

Polonizacja¹ i popularyzacja Bazy Danych Przestrzennych SpatiaLite na licencji Wolnego Oprogramowania

Polonization and popularization of Spatial Database SpatiaLite on Open Source License

Michał LUPA, Artur KRAWCZYK

Słowa kluczowe: bazy danych, bazy danych przestrzennych, GIS, SpatiaLite, Wolne Oprogramowanie
Key words: databases, spatial databases, GIS, SpatiaLite, Open Source

Otrzymano: 30 lipca 2012; **Zaakceptowano:** 2013
Received: 30 July 2012; **Accepted:** 2013

The article describes the issue of polonization and popularization of spatial databases, which were the subject of the project of Scientific Association of Computer Graphics And Geomatics, funded under the competition of "Grant Rektora 2012" in University of Science and Technology AGH. The polonization and popularization issues have been related to SpatiaLite database - the open source DBMS, which structure, architecture

and possibilities are presented in the paper. Furthermore, there are included advantages and disadvantages of this type of software, with a special focus on geoprocessing operations. In addition, except from sample operations of the spatial data, the article also describes the possibilities of remote sensing data processing and raster analysis.

Bazy danych są gałęzią informatyki, której prężny rozwój można zaobserwować już od początku lat sześćdziesiątych ubiegłego wieku. Wraz z rozwojem technologicznym relacyjnych baz danych wzrosła również liczba ich zastosowań w procesie przetwarzania różnego typu danych (pliki multimedialne, zdjęcia rentgenowskie, itd.) Relacyjne bazy danych znalazły również zastosowanie w Systemach Informacji Geograficznej (ang. *GIS – Geographic Information Systems*). W początkowym okresie służyły one wyłącznie jako magazyn danych, ponieważ przetwarzanie odbywało się po stronie programów GIS. Przełomowym momentem okazała się standaryzacja języka SQL (OGC, SQL/MM - Spatial), za sprawą której rozbudowano go o funkcje geoprzestrzenne oraz zdefiniowano typy danych geometrycznych. W następstwie tego odnotowano wzrost zainteresowania przetwarzaniem danych przestrzennych po stronie bazy, który za-

owocował utworzeniem tak zwanych „przestrzennych rozszerzeń” wśród najpopularniejszych systemów zarządzania bazą danych (SZBD, ang. *Database Management System, DBMS*). W ten sposób powstały bazy danych przestrzennych, które często określane są terminem „Geobazy”. Jedną z takich baz jest SpatiaLite, której polonizację i popularyzację wśród użytkowników Wolnego Oprogramowania zakłada projekt w ramach „Grantu Rektorskiego” AGH 2012².

Głównym celem Grantu Rektora, realizowanego w ramach Koła Naukowego Grafiki Komputerowej

¹ Polonizacja – terminem tym określono tłumaczenia na język polski dokumentacji technicznej oraz plików pomocy bazy.

² Grant Rektora AGH jest konkursem wyróżniającym najlepsze, najciekawsze oraz nowatorskie prace realizowane w ramach studenckich kół naukowych uczelni. Kwoty dofinansowania projektów są stosunkowo niewielkie, jednak często pozwalają na sfinansowanie wszystkich zakładanych wydatków. W 2009 roku odbyła się pierwsza edycja konkursu, który z roku na rok cieszy się coraz większą popularnością i wpisał się już na stałe w kalendarz działalności każdego koła naukowego.

i Geomatyki AGH było utworzenie portalu tematycznego www.spatialite.agh.edu.pl, zawierającego szereg kursów, zarówno dla początkujących, jak i zaawansowanych użytkowników języka SQL, którzy będą chcieli poszerzyć swoje umiejętności o metody geoprzetwarzania. Ponadto wykonane zostało tłumaczenie dokumentacji technicznej języka SQL oraz plików pomocy.

Spatialite jest rozwiązaniem posiadającym unikalną architekturę, która odrzuca model klient-serwer, bazując na rozwiązaniu zawierającym cały „silnik” SQL bezpośrednio w aplikacji. Wskutek tego baza traktowana jest jako pojedynczy proces lub jako komponent – na przykład *.dll*. Całość zaprojektowano w oparciu o silnik bazodanowy – SQLite, rozszerzony o możliwości składowania i przetwarzania danych geometrycznych. Przyjęte rozwiązania pozwoliły znacznie uprościć zarządzanie bazą, ponieważ administracja ogranicza się do operacji wykonywanych na plikach (otwórz, utwórz, zamknij, usuń). Kompletna baza danych, zawierająca często bardzo dużą ilość informacji również jest plikiem. Typy danych oraz funkcje geoprzestrzenne zaimplementowano zgodnie ze standardem OGC [OGC – The Open Geospatial Consortium, Spatialite Cookbook].

