VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

LOKACE STANICE V SÍTI INTERNET POMOCÍ ANALÝZY DOMÉNOVÝCH NÁZVŮ

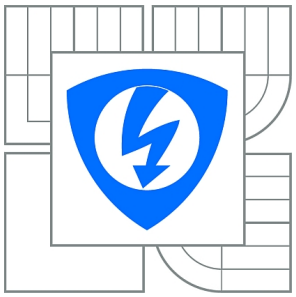
BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE ONDŘEJ JELÍNEK
AUTHOR

BRNO 2013

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

DEPARTMENT OF TELECOMMUNICATIONS

LOKACE STANICE V SÍTI INTERNET POMOCÍ ANALÝZY DOMÉNOVÝCH NÁZVŮ

GEOLOCATION OF INTERNET NODES USING DOMAIN NAME ANALYSIS

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

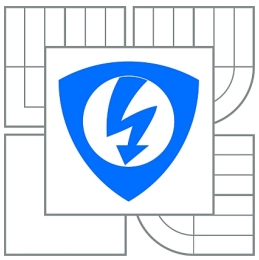
AUTOR PRÁCE — ONDŘEJ JELÍNEK

AUTHOR

VEDOUCÍ PRÁCE — doc. Ing. DAN KOMOSNÝ, Ph.D.

SUPERVISOR

BRNO 2013



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky a komunikačních technologií

Ústav telekomunikací

Bakalářská práce

bakalářský studijní obor
Teleinformatika

Student:	Ondřej Jelínek	***ID:***	136530
Ročník:	3	***Akademický rok:***	2012/2013

NÁZEV TÉMATU:

Lokace stanice v síti Internet pomocí analýzy doménových názvů

POKYNY PRO VYPRACOVÁNÍ:

Seznamte se s principy vyhodnocování fyzické polohy stanic v síti Internet. Nastudujte systém doménových jmén používaný v Internetu (DNS – Domain Name System). Realizujte aplikaci, která bude zjišťovat polohu stanice pomocí analýzy doménového jména. Pomocí reálných doménových názvů ověřte správnou činnost navržené aplikace.

DOPORUČENÁ LITERATURA:

[1] PUŽMANOVÁ, R. TCP/IP v kostce. 1. vyd. České Budějovice : Kopp, 2004. 607 s. ISBN 80-7232-236-2.

[2] DOSTÁLEK L. et al.: Velký průvodce protokoly TCP/IP: Bezpečnost. 2. aktualizované vydání. Computer Press, 2003. ISBN 80-7226-849-X.

[3] NEMETH, E., SNYDER, G., HEIN T. Linux - Kompletní příručka administrátora. Computer Press, 2004. 880 s. ISBN: 80-722-6919-4.

Termín zadání:	11.2.2013	***Termín odevzdání:***	5.6.2013

Vedoucí práce: doc. Ing. Dan Komosný, Ph.D.
Konzultanti bakalářské práce:

prof. Ing. Kamil Vrba, CSc.
Předseda oborové rady

UPOZORNĚNÍ:

ABSTRAKT

Cílem této bakalářské práce je na základě získaných informací vytvořit aplikaci pro lokaci stanic v síti pomocí analýzy doménových jmen. Teoretická část se věnuje metodám určování fyzické polohy stanice v síti Internet (tzv. geolokace) a systému DNS. Nejprve je nutné vybrat vhodnou metodu pro určení polohy stanice. Dalším krokem je pochopit systém DNS a strukturu a funkci doménových názvů. Nakonec je pomocí programovacího jazyka JavaFX Script vytvořena aplikace, která je schopná analyzovat domény. Vytvořené aplikaci a získaným výsledkům se věnuje druhá část práce. Zabývá se strukturou aplikace a její funkcí. Nakonec jsou zhodnocena získaná data a přesnost aplikace je porovnána s přesností jiných metod určení polohy stanice.

KLÍČOVÁ SLOVA

Databáze, DNS, doména, geolokace, Java, Internet, IP, server

ABSTRACT

Goal of this thesis is to create an application for geolocation of hosts using an analysis of domain names. The theoretical part of the thesis focuses on estimation of geographic location of hosts in the Internet (called geolocation) and the DNS system. This knowledge is necessary for creation of the application. The first task is to choose suitable method of geolocation. The next step is to fully undestand the DNS system and function of domain names. Finally, the application for analysis of domains in JavaFX Script programming language is created. The practical part of the thesis describes this application, its structure, function and results. Acquired data are analysed and compared to results of other geolocating methods.

KEYWORDS

Database, DNS, domain name, geolocation, Java, Internet, IP, server

JELÍNEK, Ondřej *Lokace stanice v síti Internet pomocí analýzy doménových názvů*: bakalářská práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2013. 45 s. Vedoucí práce byl doc. Ing. Dan Komosný, Ph.D.

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Lokace stanice v síti Internet pomocí analýzy doménových názvů“ jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno ....

(podpis autora)

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu bakalářské práce panu doc. Ing. Danovi Komosnému, Ph.D. za odborné vedení, ochotu, inspirující konzultace a trpělivost při tvorbě této práce.

Brno ....

(podpis autora)

OBSAH

SEZNAM OBRÁZKŮ

ÚVOD

Už mnoho let se Internet postupně stává nedílnou součástí našeho života. V současnosti je pro nás téměř nemožné si představit život bez něj. Dávno již neslouží pouze k předávání informací či zábavě, funkce jako je obchod přes internet či vzdálená správa nejsou dnes ničím výjimečným. S rostoucím využitím Internetu a jeho rozšířením do všech oblastí světa se objevily i nároky na zajištění bezpečnosti.

Jedním z prvků týkajících se bezpečnosti Internetu, kterým se v první kapitole tato práce zabývá, je metoda zvaná geolokace. Jedná se o metodu, která různými postupy zjišťuje geografickou polohu konkrétní stanice v síti. V rámci Internetu se používá především IP geolokace. Tyto metody jsou založeny na práci s IP adresou, která je přidělena každé stanici a pro každou stanici je unikátní. Existují metody aktivní a pasivní, z nich jsou v současnosti nejpoužívanější ty pasivní. Jsou založené na velké databázi, kde je pro každou IP adresu či adresní rozsah uložen záznam obsahující geografické údaje, z nichž lze určit přibližnou polohu stanice. Pro běžného uživatele stačí připojit se k Internetu, najít stránku, která je propojená s danou databází a zadat hledanou IP adresu. Pokud se adresa nachází v databázi, stránka nám vypíše odpovídající záznam.

V Internetu se kromě IP adres používají pro adresaci i takzvané domény, které jsou stěžejním tématem této práce. Jedná se o překlad konkrétní IP adresy na textový řetězec, který je použit jako název dané stránky. Nejznámějším příkladem u nás je google.com, seznam.cz a mnoho dalších. Tento způsob orientace v síti byl zaveden především pro pohodlí uživatelů. Domény jsou součástí systému DNS, pod který spadá vše, co se týká doménových jmen. Systému DNS a všemu, co k němu patří, se věnuje druhá kapitola práce. Pod DNS patří servery, které provádí překlad IP adres na domény a obráceně, software nutný pro provoz těchto serverů i struktura a vlastnosti samotných doménových názvů. V dnešní době platí, že každý, kdo chce být na Internetu úspěšný, musí mít i dobře zvolenou doménu, nejlépe nějakou jednoduchou či atraktivní, například pro české letecké společnosti mající zájem o prodej letenek přes síť, je doména letenky.cz naprosto ideální. V souvislosti s touhou po co nejlepší doméně se objevil obchod s doménovými jmény. Není totiž možné vybrat si a zaregistrovat jakoukoliv doménu. Tomuto tématu se věnuje poslední část druhé kapitoly.

Třetí kapitola se zabývá výhradně praktickou částí bakalářské práce. Je zde podrobně popsána vytvořená aplikace pro lokaci stanic v síti pomocí analýzy domén. Jednotlivé podkapitoly se zabývají hlavními funkčními bloky aplikace, jejich strukturou a účelem. V závěrečné podkapitole praktické části jsou detailně rozebírány výsledky vytvořené aplikace. Vytvořený program jednoznačně splnil zadání a to s velmi dobrou přesností určení polohy stanice, která mohla být ovlivněna někte-

rými, v podkapitole diskutovanými, činiteli. Tato práce nebyla jediná, která se zabývala využitím geolokace. Díky tomu bylo možné porovnat metodu lokace stanic pomocí analýzy domén s jinými metodami. Mezi tyto metody patří:

- pasivní metoda Whois - Jan Henek,
- pasivní metoda MaxMind GeoIP - Bc. Josef Pokorný,
- metoda Constraint-Based Geolocation (CBG) - Bc. Michael Horák.

Výsledky těchto metod a jejich vzájemné srovnání je uvedeno v závěrečné podkapitole této práce.

1 LOKACE STANICE V SÍTI INTERNET

Pro zjištění geografické polohy konkrétního objektu (počítač, mobilní telefon, notebook . . .) se používají různé metody v závislosti na tom, jakým způsobem a jaký typ dat používají. Obecně se metoda určení pozice objektu nazývá geolokace. Například u mobilního telefonu vybaveného GPS se využívají data z GPS modulu, kdežto v síti Internet se používají IP adresy, případně domény. Lokací objektu (stanice) pomocí IP adresy a hlavně pomocí domény se tato práce zabývá.

1.1 Lokace pomocí IP

IP geolokace je metoda používaná v síti Internet pro určení geografické polohy stanice. K tomu využívá protokol IP a především IP adresu námi hledaného objektu. IP geolokaci můžeme rozdělit na pasivní metody a aktivní metody. Navzájem se liší nejen způsobem zjišťování polohy cíle, ale také náročností na provoz a údržbu či přesností získaného výsledku. IP geolokace má mnohá využití, jako je ochrana proti podvodům při platbách kreditní kartou, cílená online reklama nebo lepší rozložení datové zátěže v síti.

1.1.1 Pasivní metody

Tyto metody používají k určení pozice objektu jeho IP adresu. IP adresy a jejich rozsahy jsou přidělovány organizací IANA (Internet Assigned Numbers Authority), přičemž každá adresa je unikátní. Adresy jsou přidělovány hierarchicky – koncový uživatel dostane svou IP adresu přidělenou od providera, který ji získá od lokálního či národního registru (LIR – Local Internet Registry, NIR – National Internet Registry), které spadají pod registr celého regionu (RIR – Regional Internet Registry). Regionálních registrů, které spravují adresy, je celkem 5. Je to AfriNIC pro Africký region, APNIC pro Asijsko-Pacifický region, ARIN pro region Severní Ameriky, LACNIC pro region latinské Ameriky a některé Karibské ostrovy a RIPE NCC pro Evropu, Střední Východ centrální Asii. Údaje o stanicích s danou IP adresou jsou shromažďovány různými databázemi. Jejich nevýhodou je, že správnost údajů v nich obsažených závisí na častých manuálních aktualizacích. Některé z nich jsou veřejné, jiné soukromé, odlišují se od sebe taky strukturou uložených záznamů, přesností samotných záznamů nebo množstvím IP adres a domén, které obsahují. Dvě nejznámější veřejné databáze jsou Whois a GeoIP [23].

Whois

Úplně nejznámější a nejpoužívanější je v současnosti databáze Whois. V této databázi lze hledat jak konkrétní IP adresu, tak i doménu. V některých případech je výhodnější použít přímo IP adresu, jindy je výhodnější použít doménu. Přístup k databázi je velmi jednoduchý, stačí si zobrazit stránku s připojením k Whois databázi, často stránky mají řetězec Whois přímo ve svém názvu, a pak už jen zadat hledanou adresu/doménu do okna vyhledávače. Ten zjistí, jaké údaje k našemu zadání obsahuje a tento záznam vypíše. Značnou nevýhodou databáze Whois je skutečnost, že mnoho záznamů má odlišný formát nebo neobsahují všechny údaje, které bychom potřebovali. Pokud by cílem této práce bylo vytvořit program pracující se záznamy z Whois, bylo by to kvůli nejednotnému formátu záznamů velice obtížné [21].

GeoIP

Druhou velmi používanou databází je MaxMind GeoIP. Jedná se o komerční organizaci, tudíž pokud chceme využít jejich databázi, musíme si buď zakoupit licenci a postupně si dokupovat případné aktualizace, nebo za menší poplatek si koupíme určitý počet dotazů, na které nám databáze odpoví. Mezi tyto placené databáze patří GeoIP City, GeoIP Country a další. Pro ty, kteří nechtějí platit, je zde bezplatná databáze GeoLite City and Country, která ovšem obsahuje podstatně méně informací a ani jejich přesnost není taková, jako u placených databází. I přesto se ale databáze GeoLite často využívá, protože pro běžné potřeby je její přesnost a rozsah údajů v záznamu dostačující. Na stránce [11] je dostupné demo jinak placených databází GeoIP City/ISP/Organization. Pro moji IP adresu je výstup databáze uveden na obr. 1.1.

Your IP address is 46.33.96.9.

