIN } C)$, тому такі з'єднання записують, опускаючи дужки: $A \text{ JOIN } B \text{ JOIN } C$ [Коннолли, Бегг, 2003]. Для відображення результату *натурального з'єднання*

виконано вибірку за всіма атрибутами таблиці бази геопросторових даних:

```
SELECT coatuu, geom, name_ua, existing
population, permanent population, number of births,
number of deaths, number of arrivals
```

```
FROM public.cherkasy
NATURAL JOIN public.stat
```

Внутрішнє з'єднання, до якого належить натуральне, повернуло записи, для яких виконувалась умова з'єднання, тобто записи, в яких збігаються значення визначених атрибутів (табл. 3).

Виконання операції з'єднання уможливило побудову різноманітних тематичних карт за різними приєднаними атрибутами. Наприклад, на рис. 3 подано тематичну карту районів Черкаської області за атрибутом "Кількість живонароджених".

Зовнішнє з'єднання повертає всі записи, які повертає внутрішнє з'єднання, а також усі записи з одного чи двох наборів даних, що беруть участь у з'єднанні. Відсутні значення, для яких не існує відповідності, замінюються значеннями *Null*.

Зовнішні з'єднання можна розділити на декілька груп залежно від того, які саме додаткові записи входять до них: ліві, праві та повні. **Ліве зовнішнє з'єднання** (left outer join) повертає всі значення з лівої таблиці й додає значення атрибутів із правої таблиці або NULL, якщо немає збігу за предикатом з'єднання, а **праве зовнішнє**

з'єднання (right outer join) – повертає усі значення із правої таблиці. **Повне зовнішнє з'єднання** (full outer join) сполучає результати лівого та правого зовнішніх з'єднань. Результуюча таблиця містить усі записи із обох таблиць, позначаючи NULL-значеннями відсутність збігів із кожної таблиці [Коннолли, Бегг, 2003]. Нижче в табл. 4–6 подано результати запитів із використанням операцій лівого (LEFT OUTER JOIN), правого (RIGHT OUTER JOIN) та повного зовнішнього з'єднання (FULL OUTER JOIN).

Важливо зазначити, що картографічне зображення цих результатів ідентичне (рис. 3), проте різна послідовність приєднання атрибутів таблиць, які з'єднуються, внаслідок чого запити повертають різну кількість записів. Наприклад, у разі лівого та повного з'єднання ураховують значення першої таблиці *public.cherkasy*. Результати цих запитів містять значення NULL, оскільки атрибути КОАТУУ для чотирьох новостворених районів не вказано, тому з'єднання не відбулось.

