

# СПОСОБИ ОПТИМІЗАЦІЇ ПРОЦЕСУ ОБРОБКИ ВЕЛИКИХ ДАНИХ ЗА ДОПОМОГОЮ РОЗПАРАЛЕЛЮВАННЯ

Л. Л. Шуміло<sup>1, a</sup>

<sup>1</sup> *Інститут космічних досліджень НАН України та ДКА України*

<sup>2</sup> *Фізико-технічний інститут НТУУ «КПІ»*

## Анотація

У даній роботі представлені способи розпаралелення процесів обробки даних в мовах програмування Python та C. Тестування методів проводився за допомогою 60-ти метрових та 20-ти метрових супутникових знімків Sentinel-1.

*Ключові слова:* розпаралелювання, Python, csharp, threading, subprocess, multiprocessing, Parallel

## Вступ

Під час роботи з великими даними часто виникають дуже ресурсоємкі завдання такі як побудова часового ряду супутникових знімків, необхідних для розв'язання багатьох прикладних задач таких як класифікація сільськогосподарських посівів, прогнозування врожайності та інші [1-3]. Тому виникає бажання, враховуючи кількість супутникових даних за останні роки та час їх обробки для складання часового ряду, оптимізувати використання обчислювальних ресурсів та пришвидшити цей процес.

## Постановка задачі та її розв'язок

В даній роботі запропоновані способи оптимізації процесу побудови часового ряду у програмах написаних на мовах python та C за допомогою розпаралелювання, що дозволяють здійснювати роботу з кількома каналами растрів одночасно. Для даного експерименту було використано 10 супутникових знімків Sentinel-1 за 2016 рік [4-11], та проаналізовано 3 способи розпаралелювання процесів на мові python: модуль threading, модуль subprocess та модуль multi-processing та 2 способи розпаралелювання на мові csharp: модуль threading та Parallel. Найкращим з них виявився спосіб з використанням модулю multi-processing, якщо в середньому процедура створення часового ряду на мові python працює на даній те-стовій вибірці 4760 секунд, тоді при розбитті її на 4 процеси за допомогою модуля multiprocessing ми отримуємо середній час в 4286 секунд. Основною проблемою оптимізації створення часового ряду за допомогою розпаралелювання є велика кількість звернень кожного процесу до жорсткого диску, оскільки одночасно до нього може отримати доступ тільки один процес, тим не менш було отримано хороший результат, враховуючи виграш в часі в середньому 473,5 секунд. На даному зображенні представлено порівняння часу роботи процедури створення часового ряду без розпаралелювання, з розпаралелюванням

на 2 процеси та з розпаралелюванням на 4 процеси, де перші три стовбці це максимальний час кожної процедури у відповідному порядку, наступні 3 стовбці середній час та останні три стовбці це мінімальний час роботи.

## Висновки

Використовуючи запропонований спосіб розпаралелювання процесів за допомогою модуля multiprocessing, вдалося оптимізувати та пришвидшити процес утворення часового ряду супутникових знімків на 2.5 відсотки при розбитті програми на 2 процеси та на 5 відсотків при розбитті на 4 процеси. Використання часового ряду знімків дозволило покращити точність карти класифікації, отриманої за допомогою ансамблю нейронних мереж, більше ніж на 10 відсотків у порівнянні з класифікацією по одному знімку [2].

## Список використаних джерел

1. Kussul N. Parcel based classification for agricultural mapping and monitoring using multi-temporal satellite image sequences / N. Kussul, G. Lemoine, J. Gallego, S. Skakun, M. Lavreniuk // The International Geoscience and Remote Sensing Symposium (IGARSS), — 2015. — P. 165-168.
2. Kussul N. Regional scale crop mapping using multi-temporal satellite imagery / N. Kussul, S. Skakun, A. Shelestov, M. Lavreniuk, B. Yailymov, O. Kussul // Inter. Arch. of the Photo., R. S. Spatial Inform. Sci-ences. — 2015. — P. 45-52.
3. Kolotii A. Comparison of biophysical and satellite predictors for wheat yield forecasting in Ukraine / A. Kolotii, N. Kussul, A. Shelestov, S. Skakun, B. Yailymov, R. Basarab, M. Lavreniuk, T. Oliinyk, V. Ostapenko // International Archives of the Photogrammetry, Remote Sensing Spatial Information Sciences. — 2015. — P. 39-44.
4. Kussul N. Grid technologies for satellite data processing and management within international disaster moni-