GeoIP City/ISP/Organization Results

IP Address	Country Code	Location	Postal Code	Coordinates	ISP	Organization	Domain	Metro Code
46.33.96.9	CZ	Vyskov, Jihomoravsky kraj, Czech Republic		49.2783, 17.0026	INFOS Art, s.r.o.	servers, routers	infos.cz	

Obr. 1.1: Výstup GeoIP pro IP adresu `46.33.96.9`

1.1.2 Aktivní metody

Aktivní metody geolokace jsou založeny na měření datového toku mezi stanicí se známou polohou (referenční bod) a lokalizovanou stanicí. Informace zjištěné tímto měřením se dále zpracují a z nich získáme přibližnou polohu hledané stanice. Nejčastěji metody měří zpoždění a zjišťují cestu sítí od referenčního bodu k hledané stanici. Zpoždění můžeme definovat jako dobu nutnou pro přenos jednoto datového segmentu od zdroje k cíli. Tato doba může být ovlivněna různými faktory, mezi které patří vlastnosti používaného vedení, především přenosová rychlost, momentální zatížení, ale také vlastnosti přechodových uzlů jako jsou směrovače a přepínače, jejichž rychlost přeposílání paketů i obsah vyrovnávací paměti jsou omezené. Neméně významnou veličinou je i geografická vzdálenost zdroje vysílání od cíle. Základními nástroji pro měření námi požadovaných veličin jsou `Ping` a `Traceroute`. Aktivními metodami se tato práce dále nezabývá, proto se o nich již dále nebudu rozepisovat [23].

1.2 Lokace pomocí domén

Domény a jejich využití pro geolokaci jsou stěžejním tématem této práce. Samotný doménový systém (neboli DNS - Domain Name System) si podrobně rozebereme ve druhé kapitole. Základním nástrojem pro lokaci stanice pomocí domén je aplikace `nslookup`. Tuto aplikaci můžeme jednoduše spustit v příkazovém řádku. Jak je vidět na obrázku 1.2 , `nslookup` nám z IP adresy pomocí reverzního překladu vytvoří doménu dané stanice, ze které můžeme určit například ve které zemi se daná stanice nachází. Domény ovšem můžeme zadávat i místo IP adres do vyhledávacích databází typu Whois nebo GeoIP. V některých případech je to dokonce výhodnější, než zadávat přímo IP adresu [14].

```
C:\Users\Ondra>nslookup 77.75.72.7
Server: my.router
Address: 192.168.1.1

Nazev: novinky.cz
Address: 77.75.72.7
```

Obr. 1.2: Výpis nástroje `nslookup`

1.3 Využití geolokace v praxi

Jak již bylo zmíněno dříve, využití IP geolokace je poměrně široké. Nejvyužívanější funkcí je takzvané cílení reklamy a informací. Pro reklamu je geografická poloha cílové skupiny poměrně důležitá, bylo by plýtváním například nabízet uživatelům z Evropy zboží dostupné pouze na australském trhu a naopak. Využití má i nabízení regionálních či národních výrobků cílené na zvýšení prodeje a obecné povědomosti o daném produktu právě ve státě či regionu, kde se onen produkt vyrábí. Z informací je nejčastěji cílená předpověď počasí a aktuální informace o dění v městě či regionu uživatele.

Dalším důležitým využitím je ochrana proti podvodům. Může se stát, že při platbě přes Internet dojde k narušení bezpečnosti přenášených dat a útočník tak získá například číslo i PIN naší platební karty. Pokud by byl z jiného státu či dokonce z jiného kontinentu a chtěl použít ukradené údaje k výběru peněz, geolokační metody zjistí, že požadavek přišel z lokace naprosto neodpovídající předchozím požadavkům či bydlišti majitele karty a tento podvodný požadavek odmítne [23].

2 DNS

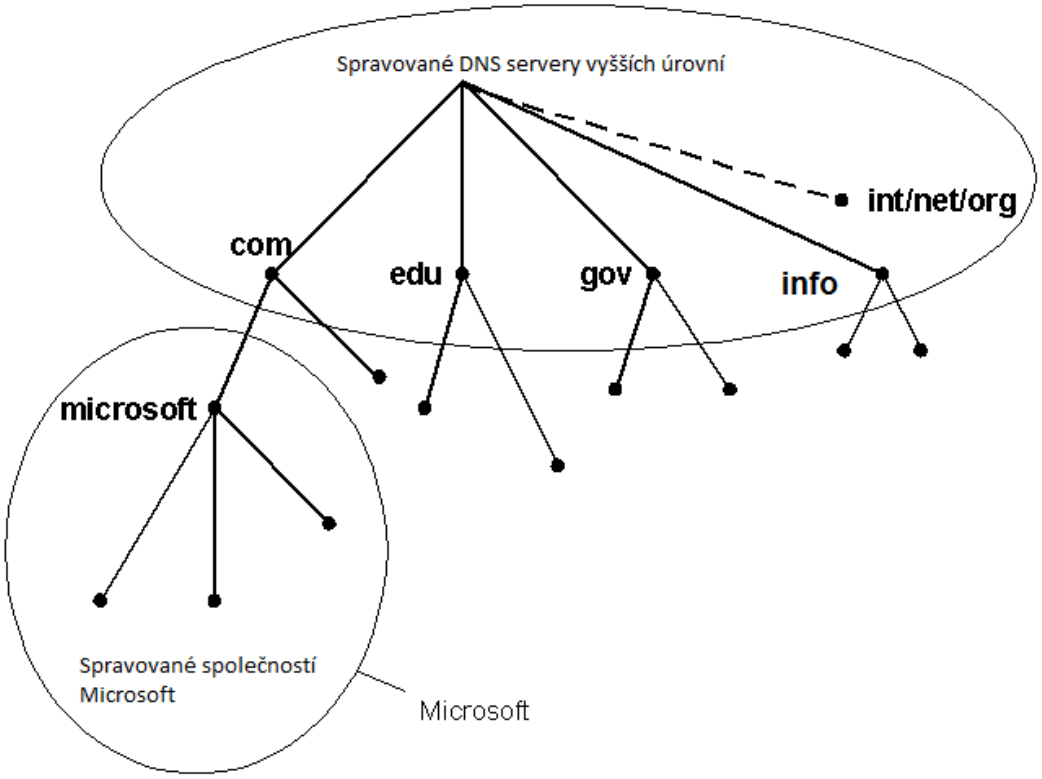
DNS neboli Domain Name System je hierarchická struktura doménových jmen, která podobně jako systém IP adres slouží k označení konkrétního místa v síti Internet. Potřeba zavedení tohoto systému vyplývá z odlišného způsobu identifikace stanic v síti. Počítače se navzájem identifikují pomocí číselných označení – IP adres. Naopak pro uživatele sítě je přirozenější identifikovat zařízení pomocí domén, protože pro většinu lidí je snadnější zapamatovat si textový řetězec, například `www.novinky.cz`, než IP adresu jako sled sedmi až dvanácti čísel (například `77.75.72.7`). V posledních letech používání vhodných domén nabírá na důležitosti. Jedná se především o prezentaci firem a jiných subjektů na Internetu, kdy krátká nebo dobře zapamatovatelná doména značně přispívá ke známosti firmy na síti a k navštěvovanosti jejich stránek [22].

2.1 Doménová jména

Jak již bylo zmíněno, doménová jména slouží, podobně jako IP adresy, k jednoznačné identifikaci konkrétního místa v síti. Na rozdíl od IP adresy zůstává doména stejná, i pokud se uložiště dat, k nimž se váže, například přesune z jednoho města do druhého. V tomto případě dojde ke změně IP adresy stanice či serveru, který doména označuje, ovšem doména samotná zůstane zachována. Doménová jména slouží především jako popis funkce dané stránky. Dobrým příkladem je již použitá doména `www.novinky.cz`, kdy hned z tvaru domény lze očekávat, že se jedná o stránku s aktuálními informacemi o dění v našem okolí. Stejné je to například u stránek `www.vutbr.cz` či `www.brno.cz`, kdy od první budeme očekávat informace o VUT v Brně a od druhé oficiální informace o městě Brně.

Všechna doménová jména v síti tvoří jednu společnou stromovou strukturu. Kořenem tohoto stromu je speciální doména, takzvaná doména nultého řádu, která se označuje tečkou a běžně se v doméně neuvádí. Název každého uzlu v této struktuře musí být unikátní. Doménové jméno je pak tvořeno cestou od kořenu k danému uzlu ve stromě, kde přechod mezi uzlem a jeho následovníkem na další úrovni je označen tečkou. Celé doménové jméno je pak tvořeno několika částmi oddělenými od sebe tečkou, např. `www.google.com`. Jednotlivé části mohou mít až 63 znaků a celková délka doménového jména může být až 255 znaků. Lze použít pouze znaky anglické abecedy, číslice a pomlčku, která ovšem nesmí být na začátku ani konci jednotlivých domén. Nejvíce vpravo se nachází doména nejvyšší úrovně, neboli TLD (Top Level Domain), která je nejobecnější. Postupem vlevo se domény více konkretizují, za tečkou se nachází doména druhého stupně a za další tečkou doména třetího stupně,

někdy taky nazývaná subdoména. Na obr. 2.1 můžeme vidět, jak zhruba vypadá struktura doménových jmen [22].



Obr. 2.1: Struktura DNS

Nyní se podíváme na jednotlivé úrovně domén trochu podrobněji.

Doména nejvyššího řádu

Doména nejvyššího řádu (TLD) je doména na nejvyšší úrovni DNS stromu, hned pod kořenem. V doménovém jméně se nachází až na konci, např. u `www.youtube.com` je doménou nejvyššího řádu doména `.com`. TLD jsou pevně stanoveny organizací IANA. Můžeme je rozdělit na tři druhy, jsou to národní TLD, které sdružují domény konkrétního státu a většinou odpovídají kódu země podle ISO 3166-1 [9], např. `.ru` pro Rusko, generické TLD, které sdružují obecné domény, např. `.com` pro komerční organizace, většinou nejsou spojené s jedním konkrétním státem, a poslední skupinou jsou infrastrukturní TLD.

Národní TLD je dvojpísmenná a označuje příslušnost doménového jména k danému státu. Státní domény jsou spravovány spravovány organizací v konkrétním státu, například v České Republice spravuje registr domény `.cz` organizace Cz.NIC. Tyto organizace, spadající pod daný stát, spravují registraci domén obsahujících

TLD daného státu a mají právo požadavek na registraci odmítnout. Využití národní TLD je ve většině případů zpoplatněno.

Generická TLD je doména společná pro určitou skupinu subjektů. Je jí mnoho druhů, vždy ale musí být nejméně třípísmenná. Použití některých z nich, jako jsou např. `.com, .org` či `.net`, není nijak omezeno. Jsou ovšem i domény vymezené pouze pro daný účel, např. `.xxx`. Některé z vymezených domén jsou dokonce garantované – nejen, že jsou vymezené pro daný účel, navíc jsou spravované organizací, která stanoví pravidla pro jejich používání a pak dohlíží na jejich provoz. Mezi tyto domény patří například domény `.museum` a `.jobs`.

Poslední skupinou jsou infrastrukturní TLD. Tyto domény jsou používány výhradně pro vnitřní potřeby internetové struktury a jsou neveřejné. V současnosti je používána pouze doména `.arpa`, kterou využívá doménový systém. Je spravována přímo organizací IANA. Původně měla sloužit pouze pro zavedení DNS struktury, kdy nově vytvořené domény byly nejprve přiřazeny pod doménu `.arpa` a až poté postupně přeřazeny do odpovídajících domén. Kromě tohoto dočasného použití se ale tato doména používá i pro zpětný překlad IP adres na doménové jméno, a proto byla zachována.

Doména druhého řádu

Doména druhé úrovně je nejdůležitější částí z hlediska orientace. Označuje, k jakému účelu stránka slouží, např. u domény `www.csob.cz` můžeme pouze z domény druhého řádu odhadnout, co se na dané stránce nachází. Také při zjišťování informací pomocí internetových vyhledávačů se často orientujeme podle domén druhého řádu, ostatní domény téměř nepoužíváme.

Doména třetího řádu

Doména třetího řádu neboli subdoména slouží k rozlišení jednotlivých webových stránek v rámci jedné domény. Pro hlavní stránku je většinou použita subdoména `www`, pro internetové obchody se často používá subdoména e-shop. Používají je také například univerzity pro odlišení stránek jednotlivých fakult nebo pro přidělení vlastní stránky různým oddělením, např. IT správa. Dobrým příkladem je samotné VUT, kdy například pro naši fakultu má doména tvar `www.feec.vutbr.cz` a pro fakultu architektonickou je to `www.fa.vutbr.cz` [1].

2.2 Funkce DNS

V úvodu této kapitoly najdeme, že DNS je stromovou strukturou obsahující doménová jména. Hlavní funkcí DNS je správa těchto domén, jejich přidělování, ukládání

do databází a řešení dotazů uživatelů. Nejčastěji provádí překlad domén na IP adresy a naopak. Ke svému fungování využívá hierarchickou strukturu serverů, kterou si podrobněji popíšeme níže. Systém DNS je provozován pomocí protokolu DNS, který používá TCP/UDP port 53 a je definován v RFC1035 [16]. Stromová struktura systému je administrativně rozdělená do zón, o které se starají jednotliví správci, přičemž daná zóna obsahuje autoritativní informace o spravovaných doménách. Při běžných situacích se dotaz i odpověď zasílá ve formě jednoho UDP paketu. UDP je jednoduchý a má minimální režii, není ovšem nijak zabezpečen proti poškození či ztrátě paketu. Pokud se celá odpověď nevleze do jednoho paketu, je vhodnější použít TCP, u nějž lze poslat více paketů odpovědi najednou.

