

PROBLEMATYKA INTEGRACJI DANYCH PRZESTRZENNYCH Z RÓŻNYCH ŹRÓDEŁ W SYSTEMIE MOBILNEJ NAWIGACJI ŚRÓDLĄDOWEJ

PROBLEMS OF SPATIAL DATA INTEGRATION FORM VARIOUS SOURCES IN MOBILE NAVIGATIONAL SYSTEM FOR INLAND SHIPPING

Grzegorz Zaniewicz¹, Witold Kazimierski², Marta Włodarczyk-Sielicka¹

¹ Marine Technology Sp. z o.o.

² Akademia Morska w Szczecinie, Wydział Nawigacyjny, Instytut Geoinformatyki

Słowa kluczowe: integracja danych przestrzennych, nawigacja śródlądowa, model danych przestrzennych, IENC

Keywords: spatial data integration, inland navigation, spatial data model, IENC

Wstęp

Problem integralności danych przestrzennych występuje wszędzie tam, gdzie dane do budowy nowego systemu pozyskiwane są z wielu źródeł. Zwykle nie mają one spójnej struktury i są pod wieloma względami zróżnicowane. Konieczne jest zatem wypracowanie jednoznacznych reguł integracji, pozwalających na opracowanie produktu finalnego. Przykładem może być tu generowanie mapy zasadniczej z różnych baz danych.

Warto zauważyć, że potrzeba integracji danych w systemach występuje coraz częściej w dobie rozrostu źródłowych baz danych i coraz powszechniejszej dostępności do informacji przestrzennej. Alternatywą dla budowy zintegrowanej bazy danych jest kopiowanie informacji źródłowych, co w większości przypadków oznaczałoby niepotrzebną redundancję danych, na którą nie ma miejsca w nowoczesnych systemach geoinformatycznych.

Potrzeba integracji danych z różnych źródeł zaistniała także przy opracowaniu produktu mapowego w systemie mobilnej nawigacji śródlądowej MOBINA^V. Ma on w swych założeniach korzystać z dostępnych powszechnie źródeł danych, na podstawie których ma być generowana mapa numeryczna. W ramach prac nad systemem opracowano zestaw reguł importu danych, a także zintegrowany model danych, spełniający wymagania i założenia systemu. Założono przy tym, że mapa ma być generowana w sposób możliwie automatyczny, a finalnie dane będą przechowywane w nowym opracowanym formacie wymiany danych MODEF (*MObinav Data Exchange Format*), bazującym na XML.

W kolejnych rozdziałach artykułu przedstawiono informacje o budowanym systemie, przedstawiono koncepcję integracji danych, scharakteryzowano źródła danych oraz przedstawiono założenia formatu MODEF. Szczególną uwagę zwrócono na problemy integracji danych, a zwłaszcza na problemy integralności geometrycznej i atrybutowej.

Prace badawcze nad systemem są realizowane w ramach projektu badawczego finansowanego przez Narodowe Centrum Badań i Rozwoju w ramach programu LIDER IV.

Mobilna nawigacja śródlądowa

Mobilna nawigacja śródlądowa, to projekt badawczy łączący w sobie zagadnienia z zakresu kartografii, geoinformatyki oraz nawigacji. Wymiernym efektem będzie prototyp systemu MOBINA, który będzie innowacyjnym systemem mobilnej nawigacji śródlądowej, odpowiadającym wymaganiom użytkowników rekreacyjnych. W myśl przyjętej hipotezy badawczej, wybór tej technologii mobilnej i metodologii kartografii mobilnej pozwoli na opracowanie produktu, który będzie miał szansę na wdrożenie i seryjną produkcję. Produkt będący efektem projektu ma uzupełniać lukę na rynku pomiędzy profesjonalnymi systemami nawigacyjnymi typu ECDIS, wymagającymi standardowych map IENC i dodatkowego sprzętu, jak AIS itd., a dostępnymi obecnie systemami nawigacji samochodowych i ich rozszerzeniami dla żeglugi. Warto zwrócić uwagę, że obszarem docelowym mają być wody śródlądowe, a nie jak w większości systemów wody morskie.

MOBINA to przykład systemu geoinformatycznego, dedykowanego dla potrzeb wybranej grupy odbiorców. Założenia systemu uwzględniają nie tylko opracowanie dedykowanego modelu mobilnej prezentacji kartograficznej, ale także odpowiedniego zestawu analiz przestrzennych. Podstawą do obu mają być dane przestrzenne pochodzące z powszechnie dostępnych źródeł. Istnieje zatem konieczność integracji tych danych. W tym celu opracowano koncepcyjny model danych w systemie oraz odpowiednie reguły importu danych.

