

УДК 004.22:005.3

к. геогр. наук, ст. н. с. Путренко В.В.,
Пархоменко А.В., НУ КПП ім. Ігоря Сікорського

ГЕОПРОСТРОВИЙ АНАЛІЗ ТА ВІЗУАЛІЗАЦІЯ ДАНИХ ПРО ВІЙСЬКОВУ КОНФЛІКТОГЕННІСТЬ

Розглянуто підходи до аналізу військових конфліктів в регіонах світу за допомогою геопросторової кластеризації набору даних про світові збройні конфлікти в ГІС. Проаналізовано набір даних UCDP GED про збройні конфлікти у світі, починаючи з 1989 року. Розроблено програмне забезпечення веб-додатку для візуалізації результатів аналізу на віртуальному глобусі в мережі Інтернет.

Ключові слова: військовий конфлікт, геопросторовий аналіз, кластеризація, геовізуалізація, 3d-глобус..

Вступ. Останніми роками в світі збільшується кількість регіональних збройних конфліктів, і постає питання аналізу причин їх виникнення та динаміки розповсюдження. Найбільш конфліктогенними регіонами світу є країни Середнього і Близького Сходу, африканського континенту. Навіть після незліченної кількості кровопролитних конфліктів фахівцям досі важко визначити, які чинники підвищують ризик збройних конфліктів і воєн, які обставини сприяють вирішенню конфлікту, а які навпаки погіршують ситуацію. Тому аналіз та візуалізація даних про військові конфлікти є актуальною темою в процесі їх вивчення.

В проектах, спрямованих на вивчення, передбачення та попередження виникнення збройних конфліктів, використання засобів розвідувального аналізу є невід'ємною частиною робочого процесу. За допомогою баз даних можна простежити зв'язки та процеси у військових конфліктах, знайти нові фактори, що їх провокують, а також способи для їх запобігання. Використання засобів геовізуалізації даних в якості інструменту для моделювання військових конфліктів, вивчення даних і створення гіпотез, створення інтерактивних карт є сучасним і ефективним підходом, що спрощує роботу з даними та доступ до них усім зацікавленим сторонам проектів.

Огляд літератури. Дослідження військових конфліктів розвиваються на основі політичних, військових та міжнародних досліджень зовнішньополітичної ситуації у світі по відношенню до України та основних макрорегіонів. Подібні дослідження виконані у роботах Костенка Т.Ф., Перепелиці Г. М., Шевченка О. [1,2,4]. Геопросторовий аналіз даних було

побудовано на використанні баз геоданих, методику організації яких розроблено в роботах Ванг Ю., Йін З. [8,9]. Питання кластеризації геопросторових даних досліджено у роботах [3,7,10]. Дослідження баз даних про військові конфлікти з точки зору кластеризації у топологічному просторі потребує окремої розробки.

Метою роботи є розробка інформаційних інструментів аналізу та візуалізації геопросторових даних про збройні конфлікти.

Основними **завданнями** роботи є аналіз існуючих джерел інформації про збройні конфлікти, вибір інтерактивного інструментарію візуалізації геопросторових даних, проведення просторово-часової кластеризації набору даних про військові конфлікти методом “Гарячих точок” та відображення його результатів у вигляді інтерактивних 3D-візуалізацій.

Виклад основного матеріалу. Як вхідні дані було взято масив від проекту Uppsala Conflict Data Program Georeferenced Event Dataset (UCDP GED) – набір геокодованих даних про конфлікти від кафедри дослідження питань миру і конфліктів університету Упсала (Швеція) [5,6]. Метою цього проекту є надання академічній спільноті найбільш повних структурованих даних про події організованого насильства у світі в період з 1989 року, задовольняючи потребу в географічно- і часово-деталізованих даних.

Для кластеризації геопросторових даних було вирішено завдання перетворення даних у формат бази геоданих, геокодування, кластеризація та візуалізація за допомогою сучасних програмних бібліотек. Як СКБД для зберігання даних було обрано MySQL, середовища обробки геоданих – QGIS.