2.3 DNS servery

Podobně jako systém DNS i jeho servery tvoří hierarchickou strukturu. Na vrcholu této struktury jsou takzvané kořenové servery (root servers). Tyto servery představují jednu z nejdůležitějších částí infrastruktury Internetu, na níž závisí spolehlivost, správnost a bezpečnost operací na Internetu. Poskytují ostatním DNS serverům kořenový zónový soubor (root zone file), který popisuje, kde se nachází autoritativní servery pro domény nejvyšší úrovně. Správu tohoto souboru zajišťuje společnost IANA. Kořenových serverů je dohromady 13, přičemž 10 z nich se nachází na území USA. Zbývající tři jsou v Londýně, Stockholmu a Tokiu. Jsou označovány písmeny A až M, která jsou i předponou jejich oficiálních názvů `root-servers.net`. Server A (`a.root-servers.net`) obsahuje kořenový zónový soubor a hlavní databázi domén nejvyšší úrovně. Zbylých 12 serverů tyto údaje přebírá právě z A serveru. Důležitosti těchto serverů odpovídá také jejich zabezpečení. Za bezproblémový provoz a správu root serverů odpovídá společnost ICANN (Internet Corporation for Assigned Names and Numbers). Seznam root serverů a organizací, které je spravují naleznete na obr. 2.2.

Ostatní servery ve struktuře obsahují vždy jen část doménové databáze, která odpovídá zóně, kterou spravují. Obsahují také údaje, na kterých severech se nachází ostatní databáze. Díky tomu je zátěž rozložena mezi mnoho serverů a značně se tak snižuje i rozsáhlost databáze, kterou daný server musí obsahovat.

Vůči doméně/zóně, kterou DNS server spravuje, může zastávat 3 různé role.

Na primárním serveru vznikají data týkající se domény. Každá doména má právě jeden primární server. Pokud je třeba udělat v doméně změnu, je nutné ji provést na primárním serveru. Primárním DNS serverem pro kořenovou doménu je již zmíněný `a.root-servers.net` a pro doménu `.cz` to je server `ns.tld.cz`.

Sekundární server je v podstatě kopií primárního. Průběžně si stahuje aktua-

Seznam root serverů

Označení serveru	IP adresa	Správce
a.root-servers.net	198.41.0.4, 2001:503:ba3e::2:30	VeriSign, Inc.
b.root-servers.net	192.228.79.201	University of Southern California (ISI)
c.root-servers.net	192.33.4.12	Cogent Communications
d.root-servers.net	128.8.10.90, 2001:500:2d::d	University of Maryland
e.root-servers.net	192.203.230.10	NASA (Ames Research Center)
f.root-servers.net	192.5.5.241, 2001:500:2f::f	Internet Systems Consortium, Inc.
g.root-servers.net	192.112.36.4	US Department of Defence (NIC)
h.root-servers.net	128.63.2.53, 2001:500:1::803f:235	US Army (Research Lab)
i.root-servers.net	192.36.148.17, 2001:7fe::53	Netnod
j.root-servers.net	192.58.128.30, 2001:503:c27::2:30	VeriSign, Inc.
k.root-servers.net	193.0.14.129, 2001:7fd::1	RIPE NCC
l.root-servers.net	199.7.83.42, 2001:500:3::42	ICANN
m.root-servers.net	202.12.27.33, 2001:dc3::35	WIDE Project

Obr. 2.2: Seznam root serverů a jejich správců

lizovaná data z primárního serveru a slouží jako jeho záloha i k rozložení zátěže a zabránění přetížení primárního serveru, Sekundárních serverů může být několik, pro každou doménu ale musí vždy být alespoň jeden. Pro přenos dat z primárního serveru na sekundární se používá tzv. zónový transfer (AXFR), který je určen k přenosu všech DNS záznamů dané zóny/domény. Nevýhodou je, že se přenáší celý obsah zóny, i když se změní jen jeden záznam. Pro vyřešení tohoto problému se používá mechanismus IXFR, který přenáší pouze změněné záznamy.

Třetím typem serverů jsou pomocné, neboli cachovací DNS servery. Na rozdíl od primárních a sekundárních, které jsou hlavními nositeli informací o doménách a poskytují autoritativní odpověď, poskytují tyto servery odpověď neautoritativní. Slouží jako vyrovnávací paměť pro snížení zátěže celého systému. Při práci s doménovými jmény by uživatel musel pro překlad tohoto jména obeslat s dotazem množství DNS serverů a to pro každou část doménového jména. To by pochopitelně vedlo k nadměrnému a zbytečnému zatížení celého systému. Proto existují cachovací servery, které uživateli zprostředkují celý systém překladu domény na IP adresu a získaná data si ukládají do paměti. Kromě výsledků si ukládají i cesty k autoritativním serverů. Díky tomu se při opakování stejných dotazů nemusí prohledávat celý systém, ale pomocný server pouze vypíše údaj ze své paměti. Aby se zabránilo neaktuálnosti dat v paměti caching serveru, mají informace v něm uložené přidělenou hodnotu Time To Live (TTL), která udává dobu, po kterou může daný záznam zůstat v paměti serveru.

Již zmíněná autoritativní [24] odpověď je poskytovaná servery, které nesou informace potřebné k práci s danou doménou. Jsou to servery primární a sekundární.

Slovo „autoritativní“ zde říká, že jde o servery, které obsahují závazné informace o dané doméně, kterými by se měla řídit všechna zařízení v síti, aby tak došla při překladu domény na IP adresu ke správném výsledku [15] [7].

2.4 Práce serverů

Hlavní náplní práce DNS serverů je řešení dotazů, které zadávají uživatelé. Tyto dotazy se týkají překladu domén na IP adresu a naopak (tzv. reverzní překlad). Umožňují tak používat k identifikaci vzdálených stanic doménová jména místo IP adres, což je pro uživatele mnohem pohodlnější.

Každá koncová stanice má ve své síťové konfiguraci zadanou i adresu lokálního DNS serveru, na který se obrací s dotazy. Tato adresa je většinou přidělena pomocí DHCP (Dynamic Host Configuration Protocol). Pokud počítač hledá určitou informaci v DNS (např. IP adresu k danému jménu), pošle dotaz právě na svůj lokální DNS server. Ten,stejně jako ostatní servery, obsahuje soubor s adresami serverů pro domény vyšší úrovně. Pokud tedy sám nezná odpověď na dotaz, obrátí se s dotazem na server vyšší úrovně.

Dotazy mohou být řešeny buďto rekurzivně nebo nerekurzivně. Rekurzivní dotaz znamená, že pokud server, kterého se dotazujeme, nezná odpověď, obrátí se na servery vyšší úrovně, sám nalezne odpověď na dotaz, uloží si ji do paměti a pošle zpět koncové stanici. Tento typ dotazu provádí cachovací servery. Naopak nerekurzivní dotazy provádí především primární servery a to z důvodu zabránění přetížení systému. V případě nerekurzivního dotazu, pokud nezná odpověď, pošle pouze seznam serverů, na které se máme obrátit, jinými slovy, pokud neobsahuje údaj odpovídající našemu dotazu, náš dotaz ho nezajímá.

Základním úkolem serverů je poskytnou informace (hlavně IP adresu) k námi zadanému doménovému jménu. Existuje ale i tzv. reverzní dotaz, který dokáže sdělit jméno, pod kterým je daná IP adresa zaregistrovaná. U tohoto typu dotazu je ovšem problém s opačným uspořádáním IP adresy a doménového jména. IP adresa má vlevo nejobecnější informace, které se postupem doprava konkretizují (zleva adresa sítě, podsítě a konkrétní stanice), naproti tomu doména má nejobecnější informace vpravo a postupem doleva se konkretizují. Tento problém řeší DNS tím, že při reverzním dotazu obrátí pořadí bajtů IP adresy. K obrácené adrese pak připojí doménu `in-addr.arpa` a výslednou doménu pak hledá standardním způsobem. Pokud bychom hledali doménu k IP adrese `142.103.76.4`, vytvoří se dotaz `4.76.103.142.in-addr.arpa`. Obrácení adresy umožňuje přeposílat dotaz správcům odpovídajících podsítí a sítí.

Data z reverzních dotazů ovšem nejsou úplně spolehlivá. Do reverzní domény se

dají napsat v podstatě libovolná jména. Správce určité sítě tedy může v reverzní zóně prohlásit o dané stanici, že se jedná například o `www.csfd.cz`. Je tedy dobré si získanou informaci ověřit standardním dotazem. Když odpovědí na něj bude námi původně vložená IP adresa, jsou data z reverzního dotazu důvěryhodná [22].

2.5 Software DNS

Aby mohly servery vůbec pracovat, musí na nich být zprovozněn příslušný software, umožňující zpracování příchozích dotazů a spravující vlastní DNS záznamy. Existuje mnoho druhů programů pro DNS servery, většina z nich je volně dostupných, některé jsou placené či dodávané společně se zakoupením jiného typu softwaru (např. Microsoft DNS). Jednotlivé programy se od sebe mohou značně lišit. Některé podporují rekurzivní dotazovaní, jiné zase mohou fungovat jako pomocné (cachovací) servery. Různá je také podpora platforem, na kterých mohou fungovat, nebo přístupnost ke kódu samotného programu (tzv. open source programy). Mezi nejvýznamnější programy patří volně dostupný BIND, Djbdns, NSD, PowerDNS, MaraDNS či Unbound a komerční MicrosoftDNS a Simple DNS Plus [19].

- BIND je jednoznačně nejpoužívanějším softwarem pro DNS servery. Nabízí velmi robustní a stabilní platformu, která je vhodná pro vysoké zatížení sítě. Kromě odolnosti vůči vysokému zatížení se vyznačuje velkou spolehlivostí. Velmi často se používá na Unixových a Linuxových serverech. Jedná se o open source software, který lze stáhnout na stránkách společnosti ISC [8].
- Djbdns není na rozdíl od ostatních jediným programem, ale celým balíčkem několika programů, z nichž každý plní specifickou funkci v rámci celého DNS serveru. Jeho nepopiratelnou výhodou oproti BINDu je vysoká míra bezpečnosti, díky které je taky v určitých prostředích upřednostňován před BINDem, u kterého se čas od času odhalí různá bezpečnostní rizika. Kromě bezpečnosti je ceněná také jeho spolehlivost a dobrý výkon, nevýhodou je ovšem uživatelsky nepříliš přátelské prostředí, které je zaměřeno spíše na přehlednou a efektivní komunikaci se softwarem, než s uživatelem. Nevýhodou může být i to, že kód programu není volně dostupný.
- NSD je software určený především pro autoritativní servery, které neprovádějí rekurzivní vyhledávání. Jeho velkou výhodou je rychlost startu a zprovoznění k plné funkčnosti, která u jiných softwarů na vytížených serverech může zabrat i hodinu. NSD je naproti tomu zprovozněno během několika minut.
- PowerDNS podporuje rekurzivní dotazování i DNSSEC. Jeho kód je plně dostupný. Největší výhodou tohoto softwaru je skutečnost, že všechny své zá-

znamy má uložené v databázi, kterou lze na rozdíl od zónového souboru snadno spravovat pomocí různých administrátorských prostředí.

- MaraDNS je open source softwarový balíček, který funguje na operačním systému Windows i Unix. Jeho hlavními přednostmi je jednoduchost, snadné ovládání a bezpečnost.
- Unbound je významným konkurentem BINDu. Je určen jak pro vysoké zatížení, tak pro méně vytížená síťová zařízení v modifikované variantě. Jeho výhodou je snadná konfigurace, velký výkon, podpora DNSSEC, bezpečnost a velmi dobrý management.
- MicrosoftDNS je dodáván společně s operačními systémy Windows, není tedy volně dostupný. Používá se především u operačního systému Windows NT, který je určen právě pro servery. Díky tomu se stal jedním z nejpoužívanějších softwarů pro DNS. Dokáže velmi dobře spolupracovat s BINDem a svoje záznamy ukládá jak do zónového souboru, tak do snadno spravovatelné databáze.
- Simple DNS Plus je dalším softwarem určeným pro systém Windows. Jedná se o komerční software, jehož cena se může lišit v závislosti na počtu zón potřebných pro daný server. Je uživatelsky přátelský a jeho zónové soubory lze spravovat jak z příkazové řádky, tak i pomocí jednoduchých webových nástrojů.

2.6 DNS záznamy

Záznamy uložené na DNS serverech, které se týkají jednotlivých domén, mohou obsahovat mnoho různých informací o těchto doménách. Základní a nejvyhledávanější informaci je IP adresa patřící k dané doméně. Kromě informací o konkrétní doméně mohou záznamy obsahovat informace o jejích subdoménách, pokud tyto nemají samostatný záznam. Mohou také odkazovat na servery, které obsahují informace o těchto subdoménách.

Všechny tyto záznamy jsou uložené v zónovém souboru. Zónové soubory obsahují všechny doménové záznamy, jejichž formát a obsažené informace jsou přesně definovány (RFC1035 [16]). Na většině serverů jsou uloženy ve formě textového souboru, ve kterém každý řádek odpovídá jednomu DNS záznamu. Změna doménových informací probíhá jednoduchou změnou textu tohoto souboru v libovolném textovém editoru. Po uložení se zavolá příkaz, přes který si server načte aktualizované hodnoty a začne je poskytovat ostatním serverům.

