

Tomasz Lehman

INTERNET I WORLD-WIDE-WEB NARZĘDZIEM DYSTRYBUCJI GEODANYCH ORAZ INTEGRACJI ROZPROSZONYCH SYSTEMÓW INFORMACJI PRZESTRZENNEJ

Streszczenie

Artykuł prezentuje możliwości wykorzystania ogólnoswiatowej sieci Internet i bazujących na niej serwisów do przesyłania danych przestrzennych oraz integracji rozproszonych Systemów Informacji Przestrzennej. We wstępie przedstawiono krótkie wprowadzenie do Internetu i WWW, następnie opisano kilka schematów możliwego wykorzystania go do transferu informacji. W dalszej części przedstawiono możliwości integracji różnych SIP w oparciu o system World-Wide-Web. Na zakończenie zasygnalizowano dalsze perspektywy rozwoju a także niebezpieczeństwa wiążące się ze stosowaniem tych technologii.

Słowa kluczowe: Internet, System Informacji Przestrzennej, Dystrybucja Geodanych

1. Wprowadzenie.

Internet to słowo „wytrych”, używane do określenia sieci komputerów o zasięgu światowym. Sieć komputerowa jest kolekcją komputerów, które mogą się wzajemnie ze sobą komunikować, to znaczy są zdolne do wymiany informacji w formie elektronicznej, zwykle bardzo szybko. Ogólnie można podzielić sieci komputerowe na sieci LAN (Local Area Networks) i WAN (Wide Area Networks). Podział ten jest często bardzo płynny, generalnie jednak przyjmuje się, że LAN łączy zbiór komputerów, które geograficznie leżą blisko siebie, natomiast WAN łączy komputery, lub wręcz sieci LAN, które znajdują się w znacznej odległości od siebie. Pojęcie *Internet* używane jest właśnie do określenia takiej najbardziej popularnej, ekstremalnie dużej, rozległej sieci WAN (Sokolewicz, 1996). Formalnie słowo *Internet* oznacza „sieć sieci”. Skupia on tysiące małych regionalnych sieci rozrzuconych po całym świecie. Transfer informacji w sieci Internet odbywa się przy pomocy

standardowego protokołu transmisji (TCP/IP - Transport Control Protocol/Internet Protocol). Protokół ten umożliwia komunikowanie się komputerów niezależnie od platformy sprzętowej i systemu operacyjnego (oczywiście każdy komputer w sieci musi mieć specjalne oprogramowanie, które umie czytać protokół TCP/IP).

Początki Internetu sięgają lat 60-tych, kiedy to w USA prowadzono badania finansowane przez Departament Obrony nad utworzeniem zdecentralizowanej sieci (bez komputera głównego, a co za tym idzie bez administratora) zdolnej do przetrzymania ataku jądrowego. Sieć taka zbudowana jest z komputerów, z których żaden nie jest wyróżniony w stosunku do pozostałych, a informacje przesyłane są z wykorzystaniem metody dynamicznego kierowania (dynamic routing), to znaczy sieć automatycznie steruje przepływem danych w obwodzie utworzonym przez komputery. Dzięki tej technice, zniszczenie dowolnego węzła sieci nie spowoduje zablokowania całej sieci. Dla rozważań będących przedmiotem niniejszego artykułu ważniejszy jednak jest fakt, że Internet ze względu na brak odgórnej kontroli zaczął się rozwijać tak spontanicznie i szybko, że obecnie stanowi największą sieć komputerową o zasięgu światowym (obecnie nie można nawet określić ile komputerów jest do niej podłączonych, szacunki sprzed roku mówią o 3-4 milionach komputerów umożliwiających dostęp do sieci ok. 30 milionom ludzi, ale dane te są już na pewno nieaktualne, gdyż rozwój Internetu następuje prawie w tempie geometrycznym).

Internet jest tworem fizycznym, platformą komunikacji, tzn. stanowi zbiór komputerów i kabli je łączących. Za przepływ informacji w tej sieci odpowiada natomiast szereg serwisów i systemów na jej bazie powstałych. Pierwszym serwisem Internetu, i chyba do dzisiaj najważniejszym i najczęściej używanym, mającym decydujący wpływ na jego popularność, była poczta elektroniczna (e-mail). Później powstawały nowe, z których najważniejsze to: Usenet czyli grupy dyskusyjne, listy dyskusyjne, Telnet umożliwiający wchodzenie na odległe komputery i przeszukiwanie ich zasobów, FTP (File Transfer Protocol) umożliwiający przesyłanie pomiędzy komputerami dowolnych plików, Gopher umożliwiający korzystanie z serwisów sieci za pomocą dynamicznych, rozwijalnych Menu.