Główną zaletą systemu powyższego jest prostota oraz pełna przenośność pomiędzy platformami. Projekt jest udostępniany na licencji wolnego oprogramowania, nieustannie rozwijany i „żywy”. Wadą przyjętych rozwiązań jest ograniczona obsługa wielodostępności, która może być jednak wytłumaczona przyjętym założeniem projektu, leżącym u podstaw paradygmatu: „jeden użytkownik, jedna aplikacja, jedna stacja robocza”. [Lupa M., 2012]

Przetwarzanie danych, wykonywane po stronie bazy przewiduje import warstw zapisanych w formacie *.shp* oraz konwersje ich do postaci tabel SQL. Następnie odpowiednie zapytania wybierające, oparte na funkcjach geoprzestrzennych, umożliwiają przeprowadzenie dowolnej analizy dyskryminacyjnej badanego obszaru.

Przykładowo region Alaski, w którym wyznaczono rzeki przepływające przez obszar tundry. W tym celu zaproponowano następujące zapytanie:

```
Z1: SELECT *FROM tundra, rivers WHERE
      Contains(tundra.Geometry, rivers.Geometry);
```

Przykładowe wyniki wykonania powyższego zapytania zobrazowano na Ryc. 1.

Istnieje również szeroki zakres zastosowań bazy w dziedzinie teledetekcji, z uwagi na możliwość składowania, a także analizy obrazów rastrowych. Metody te zawarte są w specjalnej bibliotece – RasterLite, opartej na bazie danych przestrzennych Spatialite. Podstawowym formatem tej bazy jest geotiff, czyli format tiff, uzupełniony danymi o zastosowanym układzie współrzędnych oraz wpasowaniu rastra w tym układzie. Dane składowane są z wybranym rodzajem oraz stopniem kompresji (JPEG, PNG, GIF, WAVELET), a następnie poddawane obróbce z użyciem narzędzi do przetwarzania obrazów.

Cechą charakterystyczną oraz jedną z zalet oprogramowania RasterLite są niskie wymagania dotyczące pamięci potrzebnej do pracy z bazą. Podczas przetwarzania obrazów o wysokiej rozdzielczości (przykładowo:

	PK_UID	cat	F_CODEDESC	F_CODE	AREA_KM2	Geometry	PK_UID	cat	F_CODEDESC
1	15	19.000000	Tundra	BJ110	107.237000	BLOB sz=580 GEOMETRY	1459	1459	River/Stream
2	16	20.000000	Tundra	BJ110	114.426000	BLOB sz=820 GEOMETRY	1607	1607	River/Stream
3	67	31.000000	Tundra	BJ110	5020.838000	BLOB sz=18256 GEOMETRY	871	871	River/Stream
4	67	31.000000	Tundra	BJ110	5020.838000	BLOB sz=18256 GEOMETRY	873	873	River/Stream
5	67	31.000000	Tundra	BJ110	5020.838000	BLOB sz=18256 GEOMETRY	876	876	River/Stream
6	67	31.000000	Tundra	BJ110	5020.838000	BLOB sz=18256 GEOMETRY	879	879	River/Stream
7	68	32.000000	Tundra	BJ110	1203.558000	BLOB sz=5060 GEOMETRY	939	939	River/Stream
8	73	37.000000	Tundra	BJ110	418.956000	BLOB sz=2540 GEOMETRY	913	913	River/Stream
9	86	50.000000	Tundra	BJ110	200.158000	BLOB sz=708 GEOMETRY	616	616	River/Stream
10	86	50.000000	Tundra	BJ110	200.158000	BLOB sz=708 GEOMETRY	618	618	River/Stream
11	86	50.000000	Tundra	BJ110	200.158000	BLOB sz=708 GEOMETRY	633	633	River/Stream
12	86	50.000000	Tundra	BJ110	200.158000	BLOB sz=708 GEOMETRY	2321	2321	River/Stream

Ryc. 1. Rzeki Alaski przepływające przez obszar tundry.

