VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INFORMAČNÍCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INFORMATION SYSTEMS

CENZURA NA INTERNETU

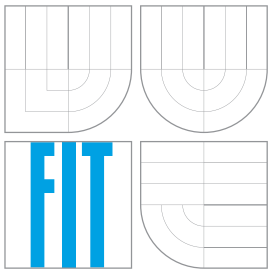
DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

Bc. JAKUB TOMAGA

BRNO 2011

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INFORMAČNÍCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INFORMATION SYSTEMS

CENZURA NA INTERNETU

CENSORSHIP IN THE INTERNET

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

Bc. JAKUB TOMAGA

VEDOUCÍ PRÁCE
SUPERVISOR

Ing. PAVEL OČENÁŠEK, Ph.D.

BRNO 2011

Abstrakt

Diplomová práce se zabývá problematikou cenzury na Internetu. Prezentuje různé technické prostředky na dosažení cenzury, stejně tak jako způsoby ověřování cenzury. Diskutuje o možnostech přístupu na blokovaný obsah a o obejití cenzury obecně. Veškerá problematika je rozebrána v širším kontextu nebo v souvislosti s Čínskou lidovou republikou.

Abstract

This Diploma thesis deals with Internet censorship. Various technical solutions for Internet censorship are presented together with censorship analysis options. Several possibilities for blocked content access and censorship circumvention in general are discussed. The topic is analyzed from the global point of view and is related to the People's Republic of China.

Klíčová slova

cenzura Internetu, obcházení cenzury, dohled, blokování, filtrování, Čína

Keywords

Internet censorship, censorship circumvention, surveillance, blocking, filtering, China

Citace

Jakub Tomaga: Cenzura na Internetu, diplomová práce, Brno, FIT VUT v Brně, 2011

Cenzura na Internetu

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením pana Ing. Pavla Očenáška, Ph.D. Uvedl jsme všechny literární prameny a publikace, ze kterých jsem čerpal.

.....
Jakub Tomaga
23. května 2012

Poděkování

Děkuji vedoucímu diplomové práce Ing. Očenáškovi, Ph.D za metodické vedení, pedagogickou a odbornou pomoc při zpracování mé diplomové práce.

Obsah

Kapitola 1

Úvod

Veľký nárast Internetu v posledných rokoch umožnil sprístupniť ohromnú časť ľudských znalostí z rôznych, mnohokrát nezvyčajných miest ako sú napr. nemocnica v odľahlých horách, ktorákoľvek miestnosť obytných priestorov alebo konferenčná miestnosť.

Na všetkých týchto miestach otvárajú možnosti pripojenia nové možnosti na zlepšenie života ľudí. V prípade zdravotných problémov vo vzdialených a nedostupných oblastiach je možné prostredníctvom Internetu zaslať výsledky testov odbornému lekárovi prakticky kdekoľvek na svete a zachrániť tak životy ľudí včasnou reakciou. Študenti môžu využívať Internetové zdroje ako základ pri riešení školských projektov alebo získavať nových priateľov bez geografických hraníc vďaka sociálnym sieťam. Pracovné výsledky je možné konzultovať s kolegami kdekoľvek na svete a vylepšovať tak aktuálne produkty spoluprácou s odborníkmi, ktorí nie sú fyzicky prítomní v kancelárii.

Internet ale neobsahuje len relevantné a užitočné informácie z hľadiska vzdelávania, pracovných záležitostí a nadväzovania nových kontaktov. Je rovnako prístupný každému, bez ohľadu na úmysly, a nie je možné vždy zabrániť zákernému správaniu sa užívateľov.

Kvôli pozitívam a negatívam, ktoré Internet postupne za roky fungovania prebral z reálneho sveta, sa objavili iniciatívy kontrolovať ako ho ľudia využívajú. Motivácií na takúto kontrolu existuje niekoľko. Medzi hlavné patrí ochrana detí od prístupu k nevhodnému obsahu, znižovanie príjmu nevyžiadanej elektronickej pošty s reklamným obsahom, využívanie pracovných staníc a internetových zdrojov zamestnancami na súkromné účely a prístup k obsahu, ktorý je obmedzovaný a regulovaný právne v danej krajine.