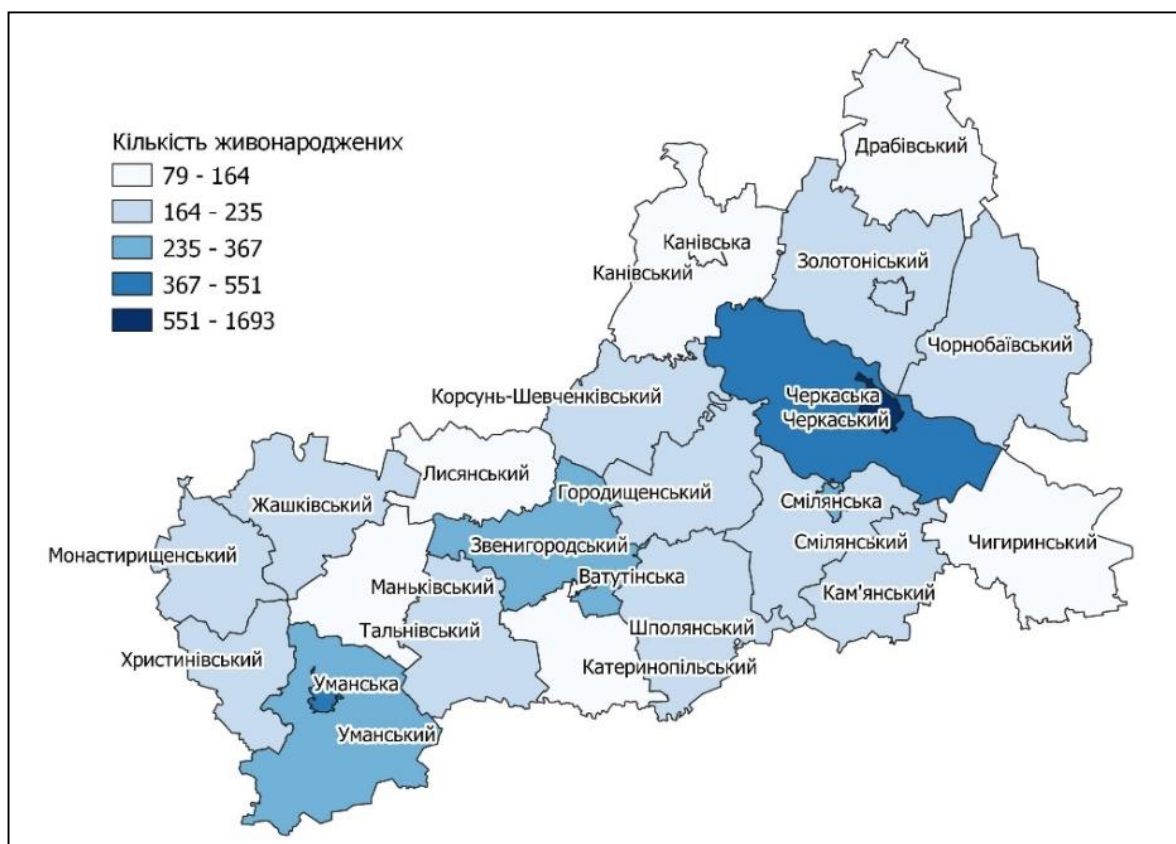


Рис. 3. Тематична карта, отримана в результаті лівого, правого та повного зовнішнього з'єднання

Результат натурального з'єднання

КОАТУУ coatuu	Геометрія geom	Назва району name_ua	Наявне населення existing population	Постійне населення permanent population	Кількість живонароджених number of births	Кількість померлих number of deaths	Кількість прибулих number of arrivals
7120600000	MULTIPOLYGON (((6392148.42197378 5478804.02538396...)))	Драбівський	32602	32806	154	717	269
7120900000	MULTIPOLYGON (((6380005.8326648 5402075.15493481...)))	Жашківський	34831	34726	235	756	430
7121200000	MULTIPOLYGON (((63356005.8326648 5412565.15493481...)))	Звенигородський	41845	41887	282	804	503
7121500000	MULTIPOLYGON (((63569775.8326648 5434895.15493481...)))	Золотоніський	39031	39339	218	824	441

SELECT coatuu, geom, name_ua, existing population, permanent population, number of births, number of deaths, number of arrivals

FROM public.cherkasy

LEFT OUTER JOIN public.stat

ON cherkasy.coatuu = stat.coatuu;

SELECT coatuu, geom, name_ua, existing population, permanent population, number of births, number of deaths, number of arrivals

FROM public.cherkasy

RIGHT OUTER JOIN public.stat

ON cherkasy.coatuu = stat.coatuu

SELECT coatuu, geom, name_ua, existing population, permanent population, number of births, number of deaths, number of arrivals

FROM public.cherkasy

FULL OUTER JOIN public.stat

ON cherkasy.coatuu = stat.coatuu

Запит із застосуванням операції тета-з'єднання має містити умову. Наприклад, значення атрибута "Постійне населення" (permanent

population) відношення В повинно бути більшим від значення атрибута "Населення" (population) відношення А або дорівнювати йому:

SELECT coatuu, geom, name_ua, existing population, permanent population, number of births, number of deaths, number of arrivals

FROM public.cherkasy

INNER JOIN public.stat **ON** cherkasy.coatuu = stat.coatuu

WHERE stat.permanent population >= population

Результат тета-з'єднання у вигляді тематичної карти районів Черкаської області за атрибутом "наявне населення" подано на рис. 4.

Еквіз'єднання. Окремий випадок тета-з'єднання, коли θ є просто рівністю. Синтаксис еквіз'єднання: $A[X = Y]B$ [Коннолли, Бегг, 2003].

Запит із застосуванням операції еквіз'єднання має містити умову. Наприклад, значення атрибута "Постійне населення" (permanent

population) відношення В повинно дорівнювати значенню атрибута “Населення” (population) відношення А:

```

SELECT coatuu, geom, name_ua, existing
population, permanent population, number of births,
number of deaths, number of arrivals
FROM public.cherkasy
INNER JOIN public.stat
ON cherkasy.coatuu = stat.coatuu
WHERE stat.permanent population = population
    
```

Результат еквіз’єднання у вигляді тематичної карти районів Черкаської області за атрибутом “Постійне населення” подано рис. 5. Еквіз’єднання має недолік: якщо воно відбувається за атрибутами з однаковими назвами, то в результатуючому відношенні з’являються два атрибути з однаковими значеннями. Позбутися цього недоліку можна, взявши проєкцію за всіма атрибутами, крім одного із тих, що дублюють, або використати операцію натурального з’єднання.