Specyfikacja źródeł danych

Podstawą systemu mobilnej nawigacji śródlądowej mają być dane przestrzenne pochodzące z powszechnie dostępnych źródeł. W ramach projektu zdecydowano o wykorzystaniu elektronicznych map nawigacyjnych dla żeglugi śródlądowej (*Inland Electronic Navigational Chart* IENC), bazy danych obiektów topograficznych BDOT10k, bazy mapy Vmap poziomu drugiego oraz powszechnie dostępnych danych zgromadzonych w ramach projektu Open Street Map. Mapy IENC zawierają szczegółowe dane, których skala kompilacji wynosi 1:2000. Są bardzo dobrym źródłem informacji o wodach śródlądowych, ale ich zasadniczą wadą jest fakt, że na terenie Polski pokrywają one jedynie część dróg wodnych dolnego odcinka rzeki Odry, od miejscowości Ognica do Szczecina, których łączna długość wynosi 97,6 km. Z tego powodu na potrzeby projektu zdecydowano się wykorzystać także inne źródła danych wejściowych. Kolejnym powodem wyboru dodatkowych źródeł informacji geograficznej jest to, że IENC zawierają dane ściśle związane z nawigacją, a głównymi użytkownikami systemu mają być turyści. Pozostałe bazy danych wejściowych zawierają uzupełniające geodane, które będą przydatne dla użytkowników końcowych systemu.

IENC

IENC można zdefiniować jako bazę danych, standaryzowaną co do zawartości, struktury i formatu, która jest przeznaczona do użytkowania w śródlądowych systemach obrazowania elektronicznych map i informacji nawigacyjnych (Inland ECDIS). Mapy elektroniczne są wydawane przez właściwy urząd lub kompetentną w tym zakresie agencję rządową. Śródlądowa mapa elektroniczna powinna zawierać informacje niezbędne do prowadzenia bezpiecznej nawigacji na akwenach śródlądowych. Dodatkowo może zawierać inne informacje, analogiczne do tych znajdujących się na mapie papierowej, które mogą być przyjęte za ważne dla potrzeb prowadzenia bezpiecznej nawigacji oraz planowania podróży. Minimalny zakres obiektów na mapie IENC powinien zawierać: linię toru żeglownego, konstrukcje przybrzeżne (falochrony, tamy podłużne, itp.), dowolny obiekt, który jest uważany za zagrożenie dla żeglugi, kontury śluz i tam, granice kanałów nawigacyjnych, jeśli są zdefiniowane, odosobnione niebezpieczeństwa, w tym przeszkody podwodne, nawodne (mosty, kable, itp.), pomoce nawigacyjne (pławy, stawy, światła, znaki nawigacyjne) oraz kilometraż lub hektometraż. Poziomym układem odniesienia IENC jest WGS84 (IENC HG, 2011).

Mapy IENC zawsze muszą spełniać wymagania standardów opracowanych przez Międzynarodową Organizację Hydrograficzną IHO (*International Hydrographic Organization*) oraz grupę harmonizującą IENC. Standard map elektronicznych dla żeglugi śródlądowej oparty jest na morskim standardzie S-57, ze wszystkimi załącznikami i aneksami, co wiąże się ze stosowaniem tego samego modelu teoretycznego i struktury danych. Standard ten wymagał jednak modyfikacji ze względu na potrzebę prezentacji na mapach IENC obiektów i danych wykorzystywanych w żegludze śródlądowej.

Mapy IENC są opracowane zgodnie ze standardem Inland ENC, wypracowanym przez europejską grupę harmonizującą (*Inland ENC Harmonization Group* – IEHG) w którego skład wchodzi:

- *Product Specification for Inland ENCs Ed. 2.3*, styczeń 2011
- *Encoding Guide for Inland ENCs Ed. 2.3.6*, lipiec 2014
- *IENC Feature Catalogue Ed. 2.3*, kwiecień 2013

Głównym celem prac IEHG jest uzgodnienie specyfikacji Inland ENC, które są odpowiednie do wszystkich znanych elektronicznych map śródlądowych oraz wymogów dotyczących danych dla prowadzenia bezpiecznej i efektywnej nawigacji na śródlądowych drogach wodnych.

BDOT

Baza danych obiektów topograficznych (BDOT10k) jest to baza danych przestrzennych o dokładności odpowiadającej mapom topograficznym w skali 1:10 000, zawierająca informacje o lokalizacji przestrzennej oraz charakterystykę określonych obiektów topograficznych. Zasób BDOT10k jest tworzony w oparciu o wektoryzację ortofotomapy cyfrowej, pomiary bezpośrednie, wykorzystanie danych BDOT500 oraz innych rejestrów prowadzonych przez instytucje publiczne.

