

Metody lokalizacji użytkowników w sieci Internet

Janusz Stal

Akademia Ekonomiczna w Krakowie

Streszczenie Geolokalizacja to metoda określania położenia geograficznego (miejscowości, regionu, czy kraju) użytkownika internetu. Artykuł ten prezentuje cele oraz korzyści wynikające z zastosowania geolokalizacji, a także metody umożliwiające identyfikację położenia geograficznego internautów.

Summary Geolocation is the technique of determining geographic location (city, region and country) of a web site visitor. This article presents some of the reasons for and benefits of using geolocation as well as several techniques identifying visitor's geographical location.

Wprowadzenie

Wraz z rozwojem internetu, coraz bardziej znacząca część prowadzonej działalności gospodarczej realizowana jest w formie elektronicznej. Dla zwiększenia jej skuteczności konieczne jest wypracowanie metod dotarcia do indywidualnego konsumenta podobnie, jak w przypadku tradycyjnych form sprzedaży, czy reklamy prowadzonej na określonym obszarze. Możliwość identyfikacji użytkownika ze względu na jego położenie pozwoli z jednej strony określić miarę popularności witryny internetowej (a co za tym idzie – podmiotu gospodarczego) na danym obszarze, z drugiej zaś – skierować do klienta (konsumenta) reklamę produktów do niego kierowaną (marketing celowy). Łatwo sobie wyobrazić, iż dla podmiotów gospodarczych prowadzących działalność na danym obszarze (hurtownie, sklepy, punkty usługowe, lokale gastronomiczne,...) znacznie skuteczniejszy będzie przekaz reklamowy skierowany jedynie do mieszkańców z obszaru, na którym działa firma, co pozwoli na precyzyjne dotarcie do grupy docelowej i, w końcowym efekcie, minimalizację kosztów.

Geolokalizacja to metoda ustalania położenia geograficznego internauty, przyłączonego do sieci. Dzięki niej możliwy jest podział użytkowników:

- wg kraju (podział na użytkowników polskich i zagranicznych),
- wg kontynentu,
- wg województw (w przypadku użytkowników z Polski),
- wg ośrodków miejskich.

Tak uzyskane informacje pozwalają na podjęcie szeregu działań, które w efekcie mogą przyczynić się do zwiększenia efektywności prowadzonej działalności gospodarczej. Potencjalne korzyści wynikające ze stosowania mechanizmów geolokalizacji można ująć następująco:

- **Statystyka.** W przypadku informacji związanych z lokalizacją użytkownika, pozwala na ustalenie, z jakich regionów (państwa, miasta) dokonywane są połączenia, a zatem jakie jest zainteresowanie produktem, czy usługą na danym terenie. Dzięki temu możliwe jest dokonanie analizy skutków przeprowadzanych kampanii reklamowych w danym rejonie (analiza skuteczności działań marketingowych) i może być bodźcem do zmiany strategii reklamowych przeprowadzanych na tych terenach.
- **Wybór języka.** W zależności od miejsca połączenia możliwe jest dostosowanie przekazu (wyświetlanej treści) do języka użytkownika, czy też wyświetlanie cen towarów i usług w walucie kraju, z którego łączy się użytkownik.
- **Geotargeting.** Pozwala na kierowanie produktu (przekazu reklamowego) bezpośrednio do grupy docelowej identyfikowanej poprzez położenie geograficzne co w szczególności umożliwia zmniejszenie kosztów prowadzonej kampanii reklamowej.
- **Ochrona przed spamem.** W przypadku komunikacji z użytkownikiem poprzez pocztę elektroniczną, możliwe jest ograniczenie liczby niepożądanego korespondencji z miejsc, z których są najczęściej wysyłane.
- **Ochrona płatności.** Płatności za produkty, czy usługi realizowane są coraz częściej w formie bezgotówkowej, z użyciem kart płatniczych. Monitorowanie i analiza miejsc, skąd są one dokonywane może umożliwić wychwycenie nieprawidłowości i przyczynić się do wykrycia transakcji realizowanych z użyciem kart skradzionych.
- **Ograniczenie dostępu.** Ustalenie możliwości dostępu do danych tylko dla użytkowników z wybranego obszaru lub też z wyłączeniem tych użytkowników.