Таблиця 4

Результат лівого зовнішнього з’єднання

КОАТУУ coatuu	Геометрія geom	Назва району name_ua	Наявне населення existing population	Постійне населення permanent population	Кількість живонароджених number of births	Кількість померлих number of deaths	Кількість прибулих number of arrivals
7121500000	MULTIPOLYGON (((6392148.42197378 5478804.02538396...)))	Золото- тоніський	39031	39339	218	824	441
7121200000	MULTIPOLYGON (((6380005.8326648 5402075.15493481...)))	Звени-го- родський	41845	41887	282	804	503
7120600000	MULTIPOLYGON (((63356005.8326648 5412565.15493481...)))	Драбівсь- кий	32602	32806	154	717	269
7120900000	MULTIPOLYGON (((63569775.8326648 5434895.15493481...)))	Жаш- ківський	34831	34726	235	756	430
NULL	MULTIPOLYGON (((6386787.50122662 5493188.28142229...)))	Черкась- кий	NULL	NULL	NULL	NULL	NULL
NULL	MULTIPOLYGON (((6450505.70254616 5545812.54783928...)))	Умансь- кий	NULL	NULL	NULL	NULL	NULL

Таблиця 5

Результат правого зовнішнього з'єднання

КОАТУУ coatuu	Геометрія geom	Назва району name_ua	Наявне населення existing population	Постійне населення permanent population	Кількість живонаро- джених number of births	Кількість померлих number of deaths	Кількість прибулих number of arrivals
7121500000	MULTIPOLYGON (((6392148.42197378 5478804.02538396...)))	Золо- тоніський	39031	39339	218	824	441
7121200000	MULTIPOLYGON (((6380005.8326648 5402075.15493481...)))	Звениго- родський	41845	41887	282	804	503
7120600000	MULTIPOLYGON (((63356005.8326648 5412565.15493481...)))	Драбівсь- кий	32602	32806	154	717	269
7120900000	MULTIPOLYGON (((63569775.8326648 5434895.15493481...)))	Жаш- ківський	34831	34726	235	756	430

Таблиця 6

Результат повного зовнішнього з'єднання

КОАТУУ coatuu	КОАТУУ coatuu	Геометрія geom	Назва району name_ua	Наявне населення existing population	Постійне населення permanent population	Кількість живонаро- джених number of births	Кількість померлих number of deaths	Кількість прибулих number of arrivals
7121500000	7121500000	MULTIPOLYGON (((6392148.42197378 5478804.02538396...)))	Золотонісь- кий	39031	39339	218	824	441
7121200000	7121200000	MULTIPOLYGON (((6380005.8326648 5402075.15493481...)))	Звенигород- ський	41845	41887	282	804	503
7120600000	7120600000	MULTIPOLYGON (((63356005.8326648 5412565.15493481...)))	Драбівський	32602	32806	154	717	269
7120900000	7120900000	MULTIPOLYGON (((63569775.8326648 5434895.15493481...)))	Жашківсь- кий	34831	34726	235	756	430
NULL	NULL	MULTIPOLYGON (((6386787.50122662 5493188.28142229...)))	Черкаський	NULL	NULL	NULL	NULL	NULL
NULL	NULL	MULTIPOLYGON (((6450505.70254616 5545812.54783928...)))	Уманський	NULL	NULL	NULL	NULL	NULL

Таблиця 7

Результат тета-з'єднання

КОАТУУ coatuu	Геометрія geom	Назва району name_ua	Наявне населення existing population	Постійне населення permanent population	Кількість живона- роджених number of births	Кількість померлих number of deaths	Кількість прибулих number of arrivals
7121500000	MULTIPOLYGON (((6392148.42197378 5478804.02538396...)))	Золо- тоніський	39031	39339	218	824	441
7121200000	MULTIPOLYGON (((6380005.8326648 5402075.15493481...)))	Звени- городський	41845	41887	282	804	503
7120600000	MULTIPOLYGON (((63356005.8326648 5412565.15493481...)))	Драбівський	32602	32806	154	717	269
7120900000	MULTIPOLYGON (((63569775.8326648 5434895.15493481...)))	Жаш- ківський	34831	34726	235	756	430

Таблиця 8

Результат еквіз'єднання

КОАТУУ coatuu	Геометрія geometry	Назва_укр name_ua	Наявне населення existing population	Постійне населення permanent population	Кількість живонароджених number of births	Кількість померлих number of deaths	Кількість прибулих number of arrivals
7121500000	MULTIPOLYGON (((6392148.42197378 5478804.02538396...)))	Золото- ніський	39031	39339	218	824	441
7121200000	MULTIPOLYGON (((6380005.8326648 5402075.15493481...)))	Звени- город ський	41845	41887	282	804	503
7120600000	MULTIPOLYGON (((63356005.8326648 5412565.15493481...)))	Драбів- ський	32602	32806	154	717	269
7120900000	MULTIPOLYGON (((63569775.8326648 5434895.15493481...)))	Жашків ський	34831	34726	235	756	430