Podstawy prawne, organizacyjne oraz standardy techniczne gromadzenia, przechowywania i udostępniania bazy danych obiektów topograficznych określa Rozporządzenie Ministra Spraw Wewnętrznych i Administracji z dnia 17 listopada 2011 roku w sprawie bazy danych obiektów topograficznych oraz bazy danych obiektów ogólnogeograficznych, a także standardowych opracowań kartograficznych (Rozporządzenie, 2011). Ostateczna struk-

tura BDOT, została opublikowana w Rozporządzeniu Ministra Administracji i Cyfryzacji z dnia 12 lutego 2013 r. w sprawie bazy danych geodezyjnej ewidencji sieci uzbrojenia terenu, bazy danych obiektów topograficznych oraz mapy zasadniczej (Rozporządzenie, 2013). Obiekty pogrupowano na trzech poziomach klasyfikacyjnych zawierających odpowiednio: 1) kategorie klas obiektów, 2) klasy obiektów, 3) określone obiekty.

Kategorie klas obiektów w bazie BDOT dzielą się na: sieć wodną, sieć komunikacyjną, sieć uzbrojenia terenu, pokrycie terenu, budowle i urządzenia, kompleksy użytkowania terenu, tereny chronione, jednostki podziału terytorialnego oraz inne obiekty. Układem współrzędnych stosowanym dla danych w bazie BDOT jest Państwowy Układ Współrzędnych Geodezyjnych 1992 (Rozporządzenie, 2011).

Vmap

Mapa Vmap poziomu drugiego, jest to baza danych geoprzestrzennych o szczególności właściwej dla mapy w skali 1:50 000, pokrywająca w całości obszar Polski. Vmap została opracowana przez Służbę Topograficzną Wojska Polskiego, a w późniejszym okresie przy współudziale GUGiK (Główny Urząd Geodezji i Kartografii). Obiekty zgrupowane są w następujących kategoriach tematycznych: granice, rzeźba terenu, hydrografia, przemysł, fizjografia, obiekty socjalno-kulturalne, transport, uzbrojenie terenu oraz roślinność.

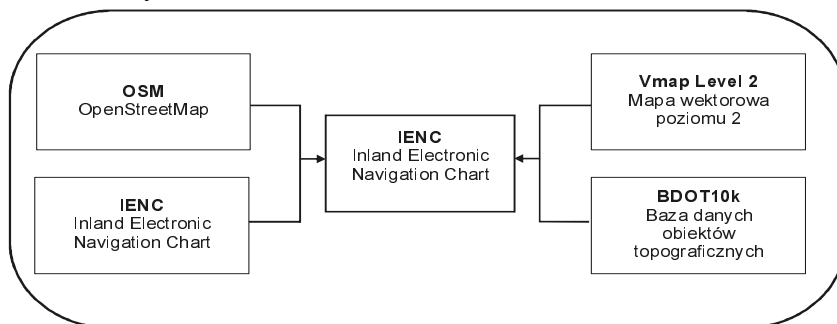
Mapy Vmap dystrybuowane są w formacie VPF (*Vector Product Format*). Model ten jest dość złożony, gdyż integruje geometrię, topologię i atrybuty w jednorodnej, relacyjnej strukturze danych. Zasięg geograficzny bazy danych podzielony jest na mniejsze jednostki zwane bibliotekami, w których strukturze wyróżnia się dziesięć warstw tematycznych. Do jednej z nich należy każda, określona w specyfikacji produktu, klasa obiektu geograficznego. Fizyczną implementacją klasy jest tablica relacyjna utrzymująca atrybuty obiektów oraz łączniki do danych przestrzennych. Obiekty geograficzne i ich atrybuty są kodowane zgodnie ze schematem FACC (*Feature and Attribute Coding Catalogue*). Poziomym układem odniesienia map Vmap jest WGS84 (Gotlib i in., 2006).

OSM

OpenStreetMap (OSM) jest projektem społeczności internetowej, mającym na celu stworzenie darmowej oraz swobodnie dostępnej mapy całej kuli ziemskiej, która jest modyfikowana przez użytkowników. Projekt OpenStreetMap został zainicjowany w 2004 roku przez Steve'a Coasta z University College of London, a jego misją jest dostarczanie zarówno gotowych map, jak i surowych geodanych *wszystkim, którzy tego potrzebują*. Użytkownicy dostarczają dane w różnych skalach oraz o różnej dokładności. Dane przechowywane są w relacyjnej bazie PostgreSQL, bez rozszerzeń przestrzennych. Podstawową jednostką jest węzeł, który może występować samodzielnie jako obiekt punktowy, bądź także jako część innego obiektu. Z węzłów budowane są drogi, które z kolei mogą być liniowe lub powierzchniowe. Wymienione typy mogą wchodzić w skład relacji, służących do logicznego połączenia obiektów, tworzących większą całość. Dane atrybutowe mają postać tzw. tagów, przechowywanych w układzie klucz: wartość, bez ograniczeń typu ani zawartości. Obowiązującym językiem w definiowaniu atrybutów jest język angielski. Układem odniesienia stosowanym w OSM jest WGS84 (Drop i in., 2013). Paradoksalnie, społecznościowy charakter bazy OSM, który jest jej największą zaletą, jest jednocześnie największą wadą. Nie wszyscy użytkownicy przestrzegają ustanowionych reguł i stosunkowo często atrybuty przyjmują dowolne wartości. Utrudnia to istotnie dalszą automatyzację przetwarzania tej bazy.