- **Przekierowania.** W przypadku pomiotów prowadzących działalność gospodarczą na znacznym obszarze, w zależności od położenia internauty, może on zostać przekierowany na serwer zlokalizowany bliżej miejsca, z którego nastąpiło nawiązanie połączenia. Pozwala to z jednej strony zmniejszyć obciążenie serwera głównego firmy, a z drugiej strony może spowodować wzrost prędkości wyświetlanych informacji na komputerze klienta.
- **Autoryzacja dostępu.** Szereg usług dostępnych za pośrednictwem sieci (np. usługi dostępu do zasobów wiedzy, słowników, encyklopedii, porad prawnych, czy finansowych), po uiszczeniu opłaty, wymaga podania identyfikatora i hasła dostępu. Monitorowanie miejsc, z których łączy się internauta pozwala zminimalizować zjawisko posługiwania się identyfikatorem i hasłem przez osoby nieuprawnione.

Metody określania lokalizacji

Każdy host w sieci internet identyfikowany jest poprzez unikalny¹ adres IP (Internet Protocol) reprezentowany jako 32 bitowa liczba binarna, zapisywana w postaci czterech oktetów z zakresu 0-255. Przydzielany jest on przez dostawców usług internetowych (ISP – Internet Service Provider) i na jego podstawie powinna zostać ustalona lokalizacja (położenie geograficzne), w jakiej znajduje się klient. Zasadniczy problem, jaki się tu pojawia to brak relacji pomiędzy adresem IP, a położeniem geograficznym hosta. Ten brak wsparcia ze strony architektury sieci powoduje, iż nie istnieje jedyna, skuteczna metoda ustalania lokalizacji hosta.

Baza WHOIS

Przydział adresów IP końcowemu użytkownikowi realizowany jest zwykle przez LIR (Local Internet Registers), natomiast nadzorem, dystrybucją, koordynacją oraz rejestracją publicznych zasobów adresów IP zajmują się RIR (Regional Internet Registers), wśród których należy wymienić: APNIC (Asia Pacific Network Information Centre), ARIN (American Registry for Internet Numbers), LACNIC (Latin American and Caribbean Internet Addresses Registry), czy też w przypadku Europy – RIPE NCC (RIPE Network Coordination Centre) [10].

Każda przydzielona przestrzeń adresowa jest rejestrowana w bazach RIR WHOIS, zawierających: przydzielony adres lub zakres adresów IP, nazwy sieci, podsieci, nazwę organizacji czy też klientów indywidualnych, dane kontaktowe (adres, miejscowość telefon), czy też dane organizacji zarządzającej. Rejestracja danych następuje w bazie WHOIS w zależności od regionu, w którym znajduje się klient (Tabela 2):

¹ z wyjątkiem hostów, którym przydzielono prywatny adres IP, najczęściej z użyciem NAT (Network Address Translation).

Tabela 2 Wykaz instytucji zarządzających przydziałem i rejestracją adresów IP.

Instytucja	Region	WWW	WHOIS
RIPE	Europa, Afryka Płn.	www.ripe.net	whois.ripe.net
ARIN	Ameryka Płn.	www.arin.net	whois.arin.net
APNIC	Azja, Oceania	www.apnic.net	whois.apnic.net
LACNIC	Ameryka Łac.	www.lacnic.net	whois.lacnic.net
AFRINIC	Afryka	www.afrinic.net	whois.afrinic.net

Źródło: opracowanie własne.