Niektoré tieto snahy dovoľujú samotným užívateľom, aby "kontrolovali"sami seba v podobe využívania rôznych spamových filtrov apod. Ostatné ale obmedzujú spôsoby ako môžu užívatelia Internet využívať a k akému obsahu môžu, resp. nemôžu pristupovať. Vládne blokovanie často končí konfliktom na úrovní nesúhlasu s blokovaním obsahu, ktorý samotní užívatelia považujú za podstatný z hľadiska ich vlastného záujmu [4].

Práve vládnym obmedzovaním prístupu k informáciám na Internete (cenzúrou) sa zaoberá táto diplomová práca. Podrobne popisuje technické pozadie cenzúry a prezentuje možnosti jej obídenia. Vo všetkých týchto aspektoch sa diplomová práca zaoberá situáciou v Čínskej ľudovej republike, ktorá je v tomto smere celosvetovo najďalej, a kde je cenzúra najrozsiahlejšia.

Kapitola 2

Cenzorské prostriedky

Celkové pozadie obmedzovania prístupu k Internetu je komplikované, politické a neustále aktívne spochybňované. Vlády majú často k dispozícii zdroje implementovať preferovanú schému monitorovania a kontroly Internetu, či už je infraštruktúra v súkromných alebo vládnych rukách. Vládne orgány tak majú možnosť priamo alebo nepriamo kontrolovať komunikáciu v bodoch, kde informácie vstupujú do krajiny alebo ju naopak opúšťajú. Rôzne právne úpravy umožňujú monitorovanie samotných užívateľov a obmedzovanie prístupu k nepovolenému obsahu na základe daných pravidiel [18].

V moderných počítačových sieťach ako je Internet sú cenzúra a sledovanie (monitorovanie komunikácie a aktivít ľudí) v praxi často prepojené. Mnoho poskytovateľov Internetu (ISP - *Internet Service Provider*) monitoruje užívateľov za účelom účtovania služieb a ochrany proti spamu (nevyžiadanej pošte). ISP často zaznamenávajú užívateľské mená spolu s IP adresami. Pokiaľ užívatelia sami nevyužívajú nástroje na zvýšenie bezpečnosti a anonymity na Internete je jednoduché na strane ISP zaznamenávať všetky informácie o tokoch, spolu s presným obsahom komunikácie jednotlivých užívateľov.

Takéto sledovanie je predpokladom na technickú cenzúru. Ale na to, aby mohlo dôjsť na strane ISP k cenzúre komunikácie, je potrebné danú komunikáciu prečítať a na základe jej obsahu aplikovať príslušné pravidlá nastavenej politiky. Intuitívnym prístupom pri znižovaní Internetovej cenzúry je skrývanie detailov obsahu komunikácie pred ISP. Zámerné pozmeňovanie komunikácie z dôvodu zníženia zrozumiteľnosti a šifrovanie sú používané vždy pri záujme skryť obsah komunikácie.

Táto kapitola sa zaoberá konkrétnymi spôsobmi cenzúry obsahu a odmietnutia prístupu k citlivým informáciám a predovšetkým vychádza z [18][4].

2.1 Blokovanie na základe URL

Jedným zo spôsobov blokovania prístupu k informáciám na webových stránkach je zabrániť prístupu na základe URL, či už celej alebo niektorej jej časti. Vo väčšine prípadov existuje záujem blokovať konkrétnu doménu úplne kvôli výhradám k obsahu, ktorý sa s danou doménou priamo spája. Jedným z najjednoduchších spôsobov takéhoto blokovania je zablokovať celé doménové meno. Niekedy dochádza k selektívnejšiemu prístupu a blokujú sa len niektoré subdomény, zatiaľ čo zvyšok domény ostáva z pohľadu prístupu nedotknutý. To je napr. prípad Vietnamu, kde vláda blokuje špecifické časti internetových stránok (ako napr. verzie BBC a Rádio Free Asia - Slobodná Ázia, ktoré obsahujú informácie vo vietnamčine), ale nedochádza k blokovaniu obsahu v angličtine [4]. K cenzúre môže dôjsť len