Fig. 1. Rivers of Alaska flowing through the area of tundra.



Ryc. 2. Kompresja rastra do formatu JPEG o wysokiej jakości, RasterLite-How-To.

Fig. 2. Raster to JPEG lossy compression, high quality, RasterLite-How-To.

o rozmiarze 500 MB każdy), komputerem, który pozwoli na efektywną pracę jest standardowy PC z pamięcią 512 MB RAM. [RasterLite Man]

Na potrzeby optymalizacji procesu przetwarzania rastrow, oprogramowanie RasterLite wyposażono w mechanizm kafelkowania (ang. *tiling*). Obrazy o dużej rozdzielczości dzielone są na kilka mniejszych, zwanych kafelkami (ang. *the individual tiles*) Wskutek tego, do pamięci ładowane są tylko te części obrazu, które w danym momencie są przetwarzane. Pozwala to w prosty sposób zwiększyć efektywność oraz znacznie ograniczyć zużycie pamięci [RasterLite Man].

RasterLite pozwala również na obsługę gridów (siatkowy model danych). Gridy nie są obrazami cyfrowymi, lecz kwadratowymi siatkami współrzędnych, w których każda komórka przechowuje określoną informację zapisaną w postaci numerycznej wartości. Gridy stosowane są bardzo często do reprezentacji Numerycznych Modeli Terenu (NMT), gdzie każda komórka odpowiada średniej wysokości terenu.

W podsumowaniu można podkreślić bardzo uniwersalny charakter opisywanej bazy danych przestrzennych. Stanowi ona bardzo cenne narzędzie pracy młodego naukowca, czy entuzjasty, użytkownika oprogramowania

GIS, do pracy z każdym typem dostępnych danych. Baza może być przenoszona i używana wprost z dysku przenośnego. Trwają prace nad możliwością uruchomienia bazy danych w wersji mobilnej (na smartfonie). Warto podkreślić uniwersalność tejże bazy danych, ponieważ w ostatnich latach znalazła ona zastosowanie między innymi w przemyśle górniczym [Krawczyk, 2011]. Świadczy to o dużym potencjale tego rozwiązania, które w przyszłości zapewne stanie się bardzo popularne.

Literatura

- Krawczyk A., 2012, Możliwości wykorzystania personalnej bazy danych przestrzennych SpatiaLite w praktyce górniczej, *Roczniki Geomatyki*, T. 10, zeszyt 2, s. 45-49, Warszawa.
- Lupa M., 2012, Optymalizacja baz danych przestrzennych, , Praca magisterska AGH.
- OGC – The Open Geospatial Consortium, <http://www.opengeospatial.org/>
- RasterLite How-To page: <http://www.gaia-gis.it/gaia-sins/rasterlite-docs/rasterlite-how-to.pdf>
- RasterLite Man page: <http://www.gaia-gis.it/gaia-sins/rasterlite-docs/rasterlite-man.pdf>
- SpatiaLite CookBook: <http://www.gaia-gis.it/gaia-sins/spatialite-cookbook/index.html>.



Mgr inż. Michał Lupa – absolwent studiów magisterskich na Wydziale Geologii, Geofizyki i Ochrony Środowiska Akademii Górniczo-Hutniczej w Krakowie, w Katedrze Geoinformatyki i Informatyki Stosowanej. Obecnie uczestnik studiów doktoranckich o specjalności geoinformatyka w macierzystej jednostce. Członek Koła Naukowego Grafiki Komputerowej i Geomatyki. Głównym obszarem badań są bazy danych przestrzennych oraz systemy GIS. mlupa@geol.agh.edu.pl



Dr inż. Artur Krawczyk – W 2002 roku uzyskał stopień doktora nauk w specjalności geoinformatyka i inżynieria środowiska Akademii Górniczo Hutniczej w Krakowie gdzie pracuje do dnia dzisiejszego. Jest aktywnym członkiem Polskiego Towarzystwa Informatyki Przestrzennej z ramienia którego kieruje Komisją Geomatyki Górniczej. Na uczelni zajmuje się opieką koła naukowego KNGK (Studenckie Koło Naukowe Grafiki Komputerowej i Geomatyki). Głównym obszarem badań jest przestrzenne modelowanie zmian środowiska spowodowanych przez podziemną eksploatację górnictwem z zastosowaniem technik ALS oraz InSAR. artkraw@agh.edu.pl, home.agh.edu.pl/artkraw