Poprzez dostępne mechanizmy wyszukiwania w bazach WHOIS możliwe jest zatem, przy znajomości adresu IP, uzyskanie informacji o danych jego dysponenta. Analiza uzyskanych w ten sposób informacji może przyczynić się do ustalenia lokalizacji (kraj, region, miejscowość), w której znajduje się host identyfikowany przez badany adres IP.

Pomimo możliwości dostępu do baz RIR WHOIS, mogą istnieć trudności w uzyskaniu stosownych informacji pozwalających na ustalenie lokalizacji. Wymienić tu należy:

- możliwość zastrzeżenia dostępu do danych lub ich ukrycia dla wybranych rekordów, na życzenie użytkownika – dysponenta adresów,
- ograniczenia prawne (brak możliwości wykorzystania uzyskanych informacji w zastosowaniach komercyjnych),
- narzucone przez instytucję nadzorującą, fizyczne ograniczenia w dostępie do danych (możliwość odczytu ograniczonej liczby rekordów w zadanej jednostce czasu, np. 100 rekordów w ciągu doby),
- rozdział adresów IP (z przydzielonej przestrzeni adresowej) użytkownikom końcowym znajdującym się w różnych lokalizacjach.

Biorąc to pod uwagę, wykorzystanie bazy WHOIS do określania lokalizacji hostów nie może stanowić jedynej, skutecznej metody. W sytuacji, gdy dla badanego adresu IP nie jesteśmy w stanie uzyskać informacji z bazy WHOIS możemy również dokonać śledzenia trasy przepływu pakietów (traceroute). Uzyskując adres IP hosta położonego najbliżej badanego, możemy dokonać odczytu jego danych z bazy WHOIS i w ten sposób z pewnym przybliżeniem uzyskać informacje o lokalizacji hosta badanego.

Domeny

Adresy IP hostów, prezentowane w postaci numerycznej, w codziennej praktyce stają się zupełnie niewygodne w użytkowaniu. Stąd możliwe jest skojarzenie z każdym z adresów IP jego słownego odpowiednika. Musi istnieć jednakże proces odwrotny, tj. zamiana nazwy na odpowiadający jej adres IP. Funkcję tę pełni DNS (Domain Name System), w skład którego wchodzi hierarchiczna, rozproszona baza danych zawierająca skojarzone nazwy oraz odpowiadające im adresy IP. Decentralizacja systemu realizowana jest poprzez delegację domen.

DNS w swej istocie nie zawiera mechanizmów identyfikujących użytkownika (właściciela domeny) ze względu na jego położenie. Z tego też powodu nie nadaje się do bezpośredniego wykorzystania. Jednakże sposób nazewnictwa przyjęty przy nadawaniu domen może nieść informacje o lokalizacji hosta.

Domeny² najwyższego poziomu (Top Level Domains TLDs) (.pl, .au, .de, ...) mogą zawierać kody krajów, które odpowiadają poszczególnym państwom³, co zgodne jest z międzynarodowym standardem ISO3166⁴. Zatem na ich podstawie możliwe jest ustalenie kraju, w którym potencjalnie może znajdować się host. Domeny niższego poziomu zawierają często nazwę miejscowości (np. .krakow.pl, .poznan.pl, ...) co umożliwi bardziej precyzyjne określenie lokalizacji hosta. Również w przypadku hostów pełniących funkcję routerów zwyczajowo umieszcza się w domenie nazwę pozwalającą na określenie jego przybliżonego położenia, np.

```
kabaty-gw-at2-0-1.warman.nask.pl
```

```
z-lodz.gw.lod.pozman.622.pol34.pl
```

```
z-war.ar9..do.war-r4.tpnet.pl
```

Należy jednak zwrócić uwagę, iż takie umieszczanie nazw miejscowości w nazwie domeny nie jest obligatoryjne. Ponadto, możliwe jest rejestrowanie nowej domeny, dla której TLDs nie będzie wyznacznikiem kraju rejestracji (możliwość rejestrowania w Polsce subdomen w domenach .com., czy .net.).