Кластеризацію даних за геопросторовими ознаками було здійснено на основі методу «гарячі та холодних» точок, що базується на аналізі геопросторової статистики.

Більшість статистичних тестів починаються з визначення нульової гіпотези. Нульова гіпотеза для інструментів аналізу структурних закономірностей — це повна просторова хаотичність або самих об'єктів або значень, пов'язаних з ними. Р-значення — це ймовірність, Z-оцінки є стандартними відхиленнями. Z-оцінки і р-значення, отримані в результаті аналізу структурних закономірностей, свідчать про те, чи можна відхилити нульову гіпотезу чи ні. Як правило, використовують один з інструментів аналізу структурних закономірностей, припускаючи, що z-оцінка і р-значення будуть свідчити про можливе спростування нульової гіпотези. Це буде говорити про те, що досліджувані об'єкти або значення, пов'язані з ними, виявляють статистично значущу кластеризацію або дисперсію. Всякий раз, коли спостерігається просторова структура, така як кластеризація ландшафту

(або просторових даних), можна побачити докази роботи деяких основних просторових процесів.

Ключова ідея полягає в тому, що значення в середині нормального розподілу представляють очікуваний результат. Коли абсолютне значення z -оцінки є великим, а ймовірності є маленькими (в хвостах нормального розподілу), спостерігається незвичний розподіл. Для методу “Аналіз гарячих точок” це означає статистично істотну “гарячу” або “холодну” точку.

Метод “Гарячі точки” розраховує статистичний показник для кожної події в наборі даних. Підсумкові p -значення (ймовірності) та z -оцінки (стандартні відхилення) говорять про те, в якій області простору кластеризуються події з високими або низькими значеннями [6]. Метод працює шляхом аналізу кожної події в контексті сусідніх за географією подій. Щоб бути статистично суттєвою гарячою точкою, подія повинна мати високе значення і бути оточена іншими подіями з також високими значеннями. Локальна сума для події та її сусідів порівнюється пропорційно із сумою всіх подій; коли локальна сума дуже відрізняється від очікуваної локальної суми, і коли ця відмінність є занадто великою, щоб бути результатом випадкового процесу, отримується статистично значуща z -оцінка.

Метод гарячих точок використовує формули:

$$G_i^{*s} = \frac{\sum_{j=1}^n w_{ij} x_j - X \overline{\sum_{j=1}^n w_{ij}}}{S \sqrt{\frac{n \sum_{j=1}^n w_{ij}^2 - \left(\sum_{j=1}^n w_{ij}\right)^2}{n-1}}} \quad (1)$$

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - \left(\overline{X}\right)^2} \quad (2)$$

де x_j — значення показника для подій j ; w_{ij} — просторова вага зв'язку між подіями i та j ; n — загальна кількість подій; X — середнє арифметичне значень подій, S — дисперсія.

Статистична величина G_i^* , для кожної події в наборі даних, є z -оцінкою. Для статистично значущих позитивних z -оцінок, чим більша z -оцінка, тим інтенсивніша кластеризація високих значень (гаряча точка). Для статистично значущих негативних z -оцінок, чим менше z -оцінка, тим інтенсивніша кластеризація низьких значень (холодна точка). На виході отримуємо новий набір даних про події із z -оцінкою, p -значеннями та рівнем достовірності G_{i_Bin} для кожної події у вхідному масиві.

Для проведення аналізу даних та їх кластеризації за методом “Гарячих точок” було використано набір UCDP GED у форматі шейп-файлів “ged50-shp” та у вигляді копії (дампу) бази даних MySQL “ged50-mysql”. В ході роботи з набором даних було застосовано програмне забезпечення QGIS з відкритим кодом та комерційну програму ArcGIS від компанії ESRI. З інструментарію ArcGIS для отримання кластеризованих даних було використано вбудований інтерпретатор Python, бібліотеку “arcpy”, функцію генерації просторово-часових вагових коефіцієнтів `GenerateSpatialWeightsMatrix_stats` та власне функцію кластеризації за методом гарячих точок `HotSpots_stats`.