Wyzwania integracji danych

Dedykowanym produktem mapowym dla systemów wspierania nawigacji na wodach śródlądowych jest standaryzowana elektroniczna mapa nawigacyjna IENC. Jej dostępność w Polsce jest jednak niewielka, gdyż obecnie obejmuje ona jedynie ok. 100 km dolnego odcinka rzeki Odry. W tej sytuacji, w związku z brakiem dostępności szczegółowych danych związanych z informacją mapową na wodach śródlądowych, opierając się na wcześniej opisanych źródłach, zaproponowano jeden integralny produkt mapowy. Pozwoli to na dostarczenie użytkownikom śródlądowym potrzebnych informacji bez konieczności opracowywania nowych, źródłowych baz danych. Schemat ideowy integracji źródeł danych przedstawiono na rysunku 1.



Rysunek 1. Schemat integracji źródeł danych

W sensie ideowym można by utworzyć wieloreprezentacyjną bazę danych, która zawierałaby wszystkie informacje o wszystkich obiektach, jednak sens praktyczny takiego działania byłby w tym przypadku niewielki. Zdecydowano się na wybór poszczególnych obiektów, istotnych dla użytkownika oraz ich odpowiednie przetworzenie. Zasadnicze problemy na jakie natrafiono opracowując reguły integracji danych to:

- problem niespójności modeli pojęciowych,
- problem niespójności geometrycznej,
- problem niespójności atrybutowej.

Specyfika każdego ze źródeł danych powoduje, że zastosowane modele pojęciowe są w wielu miejscach niespójne i te same obiekty topograficzne są modelowane w zupełnie inny sposób. Przykładem mogą być tu obszary wodne, rozbite szczegółowo w przypadku IENC na cztery klasy poligonowe (kanały, rzeki, obszar morski, jeziora) i dwie liniowe (rzeki, kanały), w przypadku BDOT10k trzy klasy poligonowe (woda morska, woda płynąca, woda stojąca) oraz trzy klasy liniowe (rzeka i strumień, kanał i rów melioracyjny), zaś w przypadku Vmap jest to jedna klasa poligonowa (wody powierzchniowe), podobnie jak w OSM (*waterways*). Zdecydowano się na opracowanie nowego modelu pojęciowego i mapowanie wybranych elementów z modeli źródłowych.

Problem niespójności geometrycznej wynika przede wszystkim z różnych metod oraz dokładności pozyskiwania danych, różnych skal kompilacji bazy oraz zastosowanych reguł generalizacji. Problem ten został szczegółowo omówiony w dalszej części artykułu, podobnie jak problem niespójności atrybutowej. Wiąże się on w pewnym sensie z różnicami w modelach pojęciowych. Różne cechy obiektów zostały uwypuklone w poszczególnych bazach, a często te same cechy zostały inaczej nazwane.

Proponowany produkt, stanowiący zintegrowaną bazę danych, będzie łączył ze sobą, pod względem geometrycznym i atrybutowym, dane zawarte w wykorzystywanych źródłach. Integracja musi odbywać się, zachowując opracowane reguły odpowiedniego wyświetlania danych oraz mapowania atrybutów obiektów przestrzennych.

Pod względem technologiczny założono, że dane będą przechowywane w dedykowanym formacie opartym na XML. Zakłada się, że w pierwszym etapie tworzenia bazy danych, system będzie automatycznie wybierał ze źródeł danych przestrzennych tylko te obiekty, które zostały zdefiniowane na poziomie budowy logicznego modelu danych. Takie podejście, pozwala na wstępną selekcję danych, przeznaczonych dla efektywnej pracy z mobilnym systemem nawigacji śródlądowej.

Pod względem przetworzenia informacji atrybutowej, należy przygotować odpowiednie reguły, integrując bazy danych wykorzystywanych źródeł z bazą danych opisywanego systemu. W rezultacie, wykorzystuje się tylko te atrybuty, które znajdują zastosowanie w opisie poszczególnych obiektów.

Problem niespójności geometrycznej danych

Problem integracji różnych danych z wielu źródeł, widoczny jest przy analizie spójności geometrycznych porównywanych warstw. W związku z tym, że poszczególne źródła danych zostały skompilowane w różnych skalach oraz z różną dokładnością pozyskiwania danych (m.in. wektoryzacja czy pomiar bezpośredni), nie jest możliwe bezpośrednie, jednoczesne połączenie danych ze wszystkich źródeł.