W marcu 1989 roku Tim Berners-Lee, pracownik CERN (europejska organizacja skupiająca fizyków wysokich energii), zaproponował projekt rozproszonego systemu multimedialnego opartego na Internecie, umożliwiającego efektywny transfer informacji wszelkich typów wewnątrz organizacji. Tak powstał World-Wide-Web (WWW), niezależny od platformy sprzętowej, oparty o architekturę klient-serwer system dający uniwersalny dostęp do wielu rodzajów dokumentów. System ten zdobywa ogromną popularność, jest ciągle udoskonalany. Dzięki temu, oraz przez fakt, że umożliwia on korzystanie praktycznie ze wszystkich dotychczas istniejących serwisów, coraz częściej utożsamiany jest on z samym Internetem. Praktycznie WWW jest rozległą,

ogromną kolekcją wzajemnie powiązanych dokumentów (stron). Oznacza to, że każdy dokument (strona) WWW, mogący zawierać kolorową grafikę, tekst, obraz video lub dźwięk, ma możliwość połączenia się z innym dokumentem WWW bezpośrednio i natychmiast. To wzajemne powiązanie określonych stron WWW realizowane jest przez hiperłączenia, fragmenty tekstu lub grafiki posiadające odnośniki adresowe do innych dokumentów (tekstu, grafiki, audio lub video), które z kolei posiadają łączenia do następnych itd.

Obecnie system WWW umożliwia również tworzenie stron z tzw. mapami aktywnymi (będącymi w istocie rastrowymi mapami obiektowymi), których określone, zdefiniowane współrzędnymi fragmenty posiadają łączenia do innych stron, ponadto istniejący schemat formularzy (WWW Form) umożliwia aktywne przesyłanie informacji również w odwrotnym kierunku, tzn. od użytkownika do komputera udostępniającego dane. Programy CGI (Common Gateway Interface) umożliwiają wyjście poza zbiór dokumentów tworzących WWW i interaktywne przeszukiwanie baz danych a nawet ich modyfikacje.

2. Dystrybucja danych przestrzennych.

Coraz więcej danych przestrzennych dostępnych jest obecnie w postaci cyfrowej (map, obrazów i zdjęć satelitarnych, fotografii lotniczych, zbiorów tekstowych z danymi opisowymi, statystycznymi i innymi). Coraz częściej zarządza się tymi danymi poprzez Systemy Informacji Przestrzennej. Rośnie również błyskawicznie potrzeba wykorzystywania danych przestrzennych w coraz szerszych sferach zastosowań. Istnieje w związku z tym pilna potrzeba zbudowania efektywnego systemu wyszukiwania istniejących w postaci cyfrowej danych przestrzennych i sprawnego ich przesyłania pomiędzy ich dostawcami a odbiorcami. Wymienione w poprzednim punkcie możliwości i wiążące się z nimi zalety Internetu i World-Wide-Web pozwalają na stwierdzenie, że mogą one być doskonałym rozwiązaniem problemu dystrybucji danych przestrzennych do ich użytkowników i pomiędzy nimi.

W zależności od stopnia skomplikowania i sprawności/wydajności wyróżnić można trzy rozwiązania architektury systemu dystrybucji geodanych (Fortunati L. 1996):

I. dane przestrzenne dla całego obszaru geograficznego wraz z różnymi warstwami tematycznymi znajdują się w jednym lub kilku plikach. Są one udostępniane np. przez serwis FTP, to znaczy mogą być w całości i w oryginalnie kopiowane przez użytkownika. Serwer udostępniający te dane nie potrzebuje SIP do operowania nimi. Użytkownik ma jedynie dostęp do listy zbiorów. Może importować cały zbiór a następnie wydobyć interesujący go podzbiór informacji używając swojego własnego SIP. Rozwiązanie to oznacza udostępnianie w sieci informacji, które mogą nie być interesu-

jące, ale z drugiej strony jest ono najprostsze i najmniej obciąża serwis je udostępniający;