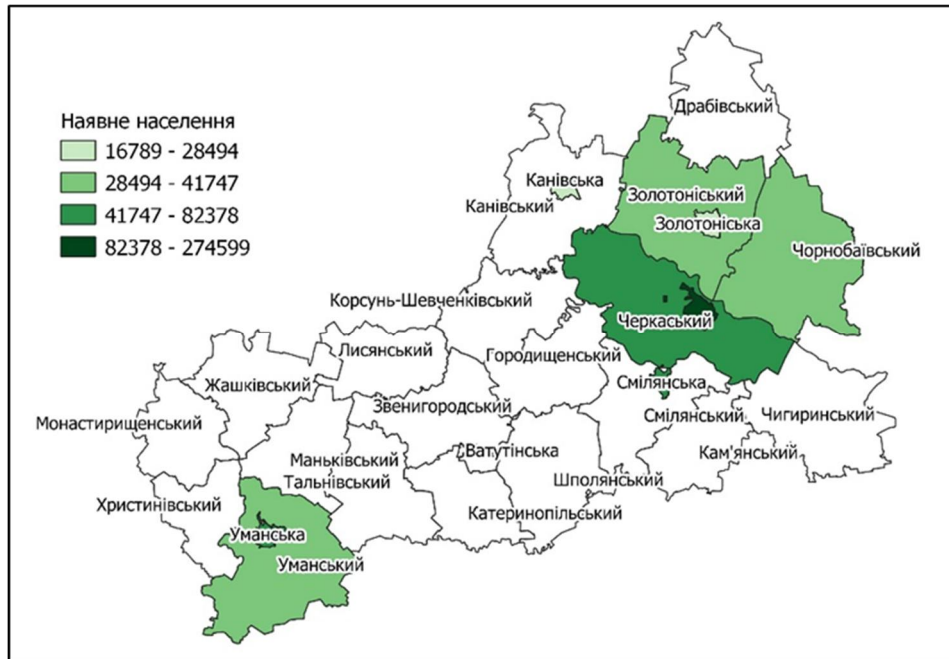


Рис. 4. Тематична карта, отримана в результаті тета-з'єднання

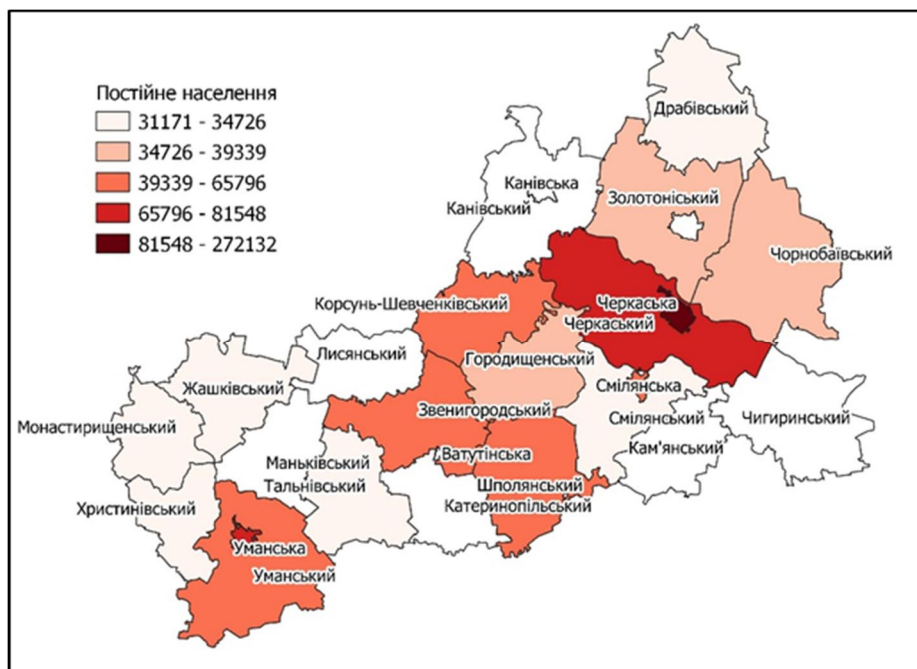


Рис. 5. Тематична карта, отримана в результаті еквіз'єднання за атрибутом "Постійне населення"

Напівз'єднання. Операція напівз'єднання визначає відношення, яке містить ті кортежі відношення А, які входять до з'єднання відношень А та В. Перевага напівз'єднання полягає в тому, що воно дає змогу зменшити кількість кортежів, які потрібно опрацювати, щоб отримати з'єднання [Конноллі та Бегг, 2003].