Na rysunku 2 przedstawiono fragmenty map przystani żeglarskich jeziora Małe Dąbie w Szczecinie pozyskane z 5 źródeł danych. Na rysunkach umieszczono tylko wybrane warstwy danych, które uznano, jako przydatne w fazie analizy dostępnych informacji. Na pierwszy rzut oka, porównując z ortofotomapą, można zauważyć różnicę pod względem ilości i jakości obiektów na poszczególnych mapach. Porównując generalizację na mapach, widoczna jest ona najbardziej w warstwie budynków na mapie IENC. Mapy OSM i BDOT10k w miarę dokładnie i szczegółowo odzwierciedlają obszar i obiekty z ortofotomapy.

Analizując obszar wodny powyższego fragmentu map, jedynie mapa IENC i BDOT10k zawierają obiekty reprezentujące pomosty do cumowania jednostek pływających. W mapie nawigacyjnej IENC przedstawione są one jako obiekty poligonowe, natomiast na mapie BDOT10k zostały zgeneralizowane do geometrii liniowej. Dodatkowo na mapie IENC, widoczne są dane punktowe, reprezentujące śródlądowe oznakowanie nawigacyjne.

Najgorzej, pod względem ilości i jakości danych, przedstawiają się dane pochodzące z Vmap L2. Odnosić można braki obiektów widocznych na innych mapach oraz wysoką generalizację linii brzegowej (rys. 3). Widoczna niespójność geometryczna linii brzegowej z Vmap L2 (pogrubiona linia) z danymi z IENC, wpływają na brak możliwości integracji danych. Dane z różnych źródeł nie pokrywają się i mogą tworzyć błędy topologiczne na linii woda-ład.

W wyniku analizy dostępnych źródeł dla budowanej aplikacji MOBINA, zaproponowano kolejność wyświetlania odpowiedniego źródła ze względu na jego dostępność na danym akwenu: 1) IENC, 2) BDOT10k, 3) OSM, 4) Vmap L2. Na przykład, jeżeli na akwenu śródlądowym użytkownik aplikacji będzie objęty zakresem map IENC to takie dane będą wyświetlane. W przypadku braku danych IENC, podkładem mapowym będzie BDOT10k (drugi w kolejności), itd.

Problem niespójności atrybutowej danych

Różne sposoby pozyskiwania danych źródłowych wiążą się również z problemem niespójności atrybutowej. Jest to w zasadzie problem oczywisty przy zastosowaniu różnych rodzajów zbiorów geodanych, jednak warto o nim wspomnieć szczególnie w przypadku danych z OSM. O ile w pozostałych, standaryzowanych zbiorach niespójność atrybutową można stosunkowo łatwo wyeliminować, o tyle dane pozyskane z OSM charakteryzują się swoją specyfiką i w wielu przypadkach wymagają ręcznej korekty.

Jak wcześniej wspomniano w OSM dane atrybutowe mają postać tagów i mogą być przypisywane do każdego elementu niezależnie. Nie występuje typowa tabela atrybutów z kolumnami i wierszami, lecz występuje układ klucz: wartość. Atrybuty nie posiadają ograniczeń co do typu czy zawartości. Istnieją jedynie pewne konwencje, które są ogólnie dostępne dla użytkowników OSM. Niestety nie każdy użytkownik OSM edytuje dane zgodnie z zaleceniami. Dodatkowo istnieją różne możliwości otagowania dokładnie tych samych obiektów. Przykładowo nieużytkowana linia kolejowa może występować jako *railway: disused* lub jako *railway: abandoned*. Obowiązującym językiem w definiowaniu atrybutów jest język angielski. Jedynie atrybut dotyczący nazwy obiektu jest wypełniany w języku polskim. Warto przy tym zauważyć, że część z danych dostępnych w OSM nie posiada wypełnionych wszystkich tagów. Zatem czasem należy uzupełniać braki w danych.

Dla wszystkich geodanych systemu mobilnej nawigacji śródlądowej pozyskiwanych z różnych źródeł wykorzystano mapowanie atrybutów na bieżąco w trakcie importu do zdefiniowanego modelu. Wszystkie algorytmy pozyskiwania danych zostały szczegółowo opisane dla każdego obiektu osobno.

Koncepcyjny model danych

W celu zbudowania bazy danych systemu MOBINA, zdefiniowany został koncepcyjny model danych, opisujący poszczególne klasy obiektów, atrybuty oraz ich dziedziczenie z warstwy abstrakcyjnej. Założono, że opracowana baza danych będzie miała charakter bazy wielorozdzielczej, w której różne reprezentacje kartograficzne obiektów będą kodowane za pomocą odpowiednich atrybutów, a wzajemna identyfikacja obiektów na różnych poziomach skalowych będzie przebiegać według tzw. wariantu atrybutowego (Gotlib, 2009).