Formát záznamu

Každý záznam (řádek textového souboru) obsahuje 5 položek [22].

- První je doménové jméno, pro které záznam vytváříme. Většinou patří do aktuálně definované domény, píše se tedy bez tečky a bude k němu doplněna aktuální doména. Pokud doplníme tečku, bere se jméno jako kompletní a nic se již doplňovat nebude.
- Druhou položkou je doba životnosti (TTL) udávaná v sekundách. Často se neuvádí a je tak nastavena implicitní hodnota. Dříve se používala hodnota 86400 sekund (to je 24 hodin), nyní se často používá pouhých 300 sekund (čili 5 minut).
- Třetí je třída, která udává rodinu protokolů, k níž se záznam vztahuje. Existují různé typy, ale používá se pouze IN pro Internet.
- Čtvrtý je typ záznamu, kterému se budu věnovat níže.
- Poslední je hodnota, vztahující se k záznamu a poskytující mu potřebné údaje. Obsahem hodnoty bývají často doménová jména.

Typy záznamů

Nejčastěji používané jsou záznamy typu A, AAAA, CNAME, MX, NS, PTR a SOA.

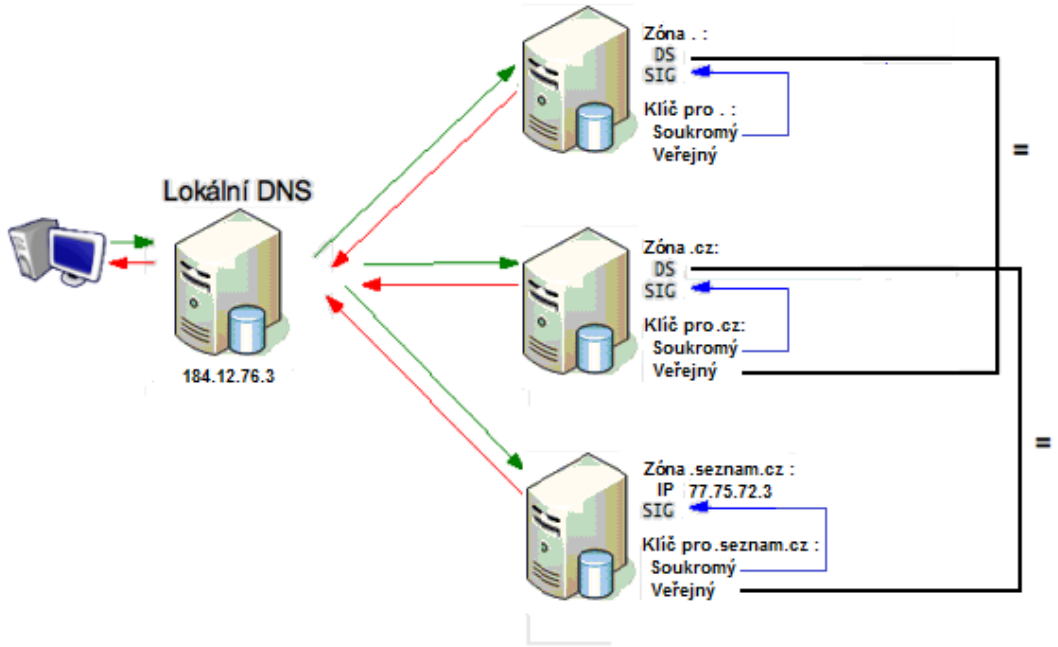
- Záznam typu A (adress record) obsahuje IP adresu, která odpovídá dané doméně.
- Záznam AAAA (IPv6 adress record) obsahuje IP adresu typu IPv6 náležící dané doméně.
- Záznam CNAME (canonical name record) definuje tzv. alias – jiné jméno pro již zavedenou doménu.
- Záznam MX (mail exchange record) udává adresu a prioritu serveru pro příjem elektronické pošty pro danou doménu. V případě více serverů platí, že čím nižší číslo, tím vyšší priorita.
- Záznam NS (name server record) obsahuje jméno autoritativního DNS serveru pro danou doménu.
- Záznam LOC (Location record) obsahuje údaje o geografické poloze stanice s konkrétním doménovým jménem.
- Záznam PTR (pointer record) je typ záznamu určený speciálně pro reverzní zóny. Obsahuje jméno počítače (získané reverzním dotazem) přidělené konkrétní síťové adrese.
- Záznam SOA (start of authority record) je speciální záznam, který zahajuje každý zónový soubor. Vyskytuje se v souboru pouze jednou a obsahuje několik specifických údajů: MNAME, RNAME, Serial, Refresh, Retry, Expire, TTL.
  - MNAME je název primárního DNS serveru pro danou zónu.

– RNAME je kontakt na správce zónového souboru, uvádí se e-mailová adresa, ve které je zavináč nahrazen tečkou.
– Serial je sériové číslo označující verzi souboru, které je nutné navýšit s každou změnou v souboru. Ostatní servery tak poznají, že došlo ke změně a stáhnou si aktuální soubor.
– Refresh udává dobu, po jejíž uplynutí se sekundární server pravidelně dotazuje primárního na aktuální verzi zónového souboru.
– Retry udává interval, po jehož uplynutí má sekundární server opakovat své dotazy, pokud se mu nepodařilo spojit s primárním.
– Expire udává dobu od posledního úspěšného připojení k primárnímu serveru, po jejímž uplynutí sekundární server označí své záznamy za neaktuální, pokud se mu do té doby nepodaří znovu připojit k primárnímu serveru.
– TTL udává dobu, po kterou záznam zůstane uložen v souboru [22].

2.7 DNSSEC

DNSSEC je rozšíření DNS, které zvyšuje jeho bezpečnost. Zajišťuje důvěryhodnost údajů získaných z DNS tím, že uživateli umožňují ověřit pravost a integritu záznamů získaných z DNS. DNS slouží k překladu domén na IP adresy (a zpět), ale není nijak chráněn proti napadení. Pokud do internetového prohlížeče zadáme doménu `www.seznam.cz`, může být přeložena na podvrženou IP adresu, ovšem v adresním řádku zůstane `www.seznam.cz`, takže uživatel nepostřehne podvod. Tímto způsobem lze získat citlivé osobní údaje, například adresu a heslo k e-mailu nebo dokonce k bankovnímu účtu.

DNSSEC používá asymetrické šifrování – jeden klíč pro zašifrování a druhý klíč na dešifrování. Držitel domény používající DNSSEC si vygeneruje dvojici soukromého a veřejného klíče. Svým soukromým klíčem pak elektronicky podepíše technické údaje, které o své doméně do DNS vkládá. Pomocí veřejného klíče je pak možné ověřit pravost tohoto podpisu. Aby byl tento klíč dostupný všem, publikuje jej držitel ke své doméně u nadřazené autority, kterou je pro všechny domény `.cz` registr domén `.cz`. I na úrovni registru domén jsou technická data v DNS podepsána a veřejný klíč k tomuto podpisu je opět správcem registru předán nadřazené autoritě. Vytváří se tak řetěz, který zajistí důvěryhodnost údajů, pokud není v žádném svém článku porušen, a všechny elektronické podpisy souhlasí [4]. Systém bezpečnosti DNSSEC je znázorněn na obr. 2.3.

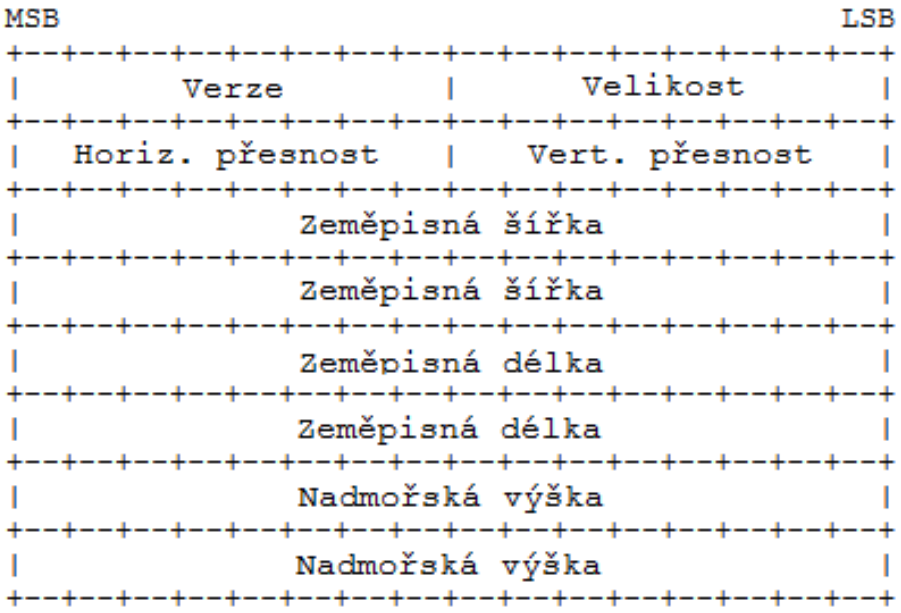


Obr. 2.3: Funkce DNSSEC

2.8 DNSLOC

Geografická poloha je již od začátku budování Internetu velmi důležitým parametrem, se kterým se při vytváření sítí musí počítat. Díky stále novým a výkonnějším technologiím tento parametr postupně ztrácí na důležitosti, přesto ale stále existují omezení, například v šířce pásma kanálů spojujících jednotlivé kontinenty, která způsobují, že ignorovat geografickou polohu stanic se vůbec nemusí vyplatit. Také vzhledem k efektivnímu využití kapacity sítí a zabránění jejich přetěžování je geografická poloha komunikujících stanic důležitou veličinou. Pro určení geografické polohy pomocí IP adres existují různé databáze, nejpoužívanější jsou již výše zmíněné WhoIS a MaxMind, které mají v databázi uložené záznamy patřící ke konkrétním IP adresám či jejich rozsahům a obsahující mnohé informace, včetně geografické polohy. Podobný záznam obsahující geografickou polohu stanice lze ale vytvořit i pro doménová jména, a právě k tomu slouží záznam LOC systému DNS.

Samotný záznam obsahuje údaje potřebné k lokaci doménového jména, tedy zeměpisnou šířku a délku společně s nadmořskou výškou. Navíc také obsahuje informaci o fyzické rozlehlosti konkrétní sítě a přesnost uvedeného údaje, obojí v metrech. Tento záznam je uložen v zónovém souboru na primárním (potažmo sekundárním) serveru dané domény [6].



Obr. 2.4: Záznam DNSLOC

2.9 Registrace domény

Podobně jako u IP adres ani u doménových jmen nelze používat jakýkoliv tvar či délku pro námi potřebovanou doménu. Omezení týkající se doménových jmen z hlediska použitých znaků jsou uvedeny v podkapitole Doménová jména. Splnění těchto podmínek je ale pouze jedním z kroků, které jsou nutné pro registraci domény. Bez registrace totiž nelze doménu efektivně používat, protože nebude zanesena do záznamů nadřazených DNS serverů a tudíž o její existenci nebude kromě nás nikdo vědět. Registrované domény jsou zaneseny do tzv. registru domén. To jest organizace, která provozuje databázi doménových jmen, spadajících pod konkrétní nadřazenou doménu. Každá doména nejvyššího řádu (TLD) je spravována samostatnou organizací, která ale spadá pod mezinárodní organizaci IANA. Tyto subjekty určují pravidla pro registraci domén spadajících pod jejich TLD a zajišťují provoz databáze těchto domén. Pro registraci domény existuje standardní postup, sestávající z několika kroků [3].

- Prvním krokem je vytvoření námi požadovaného doménového jména podle výše zmíněných podmínek. Dále si v registru domén musíme ověřit, zda tato doména již není registrovaná někým jiným.
- Následujícím krokem je výběr registrátora. Doménu nelze u správce TLD zaregistrovat, přímo, místo toho je potřeba ji zaregistrovat přes registrátora. Jedná se o firmu, které je umožněn přístup do databáze domén a která spravuje doménová jména uživatelů, kteří si je přes tuto firmu zaregistrovali. Firmy se od

sebe liší jak cenou, tak poskytovanými službami. Lze je vybírat ze seznamu registrátorů.

- Následuje registrace kontaktu. Tento kontakt obsahuje informace o osobě, která bude u dané domény uvedena v registru jako držitel domény, administrativní kontakt nebo technický kontakt. Držitel domény může být pouze jeden a může doménou jakkoliv manipulovat. Administrativní kontakt může měnit všechny údaje kromě držitele domény. Technický kontakt může měnit pouze technické parametry domény. Jako administrativní a či technický kontakt může být uvedeno až 10 osob. U všech osob se musí uvést jméno, organizace (v případě právnické osoby), adresa, e-mail a případně další informace. Tyto údaje je nutné aktualizovat při každé změně, především e-mail, přes který registr komunikuje s danou osobou.
- Předposledním krokem je registrace sady jmenných serverů. V této položce budou uvedeny technické údaje o doméně. Je nutné zadat názvy a IP adresy jmenných serverů, kam bude doménové jméno delegováno a dále zde uvést osoby v roli technického kontaktu z předchozího kroku.
- Posledním krokem je registrace samotné domény. Zde se již pouze ke zvolenému doménovému jménu přiřadí osoby vytvořené v předchozích bodech a následně se dokončí proces registrace [3].