Весь набір даних UCDP GED було розбито за ознакою приналежності до одного із 5 регіонів: Європа, Азія, Північна та Південна Америка, Африка, Близький Схід. Для кожного регіону було обраховано матрицю часово-просторових ваг з параметрами часового діапазону у 5 років та територіального віддалення від точок-сусідів до 500 км.

Використовуючи обчислені матриці ваг, було створено нові набори даних для відповідних регіонів, що містять інформацію про приналежність до гарячих або холодних точок.

Отримані дані було експортовано в окремі шейп-файли, після чого за допомогою програмного забезпечення QGIS імпортувались до бази MySQL. Результуюча база даних має 2 таблиці: власне таблицю даних UCDP GED “ged50-mysql” та таблицю розрахованих z-оцінок, p-значень та значення рівнів достовірності розподілу `Gi_Bin`. Ці таблиці зв'язані за полем унікального ідентифікатора події (рисунки 1 та 2).

Для візуалізації результатів аналізу та надання доступу до них через мережу Інтернет, на клієнтській частині використовуються засоби відкритого проекту WebGL Globe та бібліотеки візуалізації Three.js. Це дозволяє завантажувати дані з сервера у форматі JSON та забезпечувати інтерактивну взаємодію користувача з ними.

Для полегшення взаємодії користувача з додатком спосіб перемикання між роками було реалізовано у вигляді повзунка. Самі події представлені у вигляді геоприв'язаних кольорових стовпців різної величини. Висота стовпця відображає кількість жертв конфлікту, а колір — приналежність до гарячих або холодних точок. При наведенні курсору на стовпець в правому нижньому кутку екрану з'являється коротка довідка про відповідну подію.

На стороні сервера використано СКБД MySQL для збереження результуючого набору даних, веб-сервер Apache для обробки асинхронних запитів клієнтської частини та розроблені програми на мові PHP, що забезпечує бізнес-логіку роботи веб-додатку. Вигляд додатку приведено на рисунку 3.