Запит із застосуванням операції напівз'єднання має містити умову. Наприклад, значення атрибута

"Постійне населення" (permanent population) відношення В повинно дорівнювати значенню атрибута "Населення" (population) відношення А й атрибут "Кількість прибулих" (number of arrivals) відношення В повинен бути більшим від 350:

```
SELECT coatuu, geom, name_ua, existing
population, permanent population, number of births,
number of deaths, number of arrivals
```

```
FROM public.cherkasy
```

INNER JOIN public.stat **ON** cherkasy.coatuu= **WHERE** stat.permanent population = stat.coatuu population and number of arrivals > 350

Таблиця 9

Результат напівз'єднання

КОАТУУ coatuu	КОАТУУ coatuu	Геометрія geometry	Назва_укр name_ua	Наявне населення existing population	Постійне населення permanent population	Кількість живонароджених number of births	Кількість померлих number of deaths	Кількість прибулих number of arrivals
7121500000	7121500000	MULTI POLYGON (((6392148.42197378 5478804.02538396...)))	Золотонісь- кий	39031	39339	218	824	441
7121200000	7121200000	MULTI POLYGON (((6380005.8326648 5402075.15493481...)))	Звени- городський	41845	41887	282	804	503
7120900000	7120900000	MULTI POLYGON (((63569775.8326648 5434895.15493481...)))	Жашківсь- кий	34831	34726	235	756	430

У результаті виконання різних видів з'єднань відношень А і В виявлено, що для успішного їх виконання необхідно, щоб кожен об'єкт, який з'єднується, мав унікальний ідентифікатор. Крім того, кожен вид з'єднання має певні особливості: застосування умов, послідовність з'єднання та відповідно отримані результати виконання запитів.

Доведено, що операцію з'єднання (JOIN) реляційної алгебри можна використовувати для інтеграції інших тематичних геопросторових даних із базовими геопросторовими даними за допомогою географічних ідентифікаторів, які містять ці набори даних.

Наукова новизна та практична значущість

Досліджено інтеграцію наборів базових і тематичних геопросторових даних на основі моделей операції з'єднання (Join) реляційної алгебри та їх взаємодії з геокодуванням геопросторових об'єктів, яку реалізовано в сучасних ГІС та СКБД для розвитку національної інфраструктури геопросторових даних. Дослідження виконано на наборі базових геопросторових даних, а саме: відомостей про адміністративно-територіальні одиниці Черкаської області, зокрема їх меж; тематичними вибрано дані зі статистичного бюлетеня соціально-економічного становища Черкаської області за січень 2021 р. Головного управління статистики у Черкаській області Державної служби статистики України.

Висновки

Прийняття Закону України “Про національну інфраструктуру геопросторових даних” № 554 від 13.04.2020 р., Постанови Кабінету Міністрів України “Про затвердження Порядку функціонування національної інфраструктури геопросторових даних” № 532 від 26.05.2021 р. та Наказу Міністерства аграрної політики і продовольства України “Про затвердження технічних вимог до геопросторових даних, метаданих і геоінформаційних сервісів національної інфраструктури геопросторових даних” № 347 від 10.11.2021 р. сприяло розвитку Національної інфраструктури геопросторових даних (НІГД) в Україні. Уніфікація та інтеграція базових і тематичних геопросторових даних забезпечує роботу НІГД. Одним із засобів інтегрування даних є

з'єднання різноманітних баз геопросторових даних із даними, які не містять геопросторової складової, за допомогою системного ідентифікатора для їх зведення з різних джерел у НІГД.

Серед усіх методів інтеграції даних важливе місце займає геокодування як просторова прив'язка об'єктів на основі географічних ідентифікаторів, основні вимоги до якої подано у міжнародному і гармонізованому національному стандарті ДСТУ ISO 19112:2017 (ISO 19112:2003, IDT). Географічна інформація. Просторова прив'язка за географічними ідентифікаторами (Geographic information – Spatial referencing by geographic identifiers).

У статті досліджено моделі операції з'єднання Join реляційної алгебри, покладені в основу геокодування об'єктів і створення електронних газетирів, і доведено її ефективність: операція з'єднання Join забезпечує інтеграцію наборів базових і тематичних даних. Для її виконання необхідно визначити обов'язкові географічні ідентифікатори, які повинні бути серед атрибутів наборів базових та тематичних геопросторових даних та за якими виконується з'єднання. Різноманітність видів застосування операції з'єднання Join охоплює всі можливі випадки, які виникають під час їх практичного застосування. Отже, використання операції з'єднання Join передбачає на етапі проектування баз геопросторових даних визначення цих обов'язкових географічних ідентифікаторів.