Koncepcyjny model danych zawiera 34 klasy obiektów przestrzennych i jedną klasę abstrakcyjną. Jest ona swego rodzaju „super klasą”, z której dziedziczą pozostałe. W klasie abstrakcyjnej zawarto podstawowe atrybuty, między innymi opis geometrii, źródło danych (IENC, BDOT, itd.). Innymi atrybutami w tej klasie są także informacje definiujące wyświetlanie danego obiektu, pod względem skali (atrybuty: skala minimalna, skala maksymalna) oraz odpowiedniego symbolu (m.in. linki do bibliotek symboli w odpowiedniej skali).

Cały model danych, oparty został na zdefiniowanym opisie formatu danych, w którym szczegółowo określono: opisy poszczególnych obiektów, atrybuty, typy danych oraz dostępne wartości atrybutów. Na rysunku 4 przedstawiono fragment opracowanego modelu danych. Przykład przedstawia klasę obiektu reprezentującą obszary wodne. Tak jak dla każdej klasy, tak i w tym modelu, zastosowano dziedziczenie atrybutów z warstwy abstrakcyjnej. Oprócz tego, klasy posiadają swoje unikalne atrybuty opisane sześciornikowym akronimem (np. CATWAT – kategoria wody). W przykładzie przedstawiono, także stałą listę predefiniowanych wartości dla

dwóch atrybutów: CATWAT i OPSTAT. Dla każdej z klas opracowano reguły importu danych z baz źródłowych, a całość formatu zdefiniowano za pomocą XML.

Podsumowanie

Problematyka integracji danych z wielu źródeł występuje w zdecydowanej większości systemów, w których korzysta się z więcej niż jednej bazy danych. Możliwe są dwa rozwiązania: 1) kopiowanie danych źródłowych, powodujące sporą redundancję danych, 2) integracja danych źródłowych do wspólnego formatu, poprzez zdefiniowanie odpowiednich reguł importu. To drugie rozwiązanie zastosowano w systemie mobilnej nawigacji śródlądowej i jednocześnie zdecydowano się na utworzenie własnego modelu danych, zdefiniowanego w postaci schematu XML.

W artykule opisano problemy, na jakie natrafiono podczas integrowania baz zewnętrznych. Są to przede wszystkim problemy z niespójnością geometryczną i atrybutową danych. Wynikają one głównie z różnych modeli pojęciowych przyjętych w poszczególnych bazach źródłowych. Każda baza ma swoją specyfikę, oznaczającą także sposób opisu świata rzeczywistego za pomocą modeli, skalę kompilacji, generalizację, czy wreszcie dokładność geometryczną danych. W prowadzonych pracach zwykle kilka problemów występowało na raz. Ze względu na różną skalę kompilacji baz wynikały jednocześnie problemy z dokładnością w sensie geometrycznym, jak i z dokładnością odtworzenia szczegółów terenowych.

Podsumowując przeprowadzone badania można stwierdzić, że ze względu na istotne różnice pomiędzy bazami źródłowymi, zdecydowano się przeprowadzić integrację, działającą na zasadzie selekcji. Poszczególnym bazom nadano poziom istotności będący w istocie oceną bazy dla prowadzonego projektu. Podczas tworzenia bazy wynikowej zastosowano algorytm sprawdzający, które źródła istnieją dla żądanego obszaru i z nich wybiera dane do prezentacji, zaczynając od najbardziej wiarygodnego. W ten sposób ówczesnie utworzona na podstawie modelu koncepcyjnego geobaza danych została wypełniona danymi pochodzącymi z różnych źródeł. Takie podejście pozwala na pozyskanie, w sposób częściowo zautomatyzowany, możliwie dokładnych danych dla potrzeb systemu nawigacyjnego, bez konieczności wyboru zbioru danych przez użytkownika.

Literatura

- Drop P., Gajewski P., Mackiewicz M., 2013: Zastosowanie danych OpenStreetMap oraz wolnego oprogramowania do badań dostępności komunikacyjnej w skali lokalnej. *Acta Universitatis Lodzianensis Folia Geographica Socio-Oeconomica* nr 14.
- Gotlib D., Iwaniak A., Olszewski R., 2006: Budowa krajowej infrastruktury danych przestrzennych w Polsce – harmonizacja baz danych referencyjnych. Katedra Geodezji i Fotogrametrii, Akademia Rolnicza we Wrocławiu.
- Gotlib D., 2009: Wybrane aspekty modelowania wielorozdzielczych i wieloreprezentacyjnych baz danych topograficznych, *Geomatics and Environmental Engineering* vol. 3, no. 1: 25-36, AGH.
- Inland ENC Harmonization Group, 2011, Product Specification for Inland ENCs Ed. 2.3
- Inland ENC Harmonization Group, 2014, Encoding Guide for Inland ENCs Ed. 2.3.6
- Inland ENC Harmonization Group, 2013, IENC Feature Catalogue Ed. 2.3
- IHO, 2000: Transfer Standard for Digital Hydrographic Data, Ed.3.1.Special Publication No. 57
- Rozporządzenie Ministra Spraw Wewnętrznych i Administracji z dnia 17 listopada 2011 r. w sprawie bazy danych obiektów topograficznych oraz bazy danych obiektów ogólnogeograficznych, a także standardowych opracowań kartograficznych. Dz.U. 2011 nr 279 poz. 1642.