na `http://news.bbc.co.uk`, kým `http://bbc.co.uk` and `http://www.bbc.co.uk` ostanú necenzurované. Podobne je možné filtrovať stránky obsahujúce špecifický typ obsahu, kým prístup k zvyšku stránky ostane povolený. Jeden z prístupov je sledovať zložky na serveri a blokovať prístup napr. ku službe `worldservice`[1], čím dôjde k zablokovaniu správ BBC v cudzom jazyku bez zablokovania celej anglickej verzie. Niekedy je cenzúra aplikovaná na základe názvov stránok alebo vyhľadávaných výrazov, ktoré jednotlivé vyhľadávače umiestňujú priamo do URL.

Filtrovanie obsahu na základe URL je možné uskutočniť lokálne, pomocou špeciálneho softwaru nainštalovaného priamo na fyzickej stanici. Druhým variantom je využitie *proxy serveru*. Konfiguráciou siete je užívateľom zabránený priamy prístup a namiesto toho je každý komunikačný tok presmerovaný na proxy server, ktorý určuje, či HTTP požiadavky majú byť povolené. Ak áno, tak je požiadavka poslaná požadovanému serveru. Keďže týmto spôsobom je možné získať obsah celej požiadavky, je možné filtrovať webové stránky ako na základe názvu, tak aj jej obsahu (z odpovede serveru na danú požiadavku). V prípade blokovanej stránky vráti proxy server vysvetlenie prečo je daná stránka blokovaná, resp. môže predstierať, že stránka neexistuje alebo oznámiť chybu spojenia.

2.2 Filtrovanie na základe DNS a DNS spoofing

Webové prehliadače pri prístupe na webové stránky pomocou URL kontaktujú DNS (*Domain Name System*) server a požadujú preklad URL na IP adresu [9]. Ak je DNS server nakonfigurovaný tak, aby blokoval prístup, tak pri preklade porovnáva doménové mená s hodnotami v *blackliste*, zozname zakázaných URL. V prípade, kedy webový prehliadač žiada preklad zakázanej adresy, DNS server vráti nesprávnu odpoveď, resp. nevráti žiadnu odpoveď. V takom prípade počítač, ktorý o preklad žiadal, nepozná správnu IP adresu a požadovanú službu nemôže využiť. Bez správnej IP adresy nie je možné komunikáciu nadviazať a pokus o spojenie spravidla končí chybovou správou. V prípade webového prehliadača nie je možné kontaktovať správny server a zobraziť požadovanú stránku, pretože tá je z pohľadu užívateľa nedostupná. Chyby tohto typu si užívateľ môže jednoducho vysvetliť ako technický problém na strane poskytovateľa internetovej služby, ktorú sa snaží práve využiť.

Podobne je možné pri DNS preklade nanútiť nesprávnu IP adresu a tým presmerovať užívateľov na nesprávne webové stránky. Táto technika sa volá *DNS spoofing* a je využívaná napr. na presmerovanie komunikácie na neautorizované servery, ktoré umožnia kontrolovať komunikáciu[4]. Tento prístup sa využíva v prípade, kedy cenzúra nie je popieraná a užívatelia tak majú priamo možnosť sledovať rozdiel v požadovanej a zobrazenej stránke.

2.3 Filtrovanie na základe IP adresy

Dáta sú pri prenose cez Internet rozdelené na pakety, ktoré okrem samotných dát obsahujú aj informácie potrebné na prenos, ako sú zdrojová a cieľová IP adresa. Paket po sieti prechádza smerovačmi, ktoré určujú jeho výslednú cestu cez sieť. V prípade blokovania na základe IP adresy môžu smerovače zahadzovať pakety smerované na zakázané IP adresy (porovnávané voči blacklistu) alebo pre ne vrátiť chybové správy. Takéto filtrovanie zablokuje prístup na všetky služby bežiace na danom serveri, napr. webové aj e-mailové služby.

[1] World Service je služba BBC, kde je možné nájsť správy a analýzy v 27 jazykoch. Je dostupná na adrese `http://bbc.co.uk/worldservice`.