Зокрема, доцільно визначити обов'язкові географічні ідентифікатори (коди) об'єктів за офіційними загальнодержавними системами класифікації (кодифікації) об'єктів у відповідних галузевих тематичних реєстрах:

1) КОАТУУ (нині – КАТОТТГ) для об'єктів адміністративно-територіального устрою України;

2) ідентифікатори об'єктів дорожньої мережі і дорожніх споруд за реєстрами Міністерства інфраструктури України, АТ “Укрзалізниця”, Державного агентства автомобільних доріг України (Укравтодору);

3) коди річок, водойм, водотоків та басейнів річок – за класифікатором Державного водного кадастру Державного агентства водних ресурсів України;

4) коди електромереж високої напруги відповідно до реєстрів Міністерства енергетики України;

5) коди лісів та рослинності – за реєстром Державного лісового кадастру Державного агентства лісових ресурсів України;

6) ідентифікатори будівель та споруд – за реєстрами об'єктів нерухомості, реєстрами бюро технічної інвентаризації (БТІ) та містобудівного кадастру Міністерства розвитку громад та територій України;

7) ідентифікатори трубопроводів – за реєстрами Міністерства енергетики України, НАК “Нафтогаз України” тощо.

Використання операції з'єднання дасть змогу зменшити витрати на підтримку в базових геопросторових даних атрибутів обов'язкового постійного зберігання, а також забезпечить інтеграцію даних із різноманітних інформаційних ресурсів і, отже, створення різних тематичних геопросторових даних на основі сукупності відомостей, доступних із різних державних реєстрів та баз даних. Інтегрування даних на основі їх з'єднання повинно забезпечити цілісність, достовірність, інформативність і тематичну різноманітність геопросторових даних за рахунок їх інтероперабельності та сумісності із першоджерелами галузевих інформаційних ресурсів.

СПИСОК ЛІТЕРАТУРИ

- Буй, Д. Б., & Глушко, І. М. (2015). Розширення сигнатур реляційних (табличних) алгебр Кодда: сучасний стан. *Наукові записки НаУКМА. Комп'ютерні науки*, (177), 95–107.
- Географічна інформація. Просторова прив'язка за географічними ідентифікаторами: ДСТУ ISO 19112:2017 (ISO 19112:2003, IDT). [Чинний від 2017-10-01]. Київ: ДП “УкрНДНЦ”.
- Геопортал “Адміністративно-територіальний устрій України”. URL: <http://atu.gki.com.ua/> (дата звернення: 20.04.2022).
- Глушко І. М. (2013). Числення та розширення сигнатур табличних алгебр: дис. ... канд. фіз.-мат. наук: спец. 01.05.01. Київ. 142 с. <http://www.cyb.univ.kiev.ua/library/dissertations/hlushko.pdf>
- Карпінський Ю. О., Лященко А. А. (2006). Стратегія формування національної інфраструктури геопросторових даних в Україні. Київ: НДІГК. 108 с.: іл. (Сер. “Геодезія, картографія, кадастр”).
- Карпінський Ю. О., Лазоренко-Гевель Н. Ю. (2018). Методи збирання геопросторових даних для топографічного картографування. *Сучасні досягнення геодезичної науки і виробництва*, Вип. І(35). С. 204–211. <http://gki.com.ua/ua/metodi-zbirannja-geoprostorovich-danih-dlja-topografichnogo-kartografuvannja>
- Карпінський Ю. О., Лазоренко-Гевель Н. Ю. (2020). Системна модель топографічного картографування в національній інфраструктурі геопросторових даних в Україні. *ISTCGCAP*. Вип. 92, С. 24–36. <https://doi.org/10.23939/istcgcap2020.92.024>