- II. dane przestrzenne na serwerze są odpowiednio zorganizowane i udostępniane w systemie WWW. Serwer geodanych nie potrzebuje SIP do operowania nimi, a jedynie odpowiednią aplikację WWW (stronę lub zestaw stron WWW). SIP jest jednak niezbędny do zbudowania odpowiedniej struktury geodanych oraz przygotowania dla aplikacji WWW odpowiedniego graficznego interfejsu użytkownika. Użytkownik może wybrać obszar go interesujący (według predefiniowanego podziału na określone fragmenty) oraz ograniczyć wybór danych do określonych zakresów tematycznych. Chociaż system zarządzania serwisem musi przygotować strukturę danych oraz interfejs użytkownika, co czyni go bardziej skomplikowanym, tylko określona, pożądana przez użytkownika informacja zostaje załadowana do sieci a użytkownik nie musi „rozpakowywać” podzbiorów informacji. Przykład takiej architektury serwisu geodanych zobaczyć można odwiedzając stronę WWW o adresie:

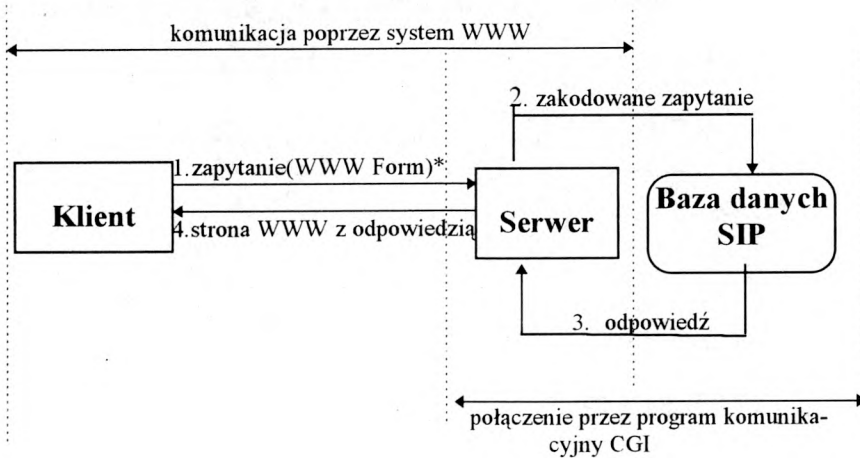
http://www.wsl.ch/rauminf/riv/datenbank/lichen/database_lichen.html;

- III. serwis geodanych wykorzystuje SIP do zarządzania danymi. System zarządzający serwisem przygotowuje w sposób dynamiczny strony WWW będące graficznym interfejsem użytkownika oraz tworzy gateway (program komunikacyjny pomiędzy serwerem a bazą danych SIP). WWW zarządza zapytaniami, koduje je (np. w standardzie GSQL - General Standard Query Language, będącego rozwinięciem znanego standardu relacyjnych baz danych SQL, utworzonego specjalnie w celu adaptacji istniejących baz danych i WWW [Sokolewicz, 1996]), a następnie przesyła zapytanie do bazy danych SIP. Dane spełniające kryteria określone w zapytaniu wracają do serwera WWW i są przez niego przekazywane dalej do użytkownika w postaci nowej strony WWW. Uproszczony schemat połączeń realizujących tą architekturę przedstawia rysunek 1.

Przykład takiej architektury serwisu geodanych zobaczyć można odwiedzając stronę WWW o adresie:

http://www-a5.igd.fhg.de/demos/oga/hist_intro.html lub:

<http://www-a5.igd.fhg.de/demos/oga/geodemo/>



*Zapytania przesyłane są do serwera z wykorzystaniem opisanego dalej (pkt 3) systemu formularzy WWW (WWW Form).