3 REALIZACE DOMÉNOVÉ LOKACE

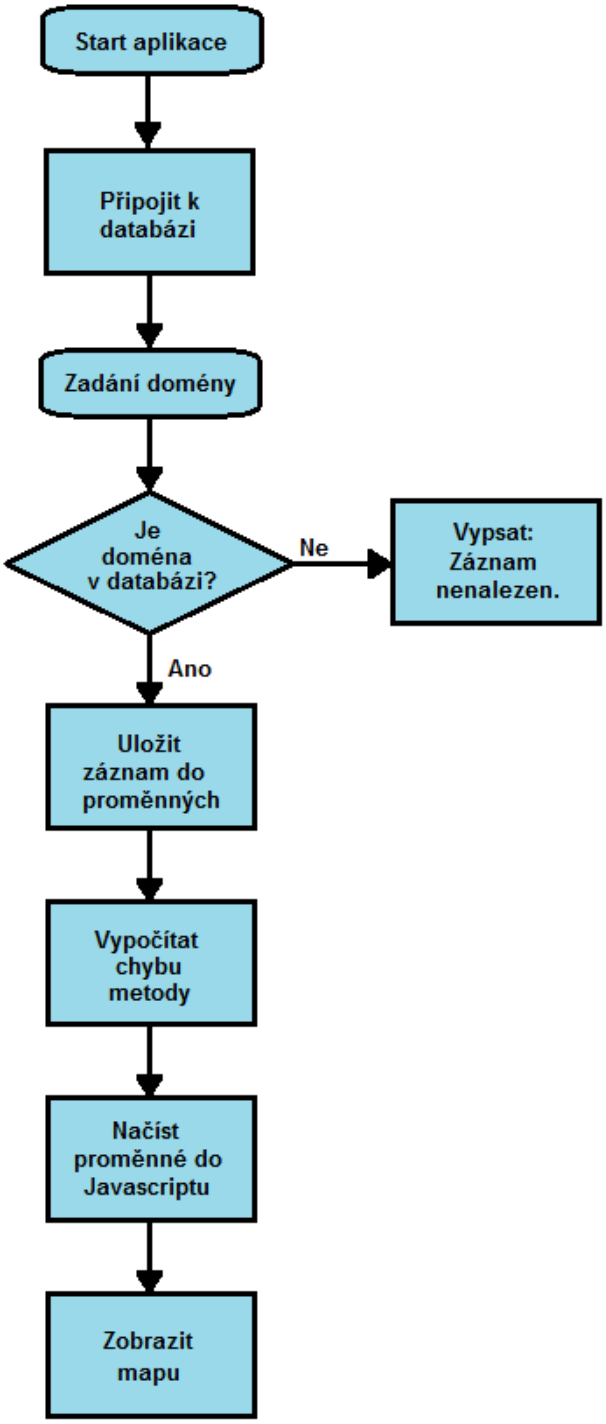
Kromě obsáhnutí dané problematiky a jejímu porozumění je hlavním výstupem této práce aplikace, vytvořená pro účely lokace stanice v síti Internet pomocí doménových jmen. V rámci semestrálního projektu byla vytvořena aplikace, která po zadání hledané domény tuto porovnala s vlastní databází a při nálezu shody vypsala geografickou polohu stanice a stát, ve kterém se nachází. Pro bakalářskou práci byla aplikace značně rozšířena. Databáze nyní obsahuje seznam domén všech evropských států, jejich nejvýznamnějších měst a organizací a u nich uvedených údajů, potřebných pro lokaci domén. Hlavním rozšířením je ale zobrazení výsledků hledání, kdy místo jednoduchého výpisu teď aplikace zobrazí hledanou stanici na mapě, s ní i skutečnou polohu stanice (pokud je známa) a u obou po najetí myší na ukazatel polohy vypíše údaje o poloze a chybě metody.

3.1 Vlastní program

Aplikace pro lokaci stanic v síti pomocí domén se skládá z několika částí, z nichž každá plní určité úkoly. Jako programovací jazyk jádra aplikace byl zvolen jazyk JavaFX Script [20]. Pro uložení potřebných dat byla vybrána databázová platforma MySQL, se kterou aplikace komunikuje jazykem SQL. Jako nejvhodnější řešení pro zobrazení výsledků bylo zvoleno zobrazení polohy pomocí GoogleMaps API [5], která pracuje s jazykem JavaScript [12].

Princip fungování celého programu je vidět na vývojovém diagramu. Nejprve je spuštěno aplikační okno, ze kterého se uživatel připojí k databázi. Následně zadá do příslušného prostoru doménové jméno hledané stanice a stiskne Vyhledat. Po stisknutí tohoto tlačítka aplikace načte zadanou doménu a porovná ji se záznamy v databázi. Pokud najde shodu, otevře nové okno, ve kterém zobrazí on-line mapu a na ní vyznačenou nalezenou polohu i s údaji o zeměpisné poloze. Pokud je známa i skutečná poloha stanice, která se může od nalezené lišit, zobrazí se na mapě také a u obou ukazatelů se kromě polohy vypíše i chyba metody, neboli vzdálenost obou ukazatelů.

Vývojový diagram lze vidět na obr. 3.1.



Obr. 3.1: Vývojový diagram aplikace

3.1.1 JavaFX

Hlavní částí aplikace je programový kód realizovaný v prostředí JavaFX. Je to softwarová platforma postavená na platformě Java, které obě vytvořila společnost Sun Microsystems [20]. Jejím hlavním užitím je tvorba tzv. bohatých internetových aplikací, které zvyšují interaktivitu webových stránek tím, že umožňují vzájemnou komunikaci mezi uživatelem a stránkou. Kromě těchto aplikací lze ale vytvořit spoustu dalších programů, jako je například přehrávač videa nebo aplikace pro práci s databázemi.

Samotný programovací jazyk JavaFX Script je koncipovaný tak, aby se programátor mohl soustředit na vytváření aplikace a nemusel se zabývat podrobnostmi kódu a jeho fungováním, například vykreslováním grafických prvků a podobně. Velkou výhodou je přenositelnost mezi různými platformami. Aplikace programované na platformě JavaFX plně podporuje jak Windows, tak MacOS i Linux. Pokud tedy vytvoříme aplikaci v prostředí Windows, nemusíme se bát, že nepůjde spustit v systému Linux. Výhodou je také podpora tříd programovaných v jazyku Java, které dokážou aplikace na platformě JavaFX bez problémů používat.

Pro tvorbu aplikace zde bylo využito prostředí e(fx)clipse. Je velmi přehledné, snadno ovladatelné a kromě samotného ověřování správnosti kódu a spuštění aplikace je lze využít i pro tvorbu speciálních souborů, které umožní použití aplikace ve webovém rozhraní. Samotný vzhled okna aplikace je pak tvořen v samostatném prostředí JavaFX Scene Builder, které je propojeno s prostředím e(fx)clipse. Vzhled aplikace je pak v podobě kódu definován v souboru `.FXML` který slouží k vytvoření grafického rozhraní. Kód se mění s tím, jak přidáváme či odebíráme prvky v Scene Builderu.

Všechny složky nutné pro používání platformy JavaFX a programování v jazyce JavaXF Script jsou volně dostupné a lze je stáhnout z oficiálních stránek bez jakýchkoliv poplatků či registrací [20].

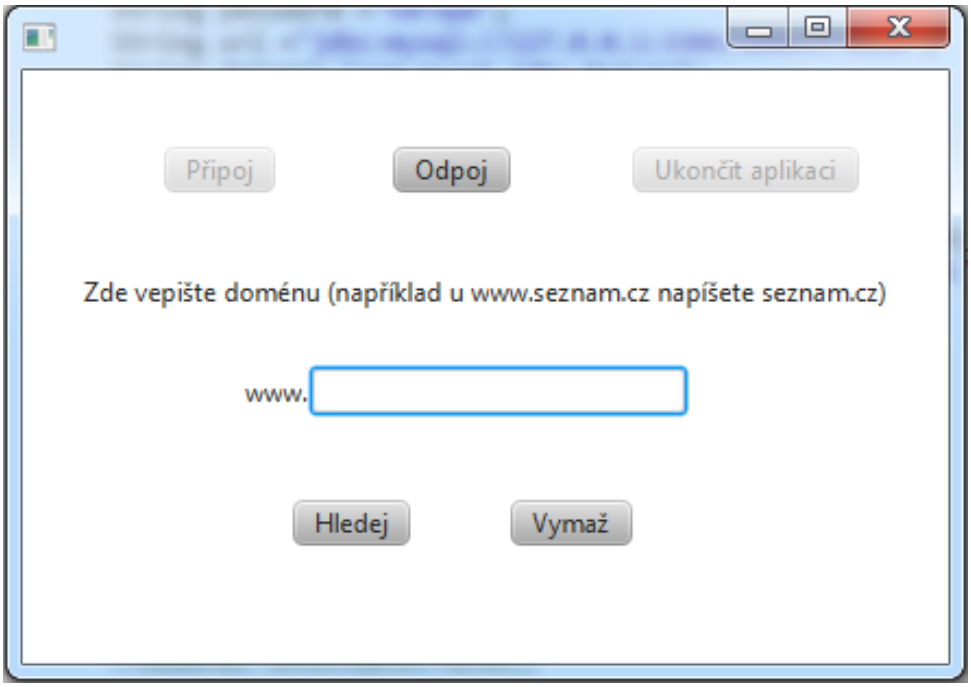
V prostředí e(fx)clipse byl vytvořen nový JavaFX projekt a v něm dvě třídy. Třída Main má za úkol samotné spuštění aplikace, tedy inicializaci okna aplikace. Vše ostatní je definováno ve třídě Kontrolak. Základem je naimportování všech potřebných balíčků, které lze ale plynule provádět při samotné tvorbě kódu. Následuje definice ovládacích a zobrazovacích prvků aplikace. Zde je potřeba zadat všechny prvky, které se mají objevit v aplikačním okně programu. Tyto prvky zahrnují tlačítka, prostor pro psaní textu a další. Poté přichází na řadu hlavní část aplikace – definice funkcí. Pro každý úkon, který má naše aplikace být schopna provést, je nutná samostatná definice. Pilířem celého programu je funkce pro připojení k databázi s uloženými údaji a funkce pro otevření nového okna a zobrazení on-line mapy v něm.

Pro přístup k databázi je nejprve nutné připojit se pomocí síťových služeb k lokálnímu serveru, na kterém je databáze spravována. Zadáním přihlašovacích údajů se zpřístupní samotná databáze a lze s ní pracovat. Následuje samotné vyhledávání, které se provádí pomocí příkazu jazyka SQL, ve kterém je databáze vytvořena. Nejprve je načten textový řetězec z příslušného prostoru okna aplikace a poté je tento řetězec porovnám s jednotlivými záznamy v databázi. Pokud je nalezena shoda, uloží se všechny údaje daného záznamu do lokálních proměnných a zavolá se funkce pro spouštění webového prohlížeče. Tato funkce nejprve nastaví nejzákladnější parametry prohlížeče, jako je velikost vyvolaného okna, jeho název či barva. Nejdůležitějším úkonem je spuštění funkce pro zobrazení mapy. Tato funkce provádí mnoho úkonů najednou. Nejprve načte lokální proměnné, do kterých byla uložena data z námi nalezeného záznamu z databáze. Tato data následně použije pro výpočet chyby metody, neboli vzdálenosti mezi zjištěnou a skutečnou polohou hledané stanice, a výsledek uloží do příslušné proměnné. K výpočtu je použit vzorec `Haversine` [17], který bere v potaz zakřivení zemského povrchu při výpočtu vzdálenosti mezi dvěma body.

Následuje vytvoření virtuálního dokumentu s koncovkou `.html`, ve kterém je definovaná celá webová stránka, která bude zobrazena. Definice stránky je provedena pomocí funkce `String.genHtml()`. Jsou zde definovány veškeré parametry samotné stránky, tedy cílové url, na které se aplikace připojí a které je základem celé stránky a všechny parametry, které tuto stránku pozměňují tak, aby vyhovovala cílům aplikace. Jedná se o url `http://maps.google.com/maps/api/js?sensor=false`, které umožní práci s on-line mapami společnosti Google [5]. Následují jednotlivé úpravy celé mapy. Především jde o stanovení geografické polohy, kterou mapa bude zobrazovat, dále vytvoření ukazatelů označujících polohu hledané stanice a také vytvoření rudé přímky mezi ukazateli, která znázorňuje přímou vzdálenost mezi oběma body a tedy i chybu určení polohy. Do všech těchto parametrů jsou dynamicky načítány lokální proměnné tak, aby zobrazená mapa odpovídala údajům zadaným do aplikace. Jakmile je tvorba stránky dokončena, je uložena do textového řetězce. Tento řetězec je načten funkcí pro zobrazení stránky a výsledkem je otevření webového prohlížeče a zobrazení polohy hledané stanice na on-line mapě.

Kromě těchto nejdůležitějších funkcí je dále definován příkaz vymazání námi zapsaného řetězce, odpojení od databáze a ukončení aplikace.