DNS

Odmienne podejście do szacowania położenia hosta zaprezentowane zostało w pracy [1], gdzie wykorzystane zostały informacje wprowadzone do baz serwera DNS (Domain Name Server). Definiując nowy typ rekordu DNS LOC (Location) dopuszcza się możliwość wprowadzania informacji o położeniu geograficznym hosta (długość i szerokość geograficzna) wskazywanego przez tą domenę. Zatem ustalając delegację domeny wprowadzane są również informacje o lokalizacji hosta wskazywanego przez tą domenę. Stąd, w dowolnym momencie, poprzez sprawdzenie zapisu w bazie danych DNS możliwe będzie uzyskanie informacji o położeniu hosta. Struktura uzyskanych informacji z DNS LOC przedstawia się następująco (dla uproszczenia ograniczono się jedynie do danych określających współrzędne geograficzne):

```
$host -t LOC domain (komenda dostępu do DNS LOC)
fu-berlin.de      LOC  52 27 19.591 N 13 17 40.978 E
nikhef.nl        LOC  52 21 24.000 N  4 57 05.000 E
ckdhr.com        LOC  42 21 43.528 N 71 05 06.284 W
yahoo.com        LOC  37 23 30.900 N 121 59 19.000 W
fukt.hk-r.se     LOC  56 10 55.000 N 15 35 39.000 E
```

gdzie przykładowo zapis „56 10 55.000 N 15 35 39.000 E” oznacza 52 stopnie, 14 minut i 5 sekund szerokości geograficznej północnej, 0 stopni, 8 minut i 50 sekund długości geograficznej wschodniej hosta wskazywanego przez domenę „fukt.hk-r.de”.

Należy jednak zauważyć, iż proponowana metoda wiąże się z koniecznością zmian w infrastrukturze sieciowej i wymaga od administratorów serwerów DNS wprowadzania informacji o położeniu geograficznym dysponentów domeny w rekordach DNS LOC. Z tego też powodu może być trudna do realizacji i dlatego obecnie niewielka liczba serwerów DNS posiada wpisy dotyczące lokalizacji hostów wskazywanych przez domeny.

Warto dodać, iż nie jest konieczne wprowadzanie informacji o każdym hoście przyłączonym do sieci. Wystarczy zrobić to dla podsieci lub też określić lokalizację najbliższego routera, co może w zupełności wystarczające dla potrzeb geolokalizacji. Dodatkową korzyścią z określenia położenia geograficznego w DNS LOC dla routerów jest możliwość śledzenia dokładnej trasy przepływu pakietów danych i możliwość jej wizualizacji.

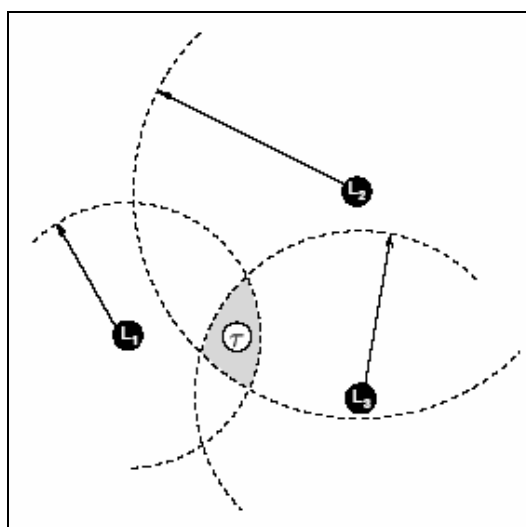
² Zdecydowana większość domen na terenie Polski rejestrowana jest przez NASK (Naukowa i Akademicka Sieć Komputerowa). <http://www.nask.pl>

³ z wyłączeniem Wielkiej Brytanii

⁴ Aktualna lista tych państw dostępna jest pod adresem <ftp://ftp.ripe.net/iso3166-countrycodes/>