Viaceré doménové mená môžu využívať jednu IP adresu a tento spôsob filtrovania zablokuje všetky, a to aj v prípade, ak plán blokovania predpokladal len obmedzenie prístupu k službe len na jednej doméne.

2.4 Filtrovanie na základe kľúčových slov

Pri filtrovaní na základe IP adries môže byť komunikácia blokovaná na základe toho odkiaľ a kam sú pakety smerované v sieti, ale nie na základe informácií, ktoré obsahujú. To predstavuje problém v prípade, keď nie je možné zostaviť kompletný zoznam IP adries obsahujúcich zakázaný obsah alebo v prípade, keď určité IP adresy obsahujú relatívne veľké množstvo povoleného obsahu. V obsahu paketov je možné vyhľadať zakázané kľúčové slová. Smerovače za normálnych okolností nenahliadajú na obsah paketov, a preto sú potrebné dodatočné prvky na dôkladnú analýzu ich obsahu.

Komunikácia, ktorá je identifikovaná ako zakázaná, môže byť ukončená priamo blokovaním paketov alebo nanútením chybových správ, ktoré zabezpečia ukončenie spojenia na oboch stranách komunikácie. Na trhu existujú prvky, ktoré túto funkcionalitu poskytujú. Alternatívne je možné použiť proxy servery tak, ako to bolo opísané vyššie.

2.5 Traffic shaping

Traffic shaping je technika využívaná sieťovými manažérmi na dosiahnutie hladšieho behu počítačovej siete prioritizáciou istého typu paketov. Tento krok umožní uprednostniť určitý typ dát pred ostatnými, doručenie ostatných sa pri veľmi nízkej priorite môže javiť ako pomalé. Priníp je podobný dopravným predpisom, v rámci ktorých sú uprednostňované vozidlá ambulancie a policajné, resp. požiarne zložky. Podobné pravidlá sa uplatňujú aj v prostredí IP sietí, ktoré si vyžadujú malé oneskorenie dát na dosiahnutie optimálnej výkonnosti (ako napríklad technológia VoIP - *Voice over IP*) [15].

Traffic shaping je využiteľný aj na oneskorenie paketov obsahujúcich určitý typ informácie. V prípade záujmu zablokovať prístup k určitej službe, je možné po identifikovaní dát, ktoré s danou službou súvisia, priradiť týmto dátam nízku prioritu a spomaliť tak odozvu služby. To môže v užívateľoch evokovať pocit nespoľahlivej alebo nedostupnej služby a znížiť tak jej obľúbenosť voči ostatným. Táto technika sa niekedy využíva na obmedzovanie používania služieb ako je BitTorrent[2] na strane ISP, aby sa tak znížila miera zdieľania súborov využitím tohto typu sietí.

2.6 Filtrovanie na základe portov

Zaradenie niektorého z portov na blacklist zabráni vo využívaní konkrétnej služby, ako napr. web alebo e-mail. Známe služby majú priradené čísla portov, ktoré priradzuje organizácia IANA[3], ale nie sú nemenné. Priradené hodnoty umožňujú smerovačom odhadnúť typ komunikácie, a preto je blokovanie napr. webovej komunikácie pomerne jednoduché, pretože typickým portom pre tento typ služby je port 80 [13].

[2] Peer-to-peer (P2P) komunikačný protokol na zdieľanie súborov (pojem BitTorrent sa používa aj v súvislosti so softwarom, ktorý túto službu sprístupňuje koncovému užívateľovi).

[3] Internet Assigned Numbers Authority - organizácia zodpovedaná za koordináciu DNS, prideľovanie IP adries a ostatných zdrojov súvisiacich s internetovými protokolmi (`http://www.iana.org`).

Prístup na jednotlivé porty môže byť kontrolovaný sieťovým administrátorom, ktorého zamestnáva spoločnosť, ktorej počítače majú podliehať kontrole, ISP, ktorý poskytuje internetové pripojenie alebo niekým iným, ako napr. vládou, ktorá má záujem o cenzúru. Blokovanie portov môže mať aj iný význam. Pomáha okrem iného zabrániť prijímaniu spamu, zablokovať peer-to-peer siete, instant messaging alebo hranie počítačových hier online.