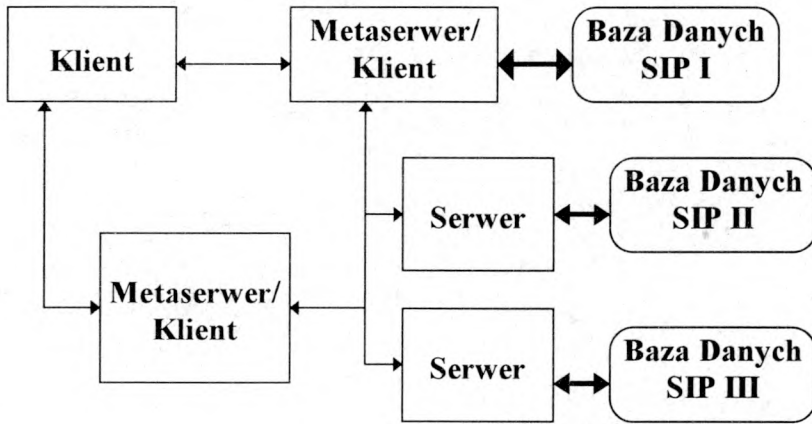
Rys. 1 Schemat integracji SIP z systemem WWW.

3. Integracja systemów rozproszonych.

Ostatni z wyżej opisanych schematów serwisu geodanych realizuje właściwie otwartą architekturę Systemu Informacji Przestrzennej. Z systemu korzystać może dowolny użytkownik, jeżeli tylko jest uprawniony do korzystania z niego i oczywiście ma dostęp do Internetu. Architektura taka umożliwia również integrację baz danych wielu rozproszonych systemów informacji przestrzennej. Skoro z takiego systemu może korzystać dowolny użytkownik, to użytkownikiem takim może być również inny SIP. W ten sposób różne systemy informacji przestrzennej mogą komunikować się między sobą, przysyłać zapytania, wymieniać dane, i to wszystko niezależnie od platformy sprzętowej i programowej. Realizacja takiej architektury przyniosłaby wiele korzyści, z których najważniejsze to niewątpliwie znaczna redukcja duplikacji prac związanych z prowadzeniem SIP, w szczególności w zakresie pozyskiwania danych oraz kapitalne zwiększenie dostępu do ogromnej liczby informacji w systemach rozproszonych. Schemat takiej architektury przedstawia rysunek 2.

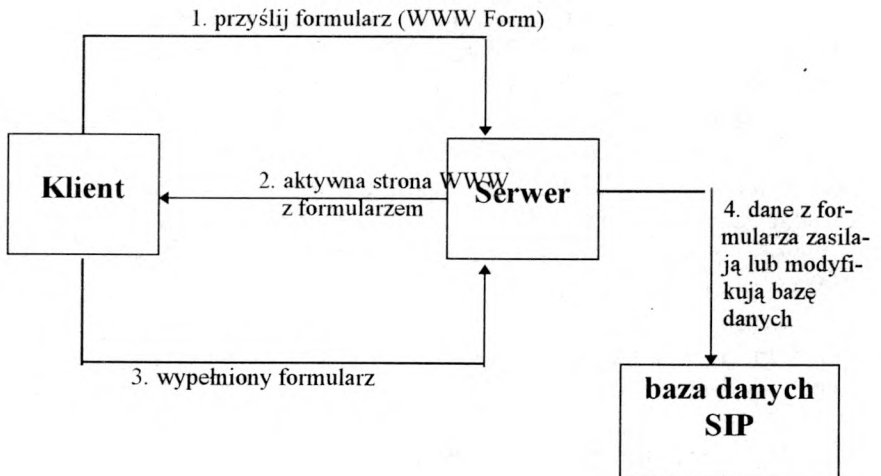
Bazy danych systemów rozproszonych można za pomocą WWW nie tylko przeglądać, ale i modyfikować. Umożliwia to obecnie system formularzy WWW (WWW Form). Odbywa się to tak, że system zarządzający bazą danych przesyła do użytkownika specjalną stronę WWW, do której użyt-

kownik może wprowadzać własne dane. Strona ta następnie wraca do serwera, a za-



Rys.2 Schemat otwartej architektury SIP.

warte w niej informacje poprzez program komunikacyjny CGI (gateway) mogą zasilić lub zmodyfikować bazę danych. Schemat tego rozwiązania przedstawia rysunek 3.