CBG

Zupełnie odmienne podejście w ustalaniu lokalizacji hostów zostało zaprezentowane w pracy [5] Metoda CBG (Constraint-Based Geolocation) określa położenie geograficzne hosta przy użyciu hostów o znanej lokalizacji, podobnie jak w przypadku systemu GPS (Global Positioning System). W przypadku GPS, położenie geograficzne odbiornika ustalane jest na podstawie pomiaru czasu przepływu sygnału z poszczególnych satelitów, natomiast CBG wykorzystuje do tego celu hosty o ustalonym położeniu geograficznym. Przy założeniu korelacji pomiędzy znaną lokalizacją hostów, a miarą opóźnienia w przesyłaniu informacji pomiędzy nimi i badanym hostem, możliwe jest wyznaczenie dystansu je dzielącego, a co za tym idzie, określenie położenia geograficznego badanego hosta (Rys. 6).



Rys. 6. Wyznaczanie położenia hosta na podstawie pomiaru odległości od hostów o znanej lokalizacji (metoda CBG (L_1 , L_2 , L_3 – hosty o ustalonej lokalizacji, τ - badany host) .

Źródło: opracowane na podstawie [5].

Przeprowadzone badania stwierdzają, iż za pomocą metody CBG możliwe jest uzyskanie precyzji w ustalaniu lokalizacji hosta na poziomie poniżej 25km w przypadku Europy Zachodniej oraz poniżej 100km w przypadku Ameryki Płn. Należy jednak zauważyć, iż:

- Prezentowane badania zostały przeprowadzone w rejonach o rozbudowanej infrastrukturze sieciowej i wysokim poziomie zaawansowania technologicznego. W przypadku zastosowania CBG w innych rejonach, miara opóźnień w przesyłaniu sygnału z badanego hosta do hostów o ustalonych lokalizacjach może być uzależniona od jakości zastosowanych rozwiązań technologicznych (przepustowość sieci).
- Zastosowane podejście pozwala określić jedynie region, w którym zlokalizowany jest badany host. W celu bardziej precyzyjnego określenia jego lokalizacji (miejscowość) autorzy proponują skorzystanie z innych metod, które pozwalały by bardziej precyzyjnie ustalić położenie hosta.

Niemniej jednak zaprezentowana metoda, wraz ze wzrostem poziomu zaawansowania technologicznego wykorzystywanego przy budowie sieci może umożliwić w przyszłości bardziej precyzyjne ustalanie lokalizacji hostów.

Ograniczenia w ustalaniu lokalizacji

Podane metody ustalania lokalizacji (położenia geograficznego) hosta bazują na jego adresie IP. Jak już zaznaczono na wstępie, nie umożliwiają w każdym przypadku jednoznacznego

wyznaczenia jego położenia. Dodatkowe przeszkody mogące utrudnić lub też całkowicie uniemożliwić ustalenie lokalizacji to:

- **Serwery proxy.** Ich zadaniem jest zmniejszenie ruchu na łączach poprzez buforowanie najczęściej pobieranych plików, udostępnianych następnie użytkownikom. Z tego też powodu, użytkownicy korzystający z serwera proxy identyfikowani są w sieci poprzez jego adres IP. Jeśli serwer proxy znajduje się w innej lokalizacji niż użytkownik, ustalenie położenia będzie błędne. Stąd, najczęściej do systemów lokalizacyjnych wprowadza się adresy IP serwerów proxy, aby w ten sposób wyeliminować takie przypadki. Należy tu nadmienić, iż dla przyspieszenia ładowania się stron internetowych, celowe wydaje się korzystanie z serwera proxy znajdującego się tej samej lokalizacji co host użytkownika.
- **Dostęp bezprzewodowy.** Szybki rozwój technologii mobilnej umożliwiającej dostęp do sieci przy wykorzystaniu urządzeń bezprzewodowych (GPRS (General Packet Radio System), EDGE (Enhanced Data for Global Evolution), czy UMTS (Universal Mobile Telecommunication Systems)) może w przyszłości utrudnić, lub wręcz uniemożliwić poprawne określenie lokalizacji hosta, szczególnie w przypadku, gdy podstawą będzie bazowanie na adresie IP hosta.
- **dynamiczne IP.** Użytkownik korzystający z dostępu do sieci musi posiadać adres IP, który może być adresem statycznym (przypisanym jednoznacznie do użytkownika) lub nadany dynamicznie, każdorazowo w momencie przyłączenia do sieci. Szczególnie w tym drugim przypadku określenie lokalizacji hosta przy użyciu adresu IP będzie niezmiernie utrudnione.