Po samotném programování bylo nutné okno aplikace vytvořit v prostředí JavaFX Scene Builder. Zde byly definovány rozměry okna, jeho vzhled a především umístěny samotné ovládací prvky. Ty bylo nutné propojit s kódem v e(fx)clipse a jejich aktivaci přiřadit odpovídající funkci. Samotné aplikační okno lze vidět na obr. 3.2.



Obr. 3.2: Okno aplikace

3.1.2 Databáze

Druhou částí aplikace je databáze. Zde jsou uloženy záznamy, obsahující doménová jména a geografickou polohu všech evropských států, jejich nejvýznamnějších měst, organizací a univerzit. Databáze byla vytvořena na platformě MySQL, která je pro podobné účely velmi vhodná. Ke zprovoznění této platformy byl použit softwarový balík Xampp, který obsahuje především serverové a klientské technologie využívající Apache HTTP server či databázi MySQL a prostředí pro práci s jazyky PHP a Perl. Podobně jako JavaFX je určen především pro tvorbu webových aplikací a jejich příslušenství a je také podporován všemi běžnými operačními systémy. Díky své kompaktnosti šetří čas vývojářům, kteří nemusí stahovat několik aplikací a složitě je navzájem propojovat a spouštět. Stačí si stáhnout volně přístupný Xampp soubor, vybrat požadovaný software a instalátor se postará o zbytek. Správa pak probíhá přes webový prohlížeč v prostředí phpMyAdmin. Pro potřeby aplikace byl nainstalován Apache server a na něm podpora MySQL databází. [2]

Databáze je tvořena tabulkou záznamů, kde u každého jsou uvedeny potřebné údaje. Na každém řádku je uvedeno ID záznamu, doména, souřadnice nalezené polohy a souřadnice skutečné polohy. Záznamy lze seřadit sestupně i vzestupně podle hodnot kteréhokoliv sloupce. Celkem je v tabulce 243 záznamů. Většina záznamů je dvouúrovňových, tedy obsahujících dvě úrovně domén. Jedná se o doménu druhého řádu, pomocí které lze určit polohu stanice s přesností na jednotlivá města a doména

nejvyššího řádu, která rozlišuje stanice s přesností na jednotlivé státy. Tento typ domén mají především města, organizace a univerzity. Některé organizace a univerzity ale mají v doménovém jméně více než jednu doménu konkrétního řádu. Nejčastěji jde o více domén druhého řádu, což může sloužit například při rozdělení univerzity na jednotlivé samostatné fakulty. V některých případech se ale může objevit i více než jedna doména nejvyššího řádu. Příkladem mohou být některé z nejznámějších univerzit ve Velké Británii. Vzhled databáze lze vidět na obr. 3.3.

domena_id	domena	Stát	Souřadnice	Souradnice_web	Sour_web_skut	Lat1	Long1	Lat2	Long2
61	x-file.de	Německo	N 49°29', E 11°06'	49.45,11.07	49.11,10.75	49.45	11.07	49.11	10.75
62	cablelink.at	Rakousko	N 47°45', E 13°9'	47.81,13.05	47.80,13.06	47.81	13.05	47.80	13.06
63	telenor.hu	Maďarsko	N 47°30', E 19°1'	47.49,19.04	47.49,19.04	47.49	19.04	47.49	19.04
64	compower.pl	Polsko	N 50°5', E 19°59'	50.07,19.94	50.06,19.94	50.07	19.94	50.06	19.94
65	minet.sk	Slovensko	N 48°45', E 21°49'	48.75,21.91	48.75,21.91	48.75	21.91	48.75	21.91
66	sbb.rs	Srbsko	N 44°45', E 20°29'	44.82,20.46	44.82,20.46	44.82	20.46	44.82	20.46
67	connecta.pl	Polsko	N 51°35', E 17°9'	51.11,17.04	51.1,17.04	51.11	17.04	51.1	17.04
68	upcbiz.ro	Rumunsko	N 44°28', E 26°0'	44.43,26.1	44.43,26.1	44.43	26.1	44.43	26.1
69	swan.sk	Slovensko	N 48°8', E 17°06'	48.14,17.11	48.14,17.11	48.14	17.11	48.14	17.11
70	telekom.rs	Srbsko	N 44°50', E 20°26'	44.82,20.46	45.82,20.47	44.82	20.46	45.82	20.47
71	multiplay.co.uk	Velká Británie	N 50°55', W 1°26'	50.91,-1.4	51.5,-0.13	50.91	-1.4	51.5	-0.13
72	universnet.ro	Rumunsko	N 44°18', E 26°06'	44.33,26.08	44.33,26.08	44.33	26.08	44.33	26.08
73	ftp.funet.fi	Finsko	N 60°05', E 24°36'	60.2,24.66	60.2,24.66	60.2	24.66	60.2	24.66
74	dist.unige.it	Itálie	N 44°23', E 8°56'	44.4,8.95	44.41,8.93	44.4	8.95	44.41	8.93
75	decor.eik.bme.hu	Maďarsko	N 47°28', E 19°06'	47,41,19,13	47.48,19.06	47,41	19,13	47.48	19.06
76	uni-essen-duisburg.de	Německo	N 51°38', E 7°	51.46,7	51.46,7	51.46	7	51.46	7
77	www.uio.no	Norsko	N 59°53', E 10°46'	59.91,10.75	59.93,10.72	59.91	10.75	59.93	10.72
78	www.icm.edu.pl	Polsko	N 52°18', E 21°06'	52.24,21.02	52.24,21.02	52.24	21.02	52.24	21.02
79	www.umu.se	Švédsko	N 63°8', E 20°18'	63.82,20.3	63.82,20.3	63.82	20.3	63.82	20.3

Obr. 3.3: Databáze domén

3.1.3 JavaScript

Jedná se o objektově orientovaný programovací jazyk určený pro vytváření webových stránek a pracující a různých platformách. V roce 1998 byl standardizován organizací ISO (International Organization for Standardization) [9]. Slovo Java v názvu je poněkud zavádějící, protože s jazykem Java má společnou pouze podobnou syntaxi. Používá se především pro tvorbu dynamických webových rozhraní a aplikací. Většinou je použit k ovládání interaktivních prvků GUI (různá tlačítka, proměnlivé obrázky a podobně). Samotný kód je často zahrnut přímo v HTML kódu dané webové stránky. Na rozdíl od jiných jazyků se JavaScript často spouští na straně klienta, tedy až po stažení webové stránky, nikoli na straně serveru [12].

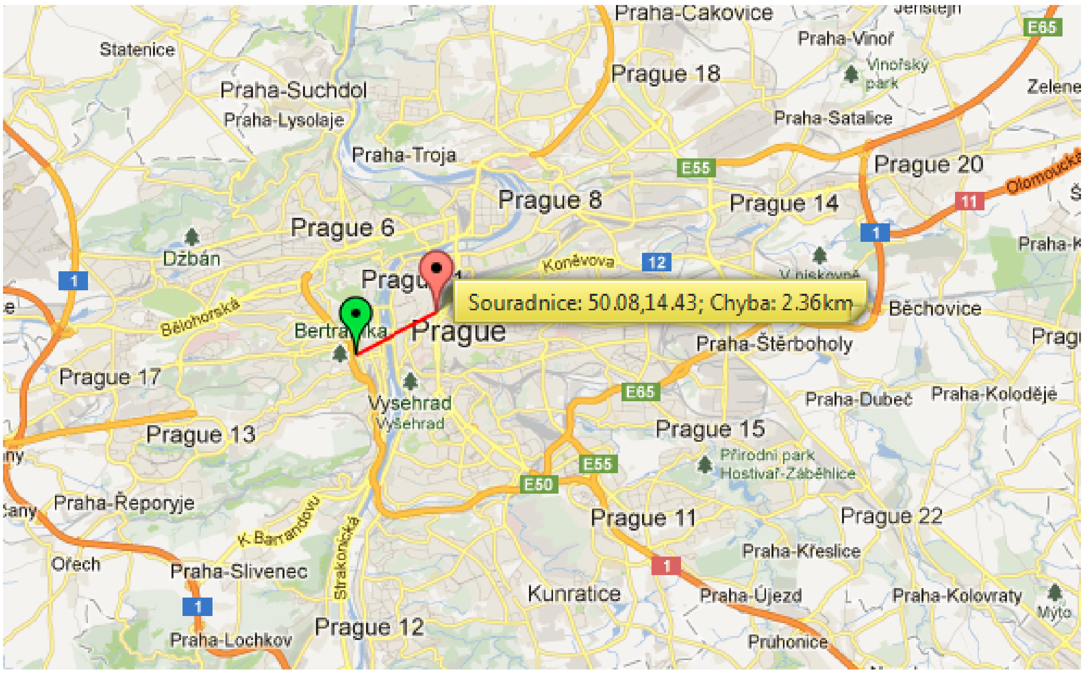
Pro účely aplikace je využita GoogleMaps API, což je aplikace umožňující fungování on-line map společnosti Google, která využívá právě jazyk JavaScript. Umož-

ňuje zprovoznit mapy na našich vlastních stránkách a všemožnými způsoby s nimi manipulovat [5].

Nejprve je nutné samotný JavaScript definovat v HTML souboru a dále stanovit cílové URL, se kterým bude aplikace pracovat. Následuje samotný kód, který postupně upravuje mapu tak, aby splňovala naše požadavky. Existuje nepřeberné množství různých úprav mapy. Základem je možnost přepínat mezi jednotlivými typy zobrazení a volba místa, které chceme zobrazit. Dále je ale možné mapu různě přibližovat, oddalovat či naklánět, označovat jednotlivé pozice ukazateli, nechat si mezi nimi vykreslit nejkratší cestu a mnoho dalších možností.

3.1.4 Výstup

Po úspěšném průběhu všech funkcí programu se otevře nové okno s on-line mapou a polohou hledané stanice. Pokud je známá i skutečná poloha stanice, objeví se na mapě také. Obě polohy jsou vyznačeny barevně odlišnými ukazateli (zelený pro skutečnou a červený pro zjištěnou polohu), mezi ukazateli se objeví červená spojnice těchto dvou bodů a po najetí myší na ukazatel se zobrazí jeho poloha a vzdálenost mezi ukazateli (neboli chyba metody). To lze vidět na obr. 3.4 a obr. 3.5. Hlavním výstupem celé aplikace je ale určení přesnosti celé metody lokace stanic v síti pomocí doménových jmen. Toho bylo dosaženo zjištěním určitého počtu poloh stanic a zanesením chyby do grafu distribuční-kumulační funkce. Graf i další informace jsou uvedeny v následující podkapitole.



Obr. 3.4: Určení polohy pro `www.seznam.cz`

Obr. 3.5: Určení polohy pro `www.mimas-nxge0.switch.ch`

3.2 Dosažené výsledky

Kromě již zmíněného vizuálního výstupu je hlavním výstupem aplikace soubor dat, ze kterých lze určit chybu metody. Jedná se o vzdálenosti zjištěných a skutečných poloh hledaných stanic. Z těchto hodnot byl vytvořen průměr, medián, percentily a především distribuční-kumulační funkce. Měření bylo provedeno na souboru předem vybraných stanic. Tento seznam jsem vytvořil společně s kolegy Henkem, Pokorným, Horákem a Mrníkem, kdy každý přidal 10 veřejně dostupných stanic, u kterých je známá jejich geografická poloha (obr. 3.6). Seznam domén a poloh těchto serverů jsem vložil do databáze mé aplikace. Pomocí postupného zadávání jednotlivých domén jsem pak zobrazil skutečné a zjištěné polohy serverů a zjistil jsem chybu metody, kterou aplikace vypočítala.

3.2.1 Chyba metody

Chybou metody je myšlena vzdálenost mezi skutečnou geografickou polohou hledané stanice a polohou, kterou dokáže určit moje aplikace. Tuto vzdálenost po zadání údajů a načtení příslušných souřadnic z databáze vypočítá aplikace. K výpočtu byl použit vzorec Haversine [17], který je vhodný pro výpočet vzdálenosti mezi dvěma body na zemském povrchu, protože do svého výpočtu zahrnuje i zakřivení Země. Získané hodnoty jsem zanesl do tabulky a vytvořil z nich graf distribuční-kumulační funkce, který znázorňuje přesnost metody.