- Класифікатор об'єктів адміністративно-територіального устрою України. http://www.ukrstat.gov.ua/klasf/st_kls/op_coatuu_2016.htm (дата звернення: 20.04.2022).
- Коннолли Т., Бегг К. (2003). Базы данных. Проектирование, реализация и сопровождение. Теория и практика / Пер. з англ. 3-е изд. Москва: Вильямс, 1440 с.: ил. Парал. тит. англ. С. 145–149.
- Лазоренко-Гевель Н. (2021). Географічні ідентифікатори як основа для інтеграції геопросторових даних. *Містобудування та територіальне планування*, 78. С. 312–326. <https://doi.org/10.32347/2076-815x.2021.78.312-326>
- Лященко А. А., & Черін А. Г. (2019). Базові моделі та методи інтеграції геопросторових даних в ГІС містобудівного кадастру. *Містобудування та територіальне планування*, 70. С. 351–365. <http://repository.knuba.edu.ua/handle/987654321/6199>
- Лященко А., Гаврилюк Є., & Смілка В. (2020). Аналіз методів унікальної ідентифікації об'єктів в наборах геопросторових даних. *Містобудування та територіальне планування*, 75. С. 217–232. <http://repository.knuba.edu.ua/handle/987654321/9512>
- Лященко А., Карпінський Ю., Гаврилюк Є., & Черін А. (2021). Методи та засоби забезпечення інтероперабельності компонентів національної інфраструктури геопросторових даних. *Містобудування та територіальне планування*, 77. С. 309–319. <https://doi.org/10.32347/2076-815x.2021.77.309-319>
- Максимова Ю. С. (2016). Створення бази даних електронного каталогу класів об'єктів для наборів профільних геопросторових даних містобудівної документації. *Містобудування та територіальне планування*, 62 (1). С. 367–376. <https://repository.knuba.edu.ua/bitstream/handle/987654321/6932/62a-368-377.pdf?sequence=1>
- Національний стандарт України ДСТУ 8774:2018 “Географічна інформація. Правила моделювання геопросторових даних”. URL: <http://gki.com.ua/ua/prinjato-nacionalni-standart-ukraïini-dstu87742018-geografichna-informacija-pravilamodeljuvannja-geoprostorovih-danij>
- Про затвердження технічних вимог до геопросторових даних, метаданих і геоінформаційних сервісів національної інфраструктури геопросторових даних: Наказ Міністерства аграрної політики та продовольства України від 10.11.2021 р. № 345. Зареєстровано в Міністерстві юстиції України 12 січня 2022 р. за № 21/37357.
- Пілічева М. О., Кінь Д. О., & Поморцева О. Є. (2018). Інтеграція топографічної і кадастрової інформації базового набору геопросторових даних земельної ділянки. *Містобудування та територіальне планування*, 66. С. 523–531.
- Постанова Кабінету Міністрів України “Про затвердження Порядку загальнодержавного топографічного і тематичного картографування” від 4 вересня 2013 р. № 661.
- Постанова Кабінету Міністрів України “Про затвердження Порядку функціонування національної інфраструктури геопросторових даних” № 532 від 26 травня 2021 р. <https://zakon.rada.gov.ua/laws/show/532-2021-%D0%BF#Text>
- Про національну інфраструктуру геопросторових даних: Закон України від 13 квітня 2020 року № 554-IX.
- Райордан Р. (2001). Основы реляционных баз данных / пер. з англ. Москва: Русская редакция. 384 с.: ил.; С. 99–104.
- Станкевич С., Титаренко О., & Голубов С. (2021). Математическая модель интеграции гетерогенных данных при оценивании нефтегазоперспективности территорий. *Херсон2021*, 86.
- Соціально-економічне становище Черкаської області: статистичний бюлетень. Головне управління статистики у Черкаській області. (2021). http://www.ck.ukrstat.gov.ua/?p=bul_soc_ek (дата звернення: 20.04.2022).
- Шипулін В. Д. (2021) Інтегрована інформаційна система нерухомості. Концепція для України: монографія. Харківський нац. університет міського господарства ім. О. М. Бекетова. <http://eprints.kname.edu.ua/57436/>
- Bhattacharya, D., & Painho, M. (2017). Smart cities intelligence system (smacisys) integrating sensor web with spatial data infrastructures (sensdi). *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4, 21–28. <https://run.unl.pt/handle/10362/28046>
- Codd, E. F. (1990). The relational model for database management: version 2. Addison-Wesley Longman Publishing Co., Inc.
- ESRI's Geodatabase Website. <http://support.esri.com/datamodels> (дата звернення: 20.04.2022).
- ESRI's Download Website. http://www.esri.com/data/download/census2000_tigerline/index.html (дата звернення: 20.04.2022).
- Franci, F., Lambertini, A., & Bitelli, G. (2014, August). Integration of different geospatial data in urban areas: a case of study. In *Second International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2014)*, Vol. 9229, p. 92290P. International Society for Optics and Photonics. <https://doi.org/10.1117/12.2066614>
- Gao, D., Jensen, C. S., Snodgrass, R. T., & Soo, M. D. (2005). Join operations in temporal databases. *The VLDB Journal*, 14(1), 2–29. <https://link.springer.com/article/10.1007/s00778-003-0111-3>
- Geocoding: Longitude and Latitude by Address. <https://gisgeography.com/geocoding/> (дата звернення: 20.04.2022).