Ustalanie położenia geograficznego

W niniejszym punkcie pracy podjęto próbę praktycznej realizacji idei określania lokalizacji hosta na podstawie informacji zawartych w nazwie domeny oraz przy wykorzystaniu informacji umieszczonych w rekordach bazy WHOIS.

Analizie poddane zostały dane uzyskane w procesie rejestracji kandydatów na studia w Akademii Ekonomicznej w Krakowie (AE) w roku akademickim 2004/2005. Każdy z kandydatów zobligowany został do wypełnienia formularza rejestracyjnego na studia, udostępnionego w sieci internet na serwerze WWW AE. W ten sposób uzyskano również informacje związane z komputerem, z którego dokonano rejestracji, tj. jego nazwę, numer IP oraz nazwę domeny. Struktura połączeń z serwerem WWW AE została przedstawiona w tabeli (Tabela 3) (ze względu na objętość danych, ograniczono się do zaprezentowania listy domen, z których połączenia były realizowane najczęściej).

Tabela 3 Struktura połączeń z serwerem AE w Krakowie w procesie rejestracji kandydatów na studia.

Liczba rejestracji z domeny		
.ppp.tpnet.pl	1679	19,3%
.ae.krakow.pl	1559	17,9%
.adsl.tpnet.pl	1374	15,8%
.internetdsl.tpnet.pl	1070	12,3%
.sdi.tpnet.pl	579	6,6%
.visp.energis.pl	126	1,4%
.unregistered.ptc.pl	116	1,3%
.crowley.pl	93	1,1%
.nowytarg.pl	90	1,0%
.chello.pl	76	0,9%
.autocom.pl	64	0,7%
.compower.pl	52	0,6%
.agh.edu.pl	50	0,6%
.net.autocom.pl	48	0,6%
.rtk.net.pl	43	0,5%
.wsh.pl	37	0,4%
.77.classcom.pl	34	0,4%
.is.net.pl	34	0,4%
.dialup.inetia.pl	29	0,3%
.brzesko.net.pl	25	0,3%

Źródło: opracowanie własne.

Już pobieżna analiza przedstawionych danych pozwala stwierdzić, iż zdecydowana większość użytkowników (kandydatów na studia) korzystała z usług dostępu do internetu (zarówno komutowanego jak i szerokopasmowego) oferowanych przez Telekomunikację Polską (TP) (domeny ppp.tpnet.pl, adsl.tpnet.pl, internetdsl.tpnet.pl, sdi.tpnet.pl – 54%). Zatem możliwość określania lokalizacji hostów znajdujących się w domenach TP należy uznać za kluczową w procesie ustalania położenia geograficznego użytkowników internetu. Tabela (Tabela 4) zawiera wykaz domen, z których nawiązywano połączenie z serwerem WWW AE realizowanych przy wykorzystaniu dostępu komutowanego oraz szerokopasmowego – usługi SDI (Szybki Dostęp do Internetu).

Tabela 4 Struktura połączeń z serwerem WWW AE realizowanych przy wykorzystaniu dostępu komutowanego oraz szerokopasmowego (SDI).