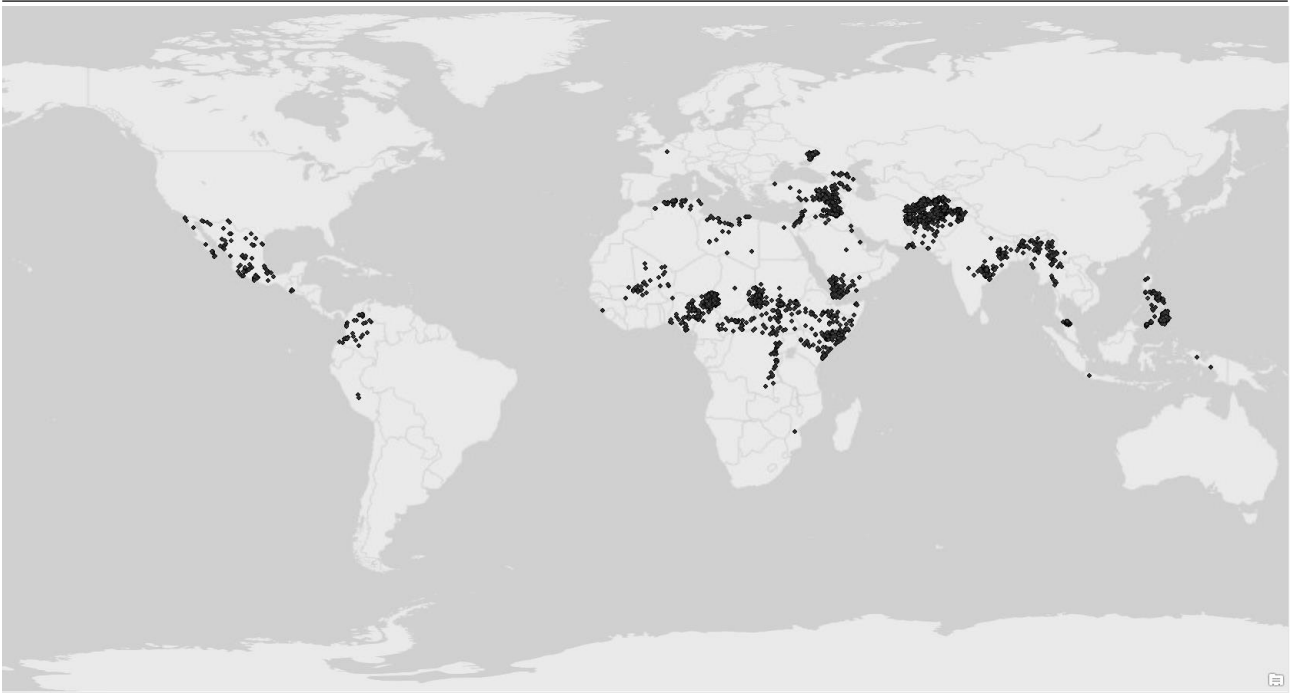


Рис. 1. Розподіл військових конфліктів у світі, 2015 р.

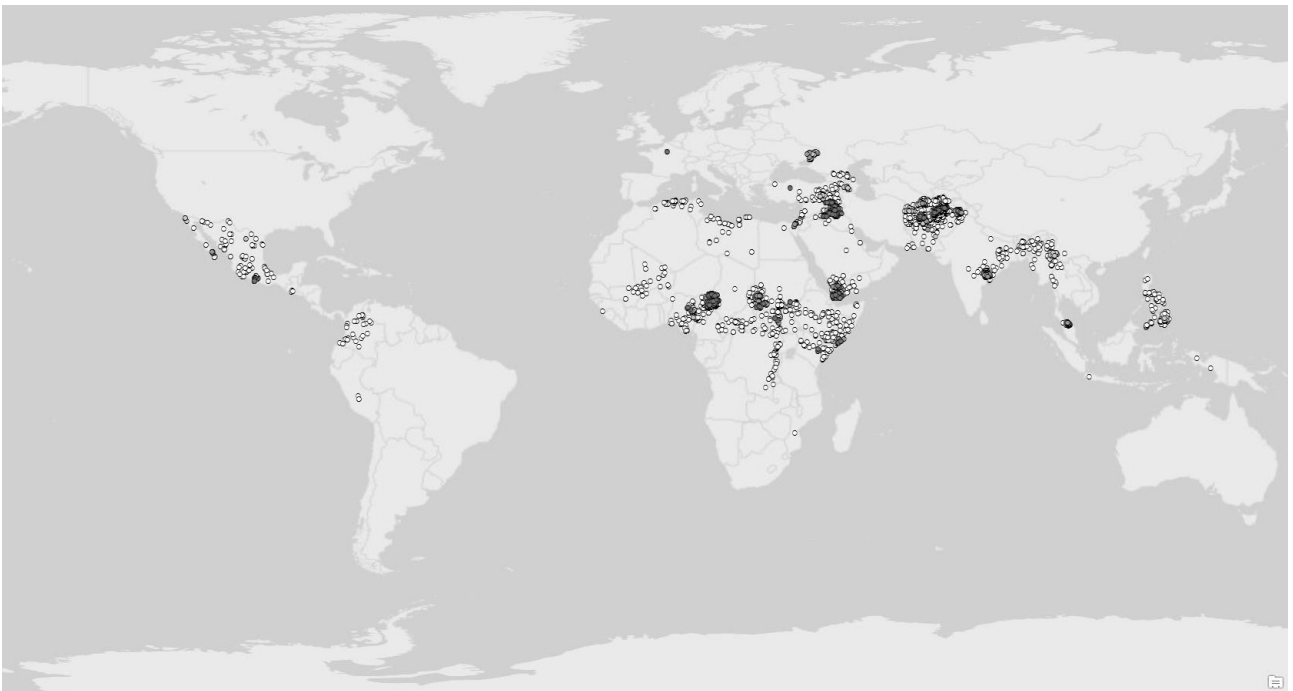


Рис. 2. Кластеризація військових конфліктів у світі, 2015 р.