Rozporządzeniu Ministra Administracji i Cyfryzacji z dnia 12 lutego 2013 r. w sprawie bazy danych geodezyjnej ewidencji sieci uzbrojenia terenu, bazy danych obiektów topograficznych oraz mapy zasadniczej. Dz.U. 2013 poz. 383.

Streszczenie

Artykuł przedstawia problematykę integracji danych przestrzennych w systemie mobilnej nawigacji śródlądowej. W tym przypadku źródłami wejściowymi danych są: śródlądowa elektroniczna mapa nawigacyjna (IENC), topograficzna baza danych BDOT10k, baza mapy Vmap poziomu drugiego oraz powszechnie dostępne dane zgromadzone w ramach projektu OpenStreetMap. Efektem finalnym ma być wielorozdzielcza i wieloreprezentacyjna baza danych przestrzennych oraz odpowiedni schemat aplikacyjny XML. Problemy integracji danych wynikają przede wszystkim z ich niespójności geometrycznej i atrybutowej. Na pierwszą największy wpływ ma różna skala kompilacji poszczególnych baz źródłowych. Dodatkowo różne sposoby pozyskiwania danych, zastosowane w różnych zbiorach również wpłynęły na to, że w wielu miejscach dane nie pokrywają się ze sobą. Niespójność atrybutowa jest w zasadzie oczywistym problemem przy różnego rodzaju zbiorach, jednak warto o niej wspomnieć szczególnie w przypadku OSM. O ile w pozostałych, standaryzowanych zbiorach niespójność atrybutową można stosunkowo łatwo wyeliminować, o tyle dane OSM charakteryzują się swoją specyfiką i w wielu przypadkach wymagają ręcznej korekty. W artykule dokonano identyfikacji problemów integracji danych przestrzennych dla potrzeb systemu mobilnej nawigacji śródlądowej, a także przedstawiono wybrane zagadnienia w ujęciu praktycznym na rzeczywistych przykładach.

Abstract

Mobile navigation for inland shipping is an example of a GIS system dedicated to selected group of users. The system concept includes elaboration of dedicated model of mobile cartographic presentation as well as a suitable set of spatial analysis. The basis for both is spatial data derived from publicly available resources. Thus, there is a need of integrating these data. For this particular reason a conceptual data model for the system was created together with suitable importing rules. The paper presents problems of spatial data integration. The input sources are Inland Electronic Navigational Chart (IENC), topographic data base BDOT10k, database of Vmap level 2 and publicly available data covered by open source project Open Street Map. The final effect to be achieved is multiresolution and multirepresentation database and suitable XML application scheme. Integration problems are caused mostly by inconsistency of geometry and attributes. The first one is mostly influenced by various compilation scales. Additionally, different methods of gathering data used caused that in many places data are not overlaying correctly. Inconsistency of attributes is in fact obvious, when using various data sources, however it is worth mentioning especially in case of OSM. In other sources attribute problems can be easily solved, as they are standardized, but OSM data have their own characteristics and they require manual correction. The paper gives identification of problems with spatial data integration, but also includes practical approach to selected problems, based on real examples.

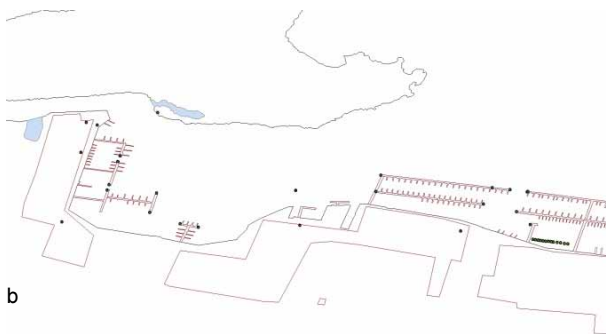
mgr inż. Grzegorz Zaniewicz
g.zaniewicz@marinetechnology.pl

dr inż. Witold Kazimierski
w.kazimierski@am.szczecin.pl

mgr inż. Marta Włodarczyk-Sielicka
m.wlodarczyk@marinetechnology.pl



a



b

Rysunek 2. Fragment obszaru przystani żeglarskich:

a – ortofotomapa (źródło: www.groportal.gov.pl),

b – IENC (źródło: RIS-Odra),

c – BDOT10k (źródło: WODGIK),

d – OSM (źródło: www.openstreetmap.pl),

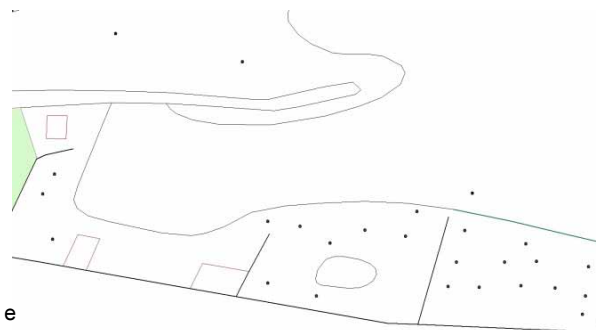
e – Vmap L2 (źródło: WODGIK)



c

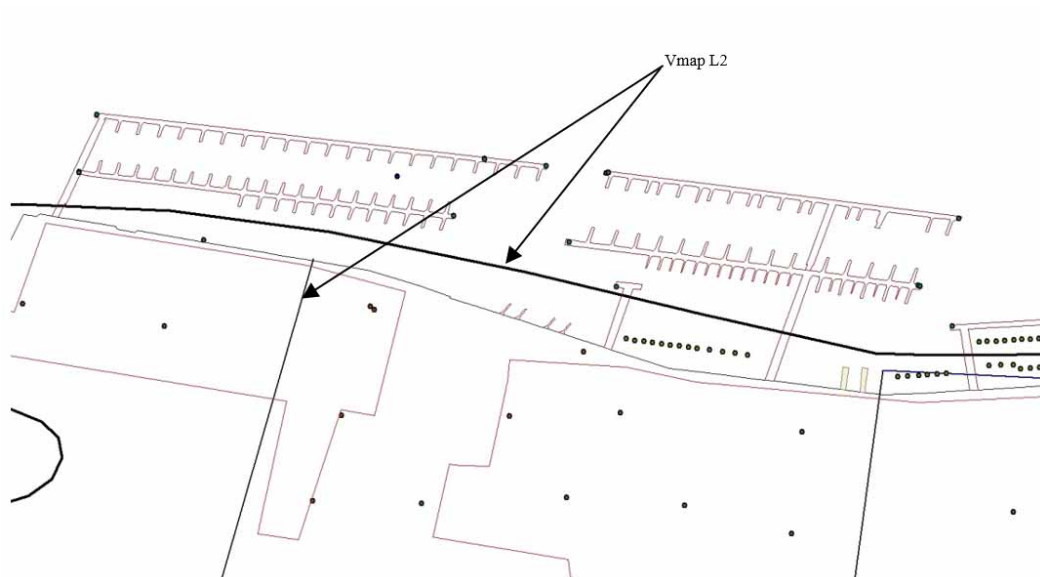


d

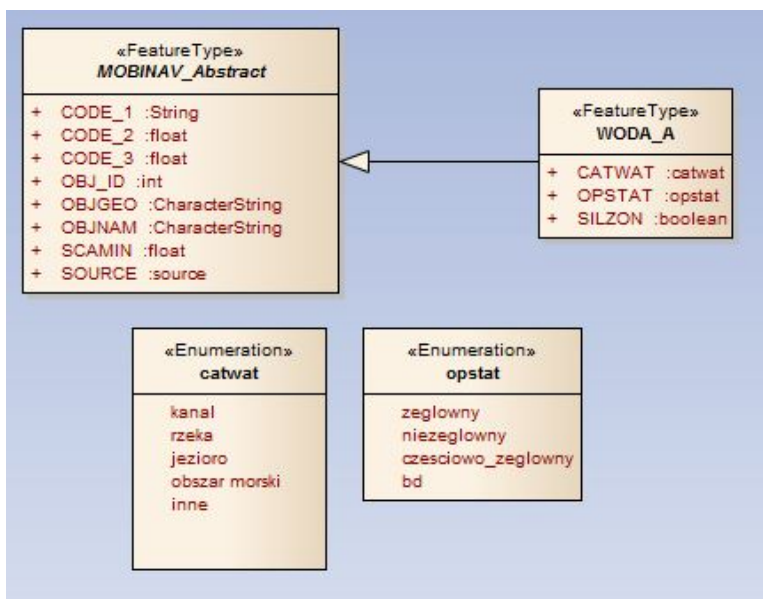


e

B



Rysunek 3. Przykład niespójności geometrycznej pomiędzy źródłami danych: IENC i Vmap L2



Rysunek 4. Fragment schematu modelu danych dla klasy WODA_A