Liczba rejestracji z domeny		Liczba rejestracji z domeny	
.krakow.cvx.ppp.tpnet.pl	341	.wadowice.sdi.tpnet.pl	25
.bielsko.cvx.ppp.tpnet.pl	264	.kielce.sdi.tpnet.pl	20
.nowysacz.cvx.ppp.tpnet.pl	264	.krakow.sdi.tpnet.pl	19
.kielce.cvx.ppp.tpnet.pl	153	.ropczyce.sdi.tpnet.pl	18
.katowice.cvx.ppp.tpnet.pl	118	.tarnow.sdi.tpnet.pl	17
.tarnow.cvx.ppp.tpnet.pl	87	.rzeszow.sdi.tpnet.pl	14
.rzeszow.cvx.ppp.tpnet.pl	85	.debica.sdi.tpnet.pl	13
.tarnobrzeg.cvx.ppp.tpnet.pl	63	.jaslo.sdi.tpnet.pl	13
.krosno.cvx.ppp.tpnet.pl	55	.przemysl.sdi.tpnet.pl	13
.czestochowa.cvx.ppp.tpnet.pl	42	.limanowa.sdi.tpnet.pl	12
.radom.cvx.ppp.tpnet.pl	29	.krosno.sdi.tpnet.pl	10
.piotrkow.cvx.ppp.tpnet.pl	27	.czestochowa.sdi.tpnet.pl	9
.chelm.cvx.ppp.tpnet.pl	23	.nowy-sacz.sdi.tpnet.pl	9
.zamosc.cvx.ppp.tpnet.pl	23	.bochnia.sdi.tpnet.pl	8
.warszawa.cvx.ppp.tpnet.pl	21	.nowysacz.sdi.tpnet.pl	8
.przemysl.cvx.ppp.tpnet.pl	19	.grybow.sdi.tpnet.pl	7

Źródło: opracowanie własne.

Przyjęta, w przypadku dostępu komutowanego TP, polityka nazewnictwa domen, w których zawarte zostały nazwy większych miejscowości, pozwala na ustalenie położenia hosta z dokładnością do kilkunastu/kilkudziesięciu kilometrów. Użyte miejscowości zgodne są z nazwami miast wojewódzkich z poprzedniego podziału administracyjnego kraju (49 województw). Większą precyzję w ustalaniu lokalizacji hosta można uzyskać w przypadku użytkowników internetu korzystających z dostępu szerokopasmowego – SDI. W tym przypadku w nazwie domeny zawarte zostały również nazwy mniejszych miejscowości, co pozwala na uzyskanie większej precyzji w określaniu położenia geograficznego internautów. Należy zauważyć, iż w chwili obecnej dominującą metodą dostępu do internetu pozostaje nadal połączenie modemowe (dial-up). Jednakże ze względu na zwiększające się wymagania użytkowników odnośnie szybkości połączeń, w niedalekiej przyszłości zdecydowana liczba połączeń realizowana będzie przy wykorzystaniu dostępu szerokopasmowego.

TP udostępnia w swojej ofercie metody szerokopasmowego dostępu do internetu oparte na technologii DSL (Digital Subscriber Line), w szczególności są to Internet DSL (IDSL) oraz ADSL (Asymmetric Digital Subscriber Line), będąca odmianą DSL, zwana popularnie Neostradą. Dokonując analizy nazw domen używanych w przypadku IDSL należy stwierdzić, iż nie zostały zawarte w nich informacje o lokalizacji hostów, korzystających z tej technologii. Jednakże ze względu na specyfikę IDSL (abonent uzyskuje tu statyczny adres/adresy IP), w rekordach bazy WHOIS znajdują się nazwy miejscowości określających lokalizację hostów. Odczyt tych danych umożliwi zatem z dużym przybliżeniem ustalenie położenia geograficznego użytkowników internetu korzystających z tego typu dostępu.

```
% This is the RIPE Whois query server #1.
inetnum:      80.55.164.20 - 80.55.164.23
netname:      PIETREWICZSUWAKI
descr:        SUWALKI
descr:        POLAND
country:      PL
```