Інтерпретація результатів кластеризації може допомогти у виявленні просторово-часових закономірностей виникнення та розвитку конфліктів. Застосований підхід дозволяє визначити «гарячі точки» конфліктів, що можуть слугувати вказівками на особливу небезпеку розвитку інцидентів в регіоні. «Холодні точки», навпаки, можуть вказувати на можливості швидкого згорання ескалації. Аналіз даних в динаміці за останні 30 років дозволяє докладніше розглянути перебіг військових конфліктів у регіонах світу.

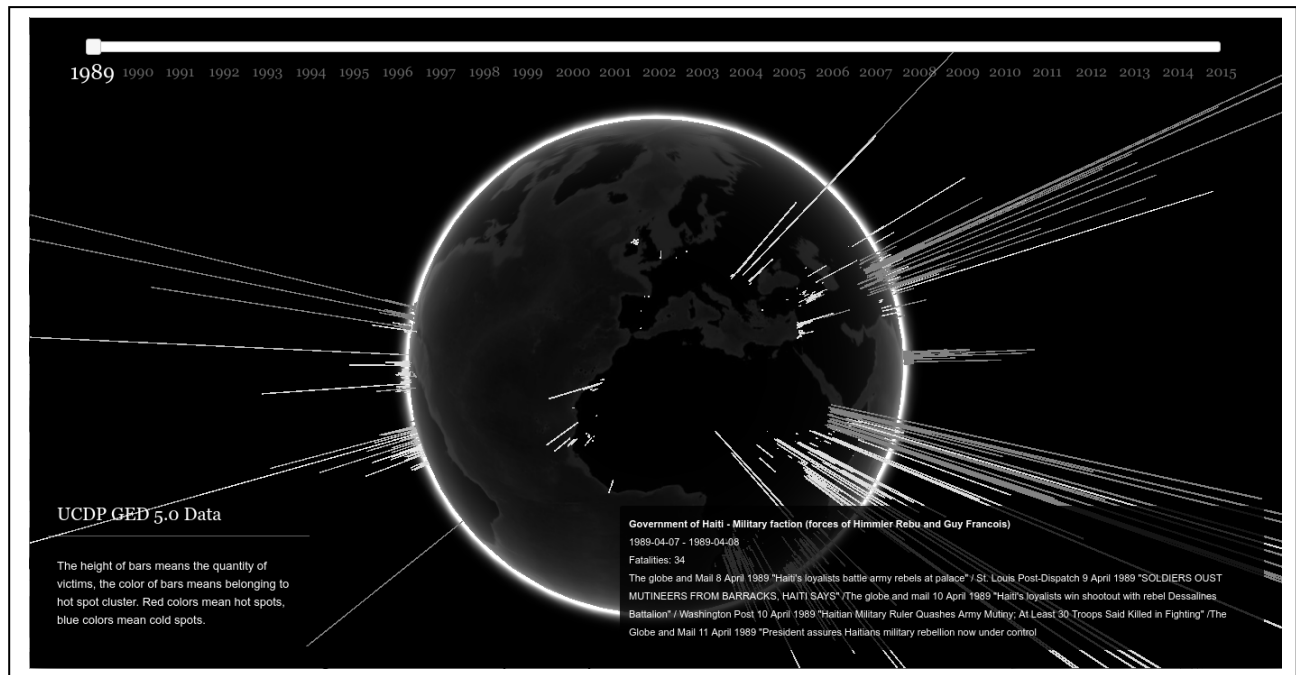


Рис. 3. Вигляд веб-додатку візуалізації результатів аналізу на 3D-глобусі.

Висновки.

Використовуючи метод аналізу гарячих точок було проведено кластеризацію даних набору UCDP GED 5.0 Упсальського університету. Проведений аналіз дозволяє отримати додаткову інформацію про характер та динаміку світових військових конфліктів протягом останніх 30 років.

За допомогою сучасних веб-технологій було розроблено засіб візуалізації результатів аналізу та інтерактивної взаємодії із ними. На тривимірний глобус були нанесені геокодовані дані про характер подій у світових збройних конфліктах та їх наслідки. Такий підхід полегшить доступ до інформації та спростить процес роботи із даними.