V prípade zablokovaného portu je všetka komunikácia na danom porte nedostupná. Pri záujme o cenzúru sa často blokujú porty 1080, 3128 a 8080, pretože na týchto portoch bežia vo väčšine prípadov proxy servery. V takomto prípade nie je možné sa priamo pripojiť na proxy server a je potrebné hľadať a následne využiť iné typy obídenia cenzúry. Ak je vo firemnom prostredí povolená komunikácia len na portoch 80 (nezabezpečený prístup na web), 443 (zabezpečený prístup na web), 110, 995 (porty e-mailového protokolu POP3[4]), 25, 465, 587 (porty e-mailového protokolu SMTP[5]), 143, 993 (porty e-mailového protokolu IMAP[6]) a 22 (SSH[7]), je pri záujme o prístup na iné služby tiež potrebné využiť niektorú z techník obídenia takéhoto typu obmedzení.

2.7 Odstavenie Internetu z prevádzky

Odstavenie Internetu (*Internet shutdown*) je extrémnym prípadom cenzúry, ku ktorému sa uchyľujú vlády v prípade citlivých politických a sociálnych udalostí. Avšak kompletné odstavenie siete (napr. od domácich, ale aj zahraničných sietí) vyžaduje rozsiahly zásah, keďže je potrebné odstaviť protokoly, komunikáciu medzi jednotlivými ISP a medzi užívateľmi. Krajiny odstavujú Internet pri pokusoch o potlačenie politických nepokojov. Takéto odstavenie následne trvá od niekoľkých hodín, po niekoľko týždňov. Užívatelia sa však môžu stále pripojiť pomocou zahraničných ISP, mobilných alebo satelitných spojení.

Odstavenie medzinárodného spojenia neznamená zablokovnie spojenia medzi lokálnymi ISP, resp. komunikácie medzi užívateľmi v rámci jedného ISP. Vyžadovalo by to dodatočné kroky na úplné izolovanie užívateľov od lokálnych sietí. Z tohto dôvodu sú takéto kroky sťažené v krajinách s viacerými ISP.

[4]Post Office Protocol - viac informácií v RFC 1939 na `http://www.ietf.org`
[5]Simple Mail Transfer Protocol - viac informácií v RFC 2821 na `http://www.ietf.org`
[6]Internet Message Access Protocol - viac informácií v RFC 3501 na `http://www.ietf.org`
[7]The Secure Shell - viac informácií v RFC 4251, 4522 4523 na `http://www.ietf.org`

Kapitola 3

Analýza a overovanie cenzúry

V dnešnej dobe je analýza cenzúry Internetu zjednodušená vďaka existencii mnohých projektov, ktorých primárnym cieľom je sprístupniť Internet z pohľadu užívateľa, ktorého prístup na Internet podlieha cenzúre každému, kto má záujem o nahliadnutie na obmedzené fungovanie tejto globálnej siete. Hlavnou motiváciou pri vzniku týchto projektov je snaha o transparentnú cenzúru a všeobecné zvyšovanie povedomia o použitých praktikách v jednotlivých krajinách.

Táto kapitola predstavuje niektoré projekty, ktoré sú zaujímavé z hľadiska cenzúry. Rovnako informuje o projektoch, ktorých primárny účel je iný, ale získané informácie je možné po interpretácii využiť na dokreslenie situácie. Okrem všeobnecných projektov sú uvedné aj niektoré špecializované na Čínu.