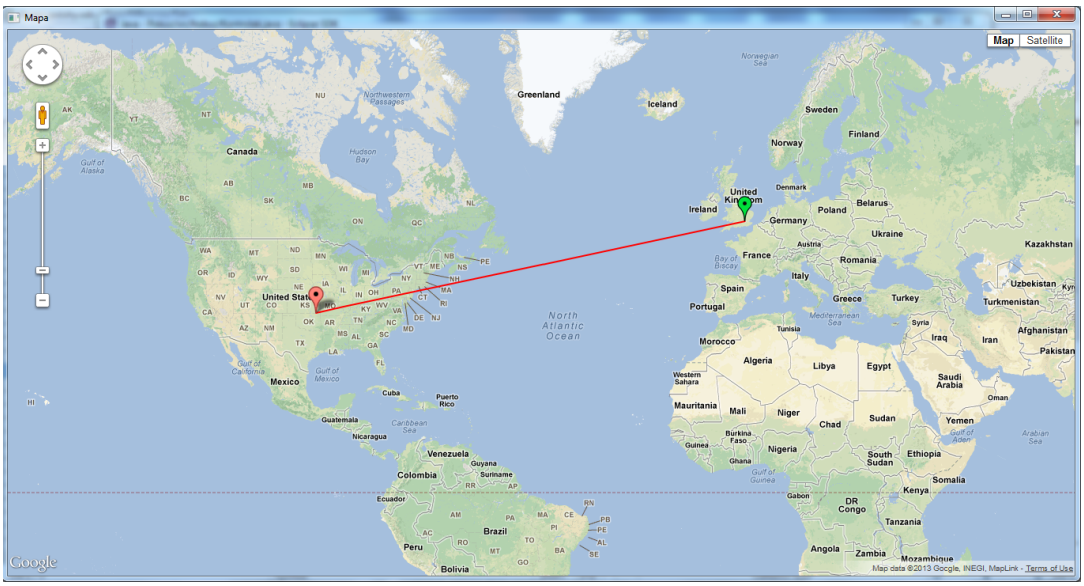
IP address	full location	latitude	longtitude	author
176.9.55.42	Německo	51.165691	10.451526	Henek Jan
91.82.84.185	Budapešť, 05, Maďarsko	47.498406	19.040758	Henek Jan
77.251.170.54	Heusden, 06, Nizozemsko	51.733004	5.138279	Henek Jan
83.222.115.86	Moskva, 48, Rusko	55.751242	37.618422	Henek Jan
80.72.40.110	Varšava, 78, Polsko	52.229676	21.012229	Henek Jan
89.215.114.195	Plovdiv, 51, Bulharsko	42.143841	24.749562	Henek Jan
109.70.149.49	Stroud, E6, Anglie (UK)	51.745734	-2.217758	Henek Jan
85.11.157.54	Sofia, 42, Bulharsko	42.696492	23.326011	Henek Jan
174.137.191.6	Amsterdam, The Netherlands	52.370216	4.895168	Mrnik Martin
46.20.125.254	London, UK	51.507335	-0.127683	Mrnik Martin
78.46.90.47	Gunzenhausen, Germany	49.114722	10.754167	Mrnik Martin
88.198.19.202	Nuernberg, Germany	49.45203	11.07675	Mrnik Martin
212.67.73.150	Praha 9	50.107473	14.502673	Mrnik Martin
5.9.59.165	Germany, Frankfurt	50.110922	8.682127	Mrnik Martin
213.153.32.170	Salzburg, Austria	47.80949	13.05501	Mrnik Martin
217.79.128.22	Budapest, Hungary	47.498406	19.040758	Mrnik Martin
194.146.252.199	Krakow, Poland	50.06465	19.94498	Mrnik Martin
213.94.75.9	Vienna, Austria	48.208174	16.373819	Mrnik Martin
94.136.136.2	Michalovce, Slovakia	48.755995	21.914858	Mrnik Martin
89.216.2.122	Belegrad, Serbia	44.820556	20.462222	Mrnik Martin
91.90.160.3	Wroclaw,Poland	51.107885	17.038538	Mrnik Martin
95.77.94.89	Bucharest, Romania	44.437711	26.097367	Mrnik Martin
213.249.64.165	Amsterdam, The Netherlands	52.370216	4.895168	Mrnik Martin
62.65.173.6	Bratislava, Slovakia	48.146239	17.107262	Mrnik Martin
77.75.76.3	Radlická 608/2, 150 00 Praha 5, Czech Republic	50.071585	14.400793	Jelínek Ondřej
212.200.163.178	Kikinda, Serbia	45.828333	20.465278	Jelínek Ondřej
85.236.100.91	London, UK	51.507335	-0.127683	Jelínek Ondřej
188.165.34.68	Paris, France	48.856614	2.352222	Jelínek Ondřej
88.190.22.159	Draveil, France	48.685388	2.408154	Jelínek Ondřej
93.115.207.238	Jilava, Romania	44.333333	26.083333	Jelínek Ondřej
81.201.56.141	Husova 58, 301 00 Plzeň	49.746106	13.364221	Pokorný Josef
193.166.3.2	Keilaranta 14, Espoo, Finland	60.205479	24.655884	Pokorný Josef
130.251.19.2	Via Balbi 5, 16126 Genova, Italy	44.415103	8.925931	Pokorný Josef
152.66.115.224	Muegyetem rkp. 9., H-1111 Budapest, Hungary.	47.481321	19.056363	Pokorný Josef
132.252.181.87	Universitaetsstr. 1-15, 45117 Essen, Germany	51.463286	7.004256	Pokorný Josef
129.240.8.200	Oslo, Norway	59.939948	10.721838	Pokorný Josef
212.87.14.41	University of Warsaw, 00-927 Warsaw, Poland	52.240402	21.019206	Pokorný Josef
130.239.141.10	Umea University, 901 87 Umea, Sweden	63.820539	20.303591	Pokorný Josef
130.235.209.220	Lund University, S-222 40 Lund, Sweden	55.71083	13.205265	Pokorný Josef
130.59.10.36	SWITCH, SWITCHmirror, CH-8021 Zurich, Switzerland	47.36865	8.539183	Pokorný Josef
77.93.192.144	Brno - Židenice	49.2	16.6333	Michael Horák
81.2.194.154	CZ	49.75	15,5	Michael Horák
91.235.52.167	SK	48.6667	19,5	Michael Horák
131.175.187.11	Milano	45.4667	9.193928	Michael Horák
129.27.201.245	Gratz	47.06938	15.450465	Michael Horák
193.136.163.66	Lisabon, Portugalsko	38.7167	9.1333	Michael Horák
109.70.148.245	London, UK	51.5142	0.0931	Michael Horák
77.47.133.22	Kyiv	50.4333	30.5167	Michael Horák

Obr. 3.6: Seznam serverů použitých pro zjištění přesnosti metody

Zjistil jsem, že nejvyšší chyba metody byla 147,61 km, což bylo způsobeno tím, že z dané domény nešlo určit polohu stanice s přesností na město, ale na celý stát. Naopak nejnižší chyba byla dosažena u domény `telenor.hu` a to 0,06 km. Tak nízká chyba metody vznikla díky tomu, že z analýzy domény se určila poloha s přesností na město a sídlo firmy bylo v blízkosti středu města. Průměrná chyba byla 24,72 km, ovšem medián vyšel 2,27 km. Tento značný rozdíl mezi průměrem a mediánem byl způsoben tím, že více než polovina vypočítaných chyb byla velmi nízké hodnoty,

řádově jednotky kilometrů, ale bylo zde i několik chyb, které přesáhly 100 km a velmi tak zvýšily výslednou hodnotu průměru.

Při pohledu na graf lze vidět, jak s rostoucí vzdáleností vzrůstá i pravděpodobnost, že chyba metody je menší, než daná vzdálenost. Například lze vidět, že s pravděpodobností 10 % bude chyba přibližně nulová, s pravděpodobností 50 % bude chyba menší nebo rovna přibližně dvěma kilometrům anebo že s pravděpodobností 90 % bude chyba menší nebo rovna zhruba 100 km. Z důvodu omezeného rozsahu mé databáze jsem musel z měření vynechat servery, které mají ve svém doménovém jméně generickou doménu nejvyššího řádu. Tyto domény jsou totiž spravovány organizacemi, které mají své sídlo v USA. Pokud bych tedy počítal chybu mezi skutečnou a zjištěnou polohou u serveru, který sídlí v Evropě, ale používá generickou TLD, vyšla by chyba metody kolem 7000 km, což by značně zkreslilo výsledky celé metody (obr. 3.7). Takových serverů bylo v seznamu celkem 10 (obr. 3.8).

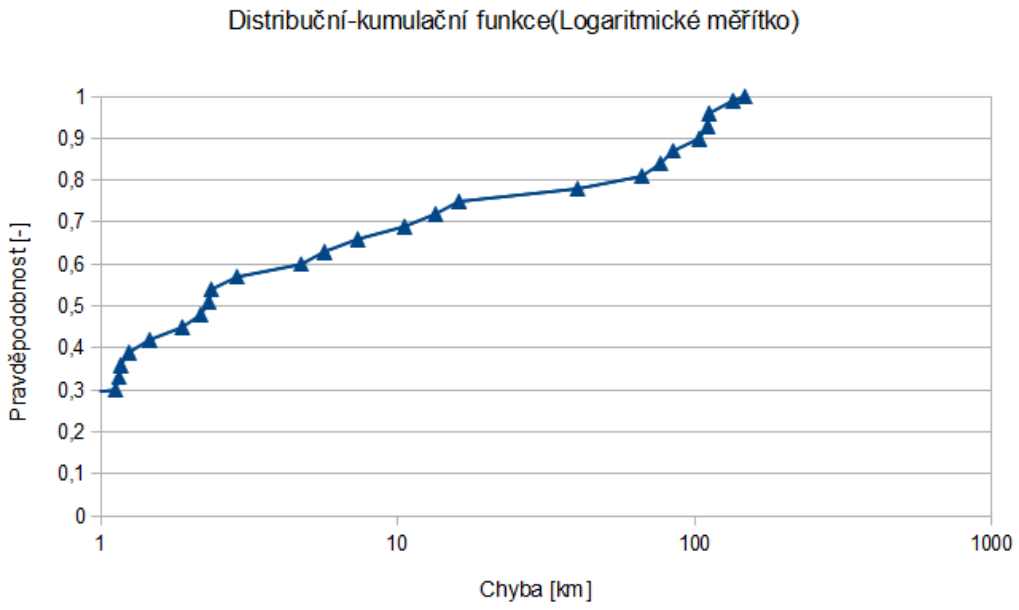


Obr. 3.7: Určení polohy stanice využívající generickou doménu

Při pohledu na graf (obr. 3.9) lze jasně vidět, že chyba metody je až do pravděpodobnosti 70 % menší nebo rovna 10 km, a teprve poté chyba se vzrůstající pravděpodobností roste dál až k hodnotě 150 km. Z toho je jasně vidět, že tato metoda lokace je pro velkou část stanic velice přesná a je tudíž vhodná pro reálné využití. Otázkou zůstává, jak by se měnila přesnost se zahrnutím domén z celého světa. Nevýhodou také zůstává, že stanice používající generickou doménu nejvyššího řádu budou způsobovat značnou chybu metody, protože v naprosté většině případů je jejich skutečná poloha vzdálená stovky až tisíce kilometrů od sídla správce generické domény. Tuto nevýhodu lokace pomocí analýzy doménových jmen bohužel nelze odstranit.

IP address	full location	latitude	longtitude	Domain name
176.9.55.42	Německo	51,16	10,45	n1ping.com
85.11.157.54	Sofia, 42, Bulharsko	42,69	23,33	sofianet.net
174.137.191.6	Amsterdam, The Netherlands	52,37	4,9	accountservergroup.com
46.20.125.254	London, UK	51,50	-0,13	accountservergroup.com
88.198.19.202	Nuernberg, Germany	49,45	11,08	flyin.org.
213.249.64.165	Amsterdam, The Netherlands	52,37	4,9	NETWORKING4ALL.COM
188.165.34.68	Paris, France	48,85	2,35	kimsufi.com
88.190.22.159	Draveil, France	48,68	2,41	loup-des-neiges.com
81.201.56.141	Husova 58, 301 00 Plzeň	49,74	13,36	icewow.pilsfree.net
81.2.194.154	Česká Republika	49,75	15,51	forpsi.com

Obr. 3.8: Seznam vyřazených serverů z USA



Obr. 3.9: Distribuční–kumulační funkce

3.2.2 Srovnání dalších metod

Určení polohy stanice v síti pomocí analýzy doménových jmen je pouze jedním z mnoha způsobů, jak tuto polohu určit. Moji kolegové Jan Henek, Bc. Michael Horák a Bc. Josef Pokorný určovali polohu stanic odlišnými metodami.

Kolega Henek využíval pro zjištění polohy stanic pasivní metodu geolokace a analýzou IP adres a s využitím databáze Whois. Tato veřejně a bezplatně dostupná databáze zahrnuje obrovské množství záznamů, obsahujících konkrétní IP adresy či jejich rozsahy. V každém záznamu jsou kromě IP uvedeny i základní údaje, jako je správce domény nejvyššího řádu, pod kterou registrovaná doména spadá, vlastník domény, kontaktní údaje vlastníka domény a především geografická poloha stanice, k níž IP adresa patří. Tato metoda geolokace vyšla v porovnání s ostatními meto-

dami jako třetí nejhorší, ovšem i přesto je velmi dobře použitelná, protože až do pravděpodobnosti 60 % je chyba metody v jednotkách kilometrů. Je to způsobeno tím, že některé záznamy v databázi mají uvedenou jako polohu pouze daný stát, ale ne už konkrétní město či přímo adresu. Maximální chyba metody je přibližně 1600 km.