Вперше запропоновано підхід до кластерного аналізу даних про військові конфлікти з урахуванням їх геопросторового положення та топології відношень. Даний підхід дозволяє врахувати взаємний вплив близькості територій з напруженою військово-політичною ситуацією на можливості широкомасштабного розгортання зіткнень. В подальшому буде розглянуто просторо-часові зв'язки у базі даних.

Бібліографічний список.

1. Костенко Т.Ф. Типи, види воєнних конфліктів та їх класифікація / Т.Ф. Костенко // Науковий вісник Дипломатичної академії України. Випуск 2 / За заг. ред. Гуменюка Б.І. – 1999. – С. 45-48.
2. Перепелиця Г.М. Воєнно-політичний конфлікт: методологія дослідження та врегулювання: автореф. дис.. д-ра політ. наук: спец. 23.00.02 "Політичні інститути та

процеси” / Інст. держави і права ім. В.М. Корецького Національної академії наук України. – К., 1996. – 38 с.

3. Путренко В.В. Використання кластерного аналізу для просторового моделювання в геоінформаційних системах / В.В. Путренко, В.А. Тихоход // Географічна наука і практика: виклики епохи: Матеріали міжнародної наукової конференції, присвяченої 130-річчю географії у Львівському університеті (м. Львів, 16 – 18 травня 2013 р.). У 3-ох томах. – Львів: Видавничий центр ЛНУ імені Івана Франка. – 2013. – Том 3. – С. 122-124

4. Шевченко О. Збройні конфлікти та шляхи їх врегулювання / О. Шевченко // Соціальна психологія. – 2005. – № 2 (10). – С. 28-37.

5. Sundberg, Ralph, and Erik Melander, 2013, “Introducing the UCDP Georeferenced Event Dataset”, *Journal of Peace Research*, vol.50, no.4, 523-532.

6. Croicu, Mihai and Ralph Sundberg, 2016, “UCDP GED Codebook version 5.0”, Department of Peace and Conflict Research, Uppsala University.

7. Getis, A. and J.K. Ord. 1992. "The Analysis of Spatial Association by Use of Distance Statistics" in *Geographical Analysis* 24(3).

8. Yin Z.C. A Multi-Scale GIS Database Model Based On Petri Net / Z.C. Yin // ISPRS Workshop on Service and Application of Spatial Data Infrastructure, XXXVI (4/W6). – Hangzhou. – 2005. – P. 271 – 275.

9. Wang Y. H. Multi-Scale Conceptual Model for GIS Geographical Features / Y. H. Wang, J. Chen, J. Jian // *Journal of China University of Mining & Technology*. – 2003. – Vol. 32(4). – P. 376–382.

10. Zhang Chao The Application of Gisway System in Emergent Thematic Mapping / Chao Zhang, Lianfa S. I., Zhuoning Chen // *Disaster Engineering science*. – 2009. – 11(8): P. 24-28 (in Chinese). [Електронний ресурс] – Режим доступу: <http://www.seiofbluemountain.com/upload/product/201105/2011fzjz37.pdf>

Аннотация

Рассмотрены подходы к анализу военных конфликтов в регионах мира с помощью геопространственной кластеризации набора данных о мировых вооруженных конфликтах в ГИС. Проанализирован набор данных UCDP GED о вооруженных конфликтах в мире, начиная с 1989 года. Разработано программное обеспечение веб-приложения для визуализации результатов анализа на виртуальном глобусе в сети Интернет.

Ключевые слова: военный конфликт, геопространственный анализ, кластеризация, геовизуализация, 3d-глобус.

Annotation

Approaches to the analysis of military conflicts in the regions of the world are analyzed using geospatial clustering of a set of data on the world's military conflicts in GIS. The UCDP GED data set with information about military conflicts in the world since 1989 has been analyzed. A web application software has been developed to visualize the analysis results on a virtual globe over the Internet.

Keywords: military conflict, geospatial analysis, clusterization, geovisualization, 3d- globe.