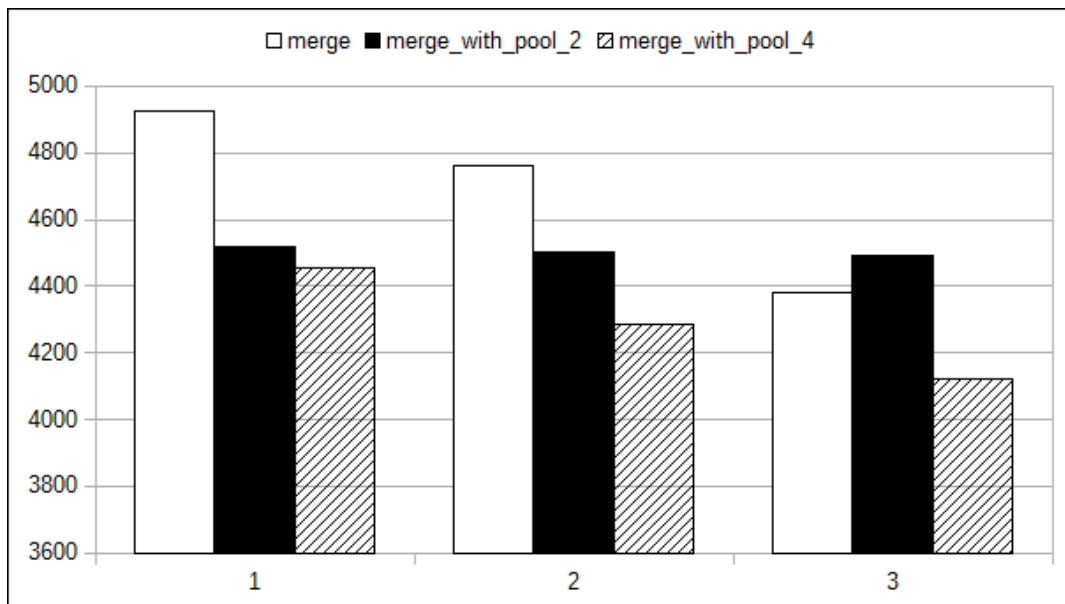


Рис. 1. 1-максимальний час, 2-середній час, 3-мінімальний час

- toring projects / N. Kussul, A. Shelestov, S. Kussul N., Skakun S., Shelestov A., Kussul O. // IEEE Skakun // Grid and Cloud Database Management. International Geoscience and Remote Sensing Symposium (IGARSS). — 2014. — P. 1497–1500.
5. Kravchenko A. Water resource quality monitoring using heterogeneous data and high-performance computations / A. Kravchenko, N. Kussul, E. Lupian, V. Savorsky, L. Hluchy, A. Shelestov // *Cybernetics and Systems Analysis*. — 2008. — Vol. 44, No. 4. — P. 616-624. DOI:10.1007/s10559-008-9032-x.
6. Bakan G.M. Fuzzy ellipsoidal filtering algorithm of static object state / G.M. Bakan, N.N. Kussul // *Problemy Upravleniya I Informatiki (Avtomatika)*.—1996. — No. 5. — P. 77-92.
7. Skakun S. Efficiency Assessment of Multitemporal C-Band Radarsat-2 Intensity and Landsat-8 Surface Reflectance Satellite Imagery for Crop Classification in Ukraine / S. Skakun, N. Kussul, A. Y. Shelestov, M. Lavreniuk and O. Kussul // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.— 2016. — Vol. 9, No 8. — P. 3712-3719. DOI: 10.1109/JSTARS.2015.2454297.
8. Kussul N. The use of satellite SAR imagery to crop classification in Ukraine within JECAM project /
9. Kussul N. Regional scale crop mapping using multi-temporal satellite imagery / N. Kussul, S. Skakun, A. Shelestov, M. Lavreniuk, B. Yailymov, O. Kussul // *International Archives of the Photogrammetry, Remote Sensing Spatial Information Sciences*. — 2015. — P. 45–52.
10. Lavreniuk M. Regional Retrospective High Resolution Land Cover For Ukraine: Methodology And Results / Lavreniuk M., Kussul N., Skakun S., Shelestov A., Yailymov B. // *International Geoscience and Remote Sensing Symposium 2015 (IGARSS 2015)*, № 15599383,— P. 3965-3968. DOI:10.1109/IGARSS.2015.7326693.
11. Mandl D. Use of the earth observing one (EO-1) satellite for the namibia sensorweb flood early warning pilot / D. Mandl, S. Frye, P. Cappelaere, M. Handy, F. Policelli, M. Katjizeu, .. , J. Silva // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.—2013. — Vol. 6, No 2. — P. 298-308.