Rys. 3. Schemat połączeń realizujących zasilenie lub modyfikację bazy danych

4. Wymagania dotyczące dalszego rozwoju.

Prawidłowy rozwój wyżej opisanych technologii w kierunku pełnej integracji WWW i SIP wymaga jednak pokonania lub ulepszenia rozwiązań niżej wymienionych problemów:

- biorąc pod uwagę, że Internet definiuje i realizuje standard transferu a nie standard danych - standard danych musi być określony aby efektywnie korzystać z możliwości sieci. Istnieje obecnie wiele standardów danych przestrzennych, takich jak niemiecki EDDBS (Einheitliche Datenbankschnittstelle), amerykański SDTS (Spatial Data Transfer Standard) czy polski SWING (System Wymiany Informacji Geodezyjnych). Standardy te są jednak zamknięte w obrębie krajów je stosujących i nie mogą się wzajemnie komunikować. Istnieje w związku z tym potrzeba określenia uniwersalnego standardu lub metod komunikacji pomiędzy już istniejącymi;
- z powyższym zagadnieniem wiąże się konieczność opracowania standardu metadanych, czyli danych o danych lub informacji opisujących dane. Jest to niezbędne, aby można było wyszukiwać odpowiednie, interesujące użytkownika dane;
- zapewnienie bezpieczeństwa danych. Ogólnodostępność i otwartość sieci będąca jej główną zaletą, w tym aspekcie stanowi poważne niebezpieczeństwo. Należy rozwinąć i udoskonalić techniki zabezpieczające, eliminujące zagrożenie nieuprawnionego przejęcia udostępnianych lub przesyłanych danych, a przede wszystkim możliwość ich zniszczenia lub uszkodzenia;
- zasady regulowania należności za udostępnione dane. Dążyć się powinno do sytuacji, aby cały proces wymiany informacji (a więc i płacenie za nie) mógł odbywać się elektronicznie, to znaczy za pośrednictwem sieci (wiąże się to również z zasygnalizowanym wyżej problem bezpieczeństwa);
- rozwój systemu WWW w kierunku realizowania funkcjonalności typowej dla interfejsu użytkownika systemu SIP, takich jak: transformacja odwzorowań i konwersja danych różnych typów, wykonywanie operacji analiz przestrzennych (buforowanie, przecięcia, nakładki tematyczne itd.)

5. Uwagi końcowe.

Powszechna akceptacja technologii WWW - w tym również a może przede wszystkim przez pracowników naukowych, pozwala oczekiwać, że znaczenie jej również w dziedzinie SIP i dystrybucji danych przestrzennych będzie coraz większe. Doprowadzi to z pewnością do możliwości korzystania z ogromnej ilości informacji przechowywanej w różnych systemach, ułatwi i

przyspieszy jej transfer, zredukuje dublowanie prac, szczególnie w zakresie pozyskiwania danych, uniezależni wreszcie użytkowników od platform sprzętowych i programowych.

Literatura:

- Fortunati L., 1996, „Geo-Data Server on WWW”, Second Join European Conference & Exhibition on Geographical Information, Barcelona, Hiszpania.
- Hägeli M., Bösch R., Wey J., 1996, „Connecting WWW, GIS and RDBMS”, Second Join European Conference & Exhibition on Geographical Information, Barcelona, Hiszpania.
- Sokolewicz M., 1996, „Information Marketing on the Internet”, First International Conference On Visual Information Systems, Melbourne, Australia.
- Orfali R., Harkey D., Edwards J., 1996, „The Essential Client/Server Survival Guide - Second Edition”, John Wiley & Sons, Inc., Nowy Jork, USA.
- Hughes K., 1993, „Entering the World-Wide-Web: A Guide to Cyberspace”, Honolulu Community College, wersja elektroniczna dostępna pod <http://pulua.hcc.hawaii.edu/> oraz informacje w wersji elektronicznej dostępne pod adresem WWW: <http://www-a5.igd.fhg.de/>

Internet and World-Wide-Web as a tool for distribution of spatial data and for integration of distributed Geographic Information Systems.

Summary:

The paper presents possibilities of using Internet and its services to transferring spatial data and to integrating distributed Geographic Information Systems (GIS). First basic information about Internet and World-Wide-Web (WWW) are given, then certain schemes of using them for distribution of spatial information are presented. Next possibilities of using WWW for integration of different GIS are described. The paper concludes with a brief glimpse of the future of this technology and dangers of its use.

Keywords: Internet, Geographic Information System, Spatial Data Distribution

Recenzował: Prof. dr hab.inż. Zbigniew Sitek