Najgorzej przedstawia się sytuacja w przypadku ADSL (neostrada). Analiza domeny nie pozwala na ustalenie lokalizacji hosta natomiast informacje zawarte w rekordach bazy WHOIS pozwalają na ustalenie najczęściej tylko nazwy obecnego województwa. Taką dokładność nie można uznać za wystarczającą i należało by podjąć inne działania, aby z większą dokładnością możliwe było ustalenie położenia hosta, a co za tym idzie, określenia położenia geograficznego internauty.

```
% This is the RIPE Whois query server #1.
inetnum:      83.30.0.0 - 83.30.255.255
netname:      NEOSTRADA-ADSL
descr:        Neostrada Plus
descr:        Katowice
country:      PL
```

Na koniec należy dodać, iż w przypadku małych, lokalnych sieci, które obejmują swoim zasięgiem niewielki obszar (najczęściej mniejszą miejscowość, dzielnicę większego miasta, czy nawet budynki położone przy jednej ulicy) praktyka pokazuje, iż możliwe jest w miarę dokładne (nazwa miejscowości oraz często również pełny adres) określenie położenia hosta na podstawie wpisów zawartych w rekordach bazy WHOIS.

Wnioski

Dynamiczny rozwój internetu powoduje, że z roku na rok coraz bardziej znacząca część działalności gospodarczej realizowana jest w formie elektronicznej. Powoła to na monitorowanie ruchu na witrynach internetowych, czy też analizę zachowań internautów odwiedzających daną witrynę. Zebrane w ten sposób informacje stanowią podstawę do wyliczenia szeregu miar i wskaźników pozwalających na tworzenie statystyk opisujących charakter ruchu na stronie. Kluczową rolę w ostatnim okresie odgrywa pozyskiwanie informacji związanych z miejscem nawiązania połączenia z witryną internetową. Dzięki temu możliwe jest podjęcie szeregu działań, w szczególności związanych z formą i treścią przekazu, co w konsekwencji prowadzić może do zwiększenia efektywności i skuteczności prowadzonej działalności gospodarczej za pośrednictwem sieci. Należy tu jednak zaznaczyć, iż ze względu na brak wsparcia dla tych działań ze strony infrastruktury sieciowej, metody ustalania lokalizacji obciążone mogą być błędem. Pomimo tego, wydaje się, iż geolokalizacja może stanowić kolejny etap w rozwoju biznesu elektronicznego.

Literatura

- [1] C. Davis, P. Vixie, T. Googwin, I. Dickinson [1996], „A means for expressing location information in the domain name system”, Internet RFC 1876, <ftp://ftp.ripe.net/rfc/>
- [2] Y. Rekhter, B. Moskowitz, D. Karrenberg, G. J. de Groot, E. Lear [1996], „Address allocation for private internets”, Internet RFC, <ftp://ftp.ripe.net/rfc/>
- [3] Instytut Badania Rynku i Opinii Publicznej SMG/KRC , <http://www.smgkrc.pl>
- [4] A. Turner [2004], „Geolocation by IP address”, Linux Journal, <http://www.linuxjournal.com/article/7856>
- [5] The Kelsey Group (TKG), <http://www.kelseygroup.com/>
- [6] B. Gueye, A. Ziviani, M. Crovella, S. Fdida [2004], „Constraint-Based Geolocation of Internet Hosts”, ACM Internet Measurement Conference
- [7] URL: <http://www.ckdhr.com/dns-loc/>
- [8] MaxMind, GeoIP, <http://www.maxmind.com/geoip/>
- [9] Quova Inc., GeoPoint, <http://www.quova.com/>
- [10] RIPE Network Coordination Centre, <http://www.ripe.net/>
- [11] Gemius S.A., GemiusTraffic, <http://www.gemius.pl>
- [12] NetGeo, Verifia Inc, <http://www.netgeo.com>