- Hansen, H. S. (1999, April). Integrating digital maps and administrative registers-Danish experiences. In 21 st Urban Data Management Symposium, 21–23.
- How to Geocode in ArcMap. <https://libraries.mit.edu/files/gis/geocoding.pdf> (дата звернення: 20.04.2022).
- Karpinskyi, Y., & Lazorenko-Hevel, N. (2020). Topographic mapping in the National Spatial Data Infrastructure in Ukraine. In E3S Web of Conferences, Vol. 171, p. 02004. EDP Sciences. <https://doi.org/10.1051/e3sconf/202017102004>
- Karpinskyi Y., Lazorenko-Hevel N., Kin D. (2020). INSPIREID implementation in the topographic database of the main state topographic map of Ukraine. *Вєб ISTCGCAP*, Vol. 91, No. 91, 20–27. <https://doi.org/10.23939/istcgcap2020.91.020>
- Lemenkova, P. (2020). Integration of geospatial data for mapping variation of sediment thickness in the North Sea. *Scientific Annals of the Danube Delta Institute*, 25, 129–138. <https://doi.org/10.7427/DDI.25.14>
- Mardani, M., Mardani, H., De Simone, L., Varas, S., Kita, N., & Saito, T. (2019). Integration of machine learning and open access geospatial data for land cover mapping. *Remote Sensing*, 11(16), 1907. <https://doi.org/10.3390/rs11161907>
- Silberschatz, A., Korth, H. F., & Sudarshan, S. (2002). Database system concepts, Vol. 5. New York: McGraw-Hill. 1376 p. <https://sncourseware.org/snctnew/files/1581236100.pdf>
- Sun, K., Zhu, Y., Pan, P., Hou, Z., Wang, D., Li, W., & Song, J. (2019). Geospatial data ontology: the semantic foundation of geospatial data integration and sharing. *Big Earth Data*, 3(3), 269–296. <https://doi.org/10.1080/20964471.2019.1661662>

Nadiia LAZORENKO

Department of Geoinformation system and photogrammetry, Kyiv National University of Construction and Architecture, 31, Povitroflotsky Ave, Kyiv, 03037, Ukraine, e-mail: nadiialg@gmail.com, <https://orcid.org/0000-0002-1572-49471>

INTEGRATION OF GEOSPATIAL DATA BASED ON THE APPLICATION OF THE JOIN OPERATION OF RELATIVE ALGEBRA

The purpose of this work is to study the integration of sets of core reference and thematic geospatial data based on the Join operation of relational algebra and its interaction with geocoding of geospatial features, which is implemented in modern geographic information systems (GIS) and database management systems (hereinafter – DBMS) for the development of the national spatial data infrastructure (hereinafter – NSDI). *Method*. The research is based on the analysis of the possibilities of applying the theory of geospatial databases and knowledge bases, international and national harmonized standards in the field of Geographic Information/Geomatics to solve the problem of integration of geospatial data using the operation JOIN relational algebra in object-relational database management systems (OR DBMS). *Results*. The paper examines the models of the Join operation of relational algebra, which underlie the geocoding of features and the creation of electronic gazetteers, and proves its effectiveness: the Join operation integrates of core reference and thematic geospatial datasets. There is a need to define the required geographic identifiers, which must be present among the attributes of the core reference and thematic geospatial datasets to perform the join. The variety of uses of the Join operation covers all possible cases that arise in their practical application. Thus, the use of the Join operation involves identifying these required geographic identifiers at the geospatial database design stage. In particular, it is expedient to determine mandatory geographical identifiers (codes) of features according to the official national systems of features classification (codification) in the relevant sectoral thematic registers, which are responsible for certain holders of thematic data in accordance with Annex 2 of the Decree of Cabinet of Ministers “The order for the functioning of the national spatial data infrastructure” of May 26, 2021, No. 532. Scientific novelty and practical significance. The integration of core reference data and thematic geospatial datasets based on JOIN operation models of relational algebra and their interaction with geocoding of geospatial features is researched, which is implemented in modern GIS and DBMS for the development of national spatial data infrastructure. The research was performed on a set of core reference spatial data, namely: information on administrative-territorial units of the Cherkasy region, including their borders; the data from the statistical bulletin of the socio-economic situation of the Cherkasy region for January 2021 of the Main Department of Statistics in Cherkasy region of the State Statistics Service of Ukraine were selected as thematic data. It has been shown that relational algebra join (JOIN) operations can be used to integrate other thematic geospatial data with core reference data using geographic identifiers that contain these datasets.

Key words: geospatial data integration; interoperability; JOIN operation; national spatial data infrastructure; core reference data; thematic geospatial data.

Надійшла 05.04.2022 р.