3.1 OpenNet Initiative

OpenNet Initiative[1] je združenie vytvorené v spolupráci troch inštitúcií: University of Toronto, Berkman Center for Internet & Society na Harvard University a SecDev Group v Ottawe. Cieľom združenia je skúmať, odhaľovať a analyzovať praktiky filtrovania Internetu a dohľadu nad jeho užívateľmi dôveryhodným a nezaujatým spôsobom. Hlavným úmyslom je odkrývať prípadné nástrahy a neúmyselné dôsledky cenzúry a pomáhať tak lepšie informovať verejnosť a právne subjekty. OpenNet Initiative využíva unikátny multidisciplinárny prístup, ktorý zahŕňa [5]:

- Vývoj a nasadenie technických prostriedkov a kľúčových metodológií na štúdium filtrovania Internetu a dohľad.
- Prepojenie siete lokálnych advokátov a výskumníkov.
- Pokročilé štúdie, ktoré skúmajú dôsledky súčasných a budúcich trendov a smerov vo filtrovaní Internetu a dohľadu nad jeho užívateľmi a ich dosah na domáce a zahraničné právo a vládne režimy.

OpenNet Initiative pravidelne zverejňuje výsledky analýzy v konkrétnych krajinách a regiónoch a poukazuje tak na všadeprítomnú cenzúru Internetu. Témou výskumu nie sú len nasadené technické softwarové prostriedky, ale aj blokovaný obsah (politika, pornografia a iné) [6].

[1] Doplňujúce informácie o neziskovom združení OpenNet Intiative je možné nájsť na `http://opennet.net/`.

3.2 WatchMouse

Produkty WatchMouse[2] testujú správanie a dostupnosť webových stránok, služieb a aplikácií využitím vlastnej infraštruktúry, ktorá zahŕňa 62 monitorovacích staníc po celom svete a sieť kontrolných uzlov v 26 krajinách. WatchMouse vykonáva kontroly z externej perspektívy za účelom napodobniť využitie služieb v reálnom čase a pomocou transakčných monitorovacích skriptov identifikuje rôzne dôvody nedostupnosti, od pomalej odozvy stránok, po monitorovanie chovania sa prihlasovacích stránok a nákupných košíkov [7].

Primárnym účelom produktov WatchMouse je umožniť overenie celkovej dostupnosti webových služieb užívateľom v rôznych krajinách. Preto je možné ich využiť pri overovaní cenzúry v jednotlivých regiónoch. Výsledkom monitorovania sú vo väčšine prípadov informácie o stave dostupnosti v konkrétnej lokalite. Interpretáciou výsledkov je možné získať prvotné povedomie o možnej cenzúre pri nedostupnosti služby vo vymedzených oblastiach.

3.3 Projekt greatfirewallofchina.org

Tento projekt je zameraný na testovanie blokovaných webových stránok v Číne. Umožňuje nahliadnuť na Internet z perspektívy čínskeho užívateľa a sledovať či sú požiadavky na konkrétne stránky blokované.

Za projektom stojí nezisková skupina kreatívnych ľudí ako sú web dizajnéri, režiséri dokumentárnych filmov a žurnalisti, ktorí sa snažia o transparentný systém Internetovej cenzúry. Hlavným cieľom projektu je vybudovať komunitu, ktorá pomôže zviditeľniť Internetovú cenzúru. Počas behu sa vytvárajú záznamy o tom koľko, resp. koľkokrát sú jednotlivé stránky cenzurované.

Technické pozadie projektu je jednoduché. Všetky požiadavky na zobrazenie obsahu stránky pomocou URL sú presmerované na webový server v Číne, kde beží podporná webová služba projektu. Konkrétne HTTP požiadavky sa na webový server posielajú priamo z Číny a výsledky sú následne poslané späť. Testovanie ja založené na jednom serveri v jednej lokalite. V prípade problémov je dostupných niekoľko záložných serverov rozmiestnených v rôznych lokalitách. Samotné výsledky sú ale v danom momentne dostupné len z jedného z nich [3].

3.4 Chinese Firewall Checker

Chinese Firewall Checker[3] je produkt, ktorý umožňuje veľmi jednoducho overiť dostupnosť webových stránok z piatich rôznych lokalít v Číne. Zaujímavosťou tejto služby je rebríček najviac overovaných stránok [2]:

1. `http://www.facebook.com`
2. `http://www.youtube.com`
3. `http://www.google.com`

[2] Detaily projektu na `www.watchmouse.com`.

[3] Otestovať službu Cinese Firewall Checker je možné na adrese `http://www.bestvpnservice.com/` v sekcii *China Firewall Test*.