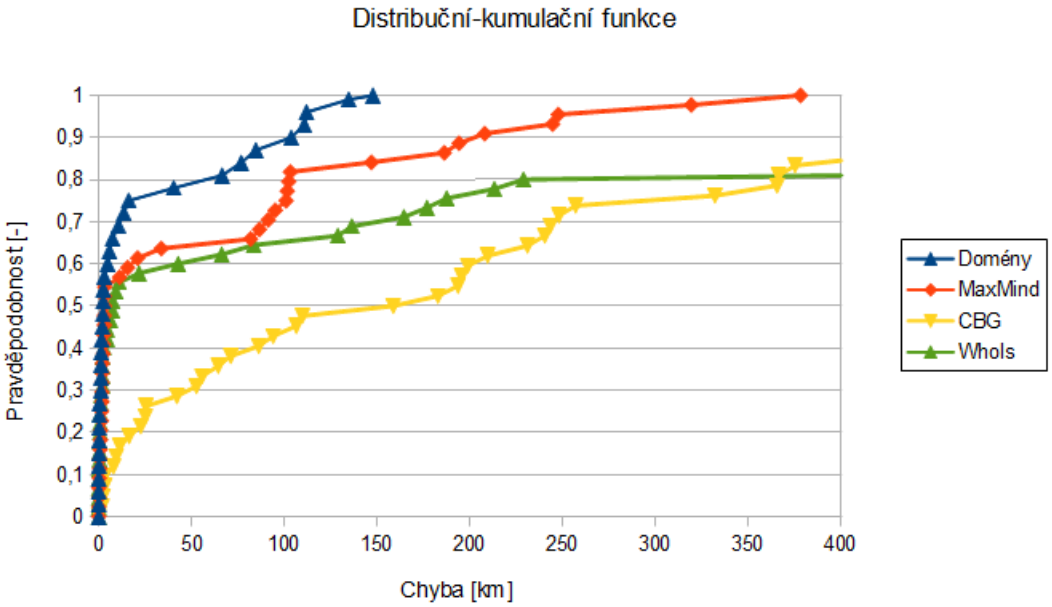
Kolega Horák se věnoval metodě CBG. Jedná se o aktivní metodu geolokace, která k určení polohy stanice využívá měření zpoždění mezi hledanou stanicí a pevně umístěnými sondami v rámci projektu Planetlab, které lze k měření použít. Pro samotnou lokalizaci každá sonda samostatně změří zpoždění k hledané stanici a z něj vypočítá poloměr kruhové oblasti, ve které se stanice bude nacházet a jejímž středem je právě daná sonda. V průniku oblastí jednotlivých sond se pak nachází hledaná stanice. Metoda CBG vyšla v celkovém porovnání jako nejhorší, s rostoucí pravděpodobností roste chyba metody nejrychleji a také maximální chyba metody, která je rovna 1783 km, je největší. Tento horší výsledek mohl být do značné míry způsobený právě zpožděním, které je měřeno. To se totiž může v čase lišit i v rámci jediné linky a v rámci celé sítě, přes kterou signál od sondy ke stanici prochází, může být zpoždění ovlivněno mnoha faktory na mnoha místech trasy spojení.

Kolega Pokorný se zabýval, podobně jako kolega Henek, pasivní metodou geolokace pomocí IP adres s využitím internetové databáze. K analýze IP adres ale použil databázi MaxMind, místo databáze Whois. Tato databáze patří společnosti MaxMind a jedná se o komerční projekt. Pro využití této databáze je třeba zakoupit licenci, jejíž cena se odvíjí od počtu námi provedených hledání v databázi. Je zde ovšem i možnost jakési demoverze, tedy omezené neplacené využití této databáze. V takovém případě je počet vyhledání záznamů, které můžeme provést, značně omezen. Díky tomu, že je databáze MaxMind komerční, je provozována jako profesionální projekt a je tedy odpovídajícím způsobem spravována. Záznamy obsahují v podstatě stejné údaje, jako databáze Whois, ale většina ze záznamů obsahuje buď přímo adresu stanice, nebo alespoň město, ve kterém se stanice nachází. Díky tomu je metoda využívající databázi MaxMind přesnější, než při využití databáze Whois. V celkovém srovnání se zařadila na druhé místo. S pravděpodobností více než 60 % je chyba metody pouze v jednotkách kilometrů a až do hodnoty přesahující 80 % je nejvyšší chyba zhruba 100 km. Maximální chyba metody je pak zhruba 380 km.

Získaná výstupní data jsem všichni zanesli do grafu distribuční kumulační funkce a všechny tyto grafy jsme sloučili do jednoho. Díky tomu lze snadno porovnat přesnost použitých metod. Jak již bylo zmíněno, jako nejpřesnější vyšla metoda lokace pomocí analýzy doménových jmen, následovaná analýzou IP adres s využitím databáze MaxMind na druhém místě a využití databáze Whois na třetím místě. Jako nejméně přesná vyšla metoda CBG.

Nevýhodou metody doménových jmen je, že pro její využití v praxi by musela být

vytvořena nová rozsáhlá databáze, která by obsahovala obrovské množství záznamů domén a k nim přiřazených údajů. Vytvoření takové databáze by bylo finančně a především časově velice náročné. Z toho důvodu je pro praktické využití lepší použít již existující databáze Whois či MaxMind. Zjištění IP adresy konkrétní stránky lze provést jednoduchým příkazem v příkazové řádce bez ohledu na používaný operační systém.



Obr. 3.10: Vzájemné porovnání přesnosti jednotlivých metod

4 ZÁVĚR

Úkolem této práce bylo vytvoření aplikace pro určení geografické polohy hledané stanice v síti Internet a to díky analýze doménových jmen. K tomu bylo potřeba nejprve porozumět dané problematice.

V první kapitole byla představena metoda určení geografické polohy hledaného cíle zvaná geolokace. Byl stručně popsán její princip, rozdělení na aktivní a pasivní metody, jejich výhody a nevýhody a především jejich využití v praxi.

Druhá kapitola se věnovala stěžejnímu tématu této práce – systému DNS a doménovým jménům, která pod tento systém spadají. Nejprve byla detailně popsána doménová jména, jejich struktura a funkce. Poté se práce zabývala již samotným systémem DNS a všemi jeho náležitostmi.

Poslední kapitola se zabývá hlavním výstupem této práce – vytvořenou aplikací. Tato aplikace byla vytvořena pomocí platformy a jazyka JavaFX a pracuje s databází záznamů, které obsahují doménová jména konkrétních stanic a jejich geografickou polohu. Těchto záznamů se v databázi nachází celkem 243. Po zadání domény hledané stanice do aplikace je zobrazena online mapa, na které je vyznačena skutečná a aplikací zjištěná poloha stanice. Tyto dvě polohy se mohou lišit, proto je mezi nimi vyznačena rudá přímka a aplikace vypočítá a vypíše chybu metody při určování polohy.

Přesnost aplikace byla změřena na souboru dat, ve kterém je uvedeno zhruba 50 reálných stanic se známou polohou. Z výsledků je jasné, že lokace stanic pomocí analýzy doménových jmen je velice přesnou metodou. S pravděpodobností 70 % je chyba metody při určování polohy rovna 10 km nebo menší a maximální chyba metody je 150 km. Tyto hodnoty ovšem vychází z konkrétní množiny stanic a při použití v praxi na různé domény by se mohla přesnost metody měnit, jelikož u použitých stanic bylo zjištěno přímo sídlo firmy či univerzity, pod kterou stanice spadá, a z toho vyplývá i velmi přesné určení polohy. U některých domén ovšem nelze zjistit všechny informace s potřebnou přesností a tudíž by se přesnost celé metody mohla aplikováním na různé domény zhoršit. Velkou nevýhodou, se kterou bohužel nelze nic dělat, je určování polohy stanice využívající generickou doménu. Správci generických domén mají sídlo v USA a pokud se tedy stanice nachází jinde, než v Severní Americe, bude chyba metody v řádech několika tisíc kilometrů.

LITERATURA

[1] ACTIVE 24. *Domény a DNS* [online]. 2010 [cit. 5. 12. 2012]. Dostupné z: <http://napoveda.active24.cz/idx.php/5/0/>.

[2] APACHE FRIENDS. *Apache Friends - XAMPP* [online]. 2013 [cit. 4. 5. 2013]. Dostupné z: <http://www.apachefriends.org/en/xampp.html>.

[3] CZ.NIC. *Jak registrovat doménu .cz* [online]. 2013 [cit. 1. 5. 2013]. Dostupné z: <http://www.nic.cz/page/313/>.

[4] CZ.NIC. *O DNSSEC* [online]. 2012 [cit. 5. 12. 2012]. Dostupné z: <http://www.dnssec.cz/page/444/jak-funguje-dnssec/>.

[5] Google Developers. *Google Maps API* [online]. 2013 [cit. 4. 5. 2013]. Dostupné z: <https://developers.google.com/maps/?hl=cs>.

[6] Christopher Davis. *DNS LOC: Geo-enabling the Domain Name System* [online]. 2001 [cit. 30. 4. 2013]. Dostupné z: <http://www.ckdhr.com/dns-loc/>.

[7] IANA. *Root Servers* [online]. 2007 [cit. 5. 12. 2012]. Dostupné z: <http://www.iana.org/domains/root/servers>.

[8] ISC. *BIND* [online]. 2013 [cit. 30. 4. 2013]. Dostupné z: <https://www.isc.org/software/bind>.

[9] ISO. *ISO 3166-1 decoding table* [online]. 2012 [cit. 5. 12. 2012]. Dostupné z: <http://www.iso.org/iso/iso-3166-1_decoding_table.html>.

[10] MaraDNS. *A small open-source DNS server* [online]. 2013 [cit. 29. 4. 2013]. Dostupné z: <http://www.maradns.org/index.html>.

[11] MaxMind, Inc. *GeoIP/IP Address Location Database* [online]. 2012 [cit. 5. 12. 2012]. Dostupné z: <http://www.maxmind.com/en/geolocation_landing>.

[12] MDN. *JavaScript - MDN* [online]. 2012 [cit. 4. 5. 2013]. Dostupné z: <https://developer.mozilla.org/en-US/docs/JavaScript>.

[13] Microsoft, Inc. *DNS* [online]. 2012 [cit. 5. 12. 2012]. Dostupné z: <http://technet.microsoft.com/en-us/library/bb742582.aspx>.

[14] Microsoft, Inc. *Používání nástroje NSlookup.exe* [online]. 2012 [cit. 5. 12. 2012]. Dostupné z: <http://support.microsoft.com/kb/200525/cs>.

[15] Microsoft, Inc. *Role serveru DNS* [online]. 2012 [cit. 5. 12. 2012]. Dostupné z: <http://technet.microsoft.com/cs-cz/library/cc753635%28v=ws.10%29.aspx>.

[16] MOCKAPETRIS, P. *DOMAIN NAMES - IMPLEMENTATION AND SPECIFICATION* [online]. 1987 [cit. 5. 12. 2012]. Dostupné z: <http://www.ietf.org/rfc/rfc1035.txt>.

[17] Movable Type Scripts. *Calculate distance and bearing between two Latitude/Longtitude points using Haversine formula in JavaScript* [online]. 2010 [cit. 4. 5. 2012]. Dostupné z: <http://www.movable-type.co.uk/scripts/latlong.html>.

[18] Network Working Group. *A Means for Expressing Location Information in the Domain Name System* [online]. 1996 [cit. 30. 4. 2013]. Dostupné z: <http://www.ietf.org/rfc/rfc1876.txt>.

[19] NTC Hosting. *Popular Open-Source DNS Software* [online]. 2009 [cit. 29. 4. 2013]. Dostupné z: <http://www.ntchosting.com/dns/software.html>.

[20] Oracle, Inc. *JavaFX* [online]. 2012 [cit. 5. 12. 2012]. Dostupné z: <http://www.oracle.com/us/technologies/java/fx/overview/index.html>.

[21] Svět Hostingu. *Co je WHOIS* [online]. 2012 [cit. 5. 12. 2012]. Dostupné z: <http://svet-hostingu.cz/2009/07/17/co-je-whois/>.

[22] ŠŤASTNÝ, Petr. *Průvodce DNS* [online]. 2012 [cit. 5. 12. 2012]. Dostupné z: <http://www.dns-info.cz/dns/index.html>.

[23] VERNER, Lukáš a KOMOSNÝ, Dan. Geolokace síťových zařízení v internetových sítích *Elektrorevue* [online]. 2011, č.33, s.7. ISSN 1213-1539. Dostupné z: <http://www.elektrorevue.cz/cz/clanky/komunikacni-technologie/40/geolokace-sitovych-zarizeni-v-internetovych-sitich/>.

[24] WEDOS. *Autoritativní DNS servery* [online]. 2010 [cit. 5. 12. 2012]. Dostupné z: <http://kb.wedos.com/dns/teorie/domeny-dns.html>.

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

API rozhraní pro programování aplikací – Application Programming Interface

CBG geolokace pomocí zpoždění – Constraint-Based Geolocation

DNS systém doménových jmen – Domain Name System

GPS globální polohový systém – Global Positioning System

GUI grafické uživatelské rozhraní – Graphical User Interface

IP internetový protokol – Internet Protocol

ISO mezinárodní organizace pro tvorbu norem – International Organization for Standardization

LIR lokální registr – Local Internet Registry

NIR národní registr – National Internet Registry

RIR regionální registr – Regional Internet Registry

TCP protokol řízení přenosu – Transmission Control Protocol

UDP protokol pro nespojové datagramové služby – User Datagram Protocol

SEZNAM PŘÍLOH

A OBSAH PŘILOŽENÉHO CD

- Aplikace (adresář)
  - Databáze (data a postup zprovoznění databáze)
  - Program (adresář obsahující soubory nutné ke spuštění aplikace)
  - Spuštění aplikace (postup pro zprovoznění aplikace)
  - Urceni polohy (soubor dat pro určení přesnosti metody)
- Bakalarska_prace (elektronická verze práce)