3.5 China Channel

China Channel[4] je doplnok do webového prehliadača Firefox, ktorý umožňuje užívateľom mimo Čínu surfovať po Internete tak, ako keby boli fyzicky prítomní v Číne. Doplnok je založený na ďalšom doplnku do Firefoxu, a to konkrétne Switch Proxy. Cieľom doplnku China Channel je odstrániť technickú bariéru pri pokusoch o surfovanie čínskeho Internetu [1].

Po inštalácii doplnku je možné sa pomocou pridaného prepínača v paneli nástrojov prepnúť do proxy režimu a následne využívať Internet obvyklým spôsobom. Nevýhodou doplnku je fakt, že je kompatibilný len s verziou Firefox 3.0 a staršími.

[4] Ďalšie informácie o doplnku a celom projekte je možné dohľadať na `http://chinachannel.fffff.at/`.

Kapitola 4

Spôsoby obídenia cenzúry

Obídenie cenzúry a zvýšenie bezpečnosti prehliadania dát na Internete spolu úzko súvisia. Úroveň potrebnej bezpečnosti závisí na vykonávaných aktivitách a na dôsledkoch týchto aktivít. Základné bezpečnostné praktiky by mali byť aplikované bez ohľadu na to, či sa niekto cíti byť ohrozený vládnym dohľadom alebo nie. Niektoré spôsoby vyžadujú viac úsilia, ale na obídenie cenzúry sú niekedy nevyhnutné [18][4].

4.1 Základné spôsoby

Existuje niekoľko techník ako prekonať blokovanie Internetu. Ak je cieľom jednoducho len získať prístup k webovým stránkam alebo internetovým službám, ktoré sú nedostupné len z istej lokality a v danom momente nie je podstatné, či sú pokusy o obídenie cenzúry detekovateľné, je možné využiť jednu z nasledujúcich techník.

4.1.1 Využitie HTTPS

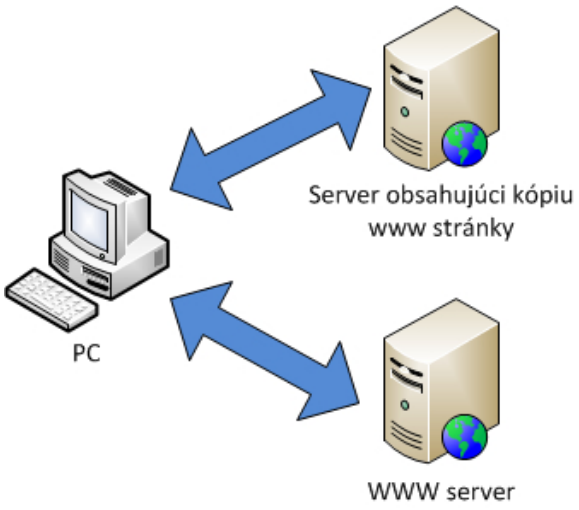
HTTPS je zabezpečená verzia protokolu HTTP využívaná na prístup k webovým stránkam [8]. V niektorých krajinách je možné využitím HTTPS pristúpiť na zablokované stránky nahradením protokolu v URL z `http://` na `https://`. Pred využitím ostatných metód na obídenie cenzúry je doplnenie *"s"* za http prvým krokom. V prípade úspechu okrem prístupu na blokovanú stránku získame zašifrovanie komunikácie medzi nami a vzdialeným serverom.

4.1.2 Využitie alternatívnych doménových mien a URL

Jeden z najrozšírenejších spôsobov na cenzurovanie webových stránok je blokovanie prístupu k ich doménovému menu, napr. `http://news.bbc.co.uk`. Avšak stránky sú často prístupné aj na iných URL, ako napríklad `http://newsrss.bbc.co.uk`. Ak je jedna z domén blokovaná, je niekedy možné nájsť jej obsah na inej doméne.

Takisto je možné skúsiť pristúpiť na špeciálne verzie, ktoré niektoré stránky vytvárajú pre zariadenia typu *smartphone*. Vo väčšine prípadov majú rovnakú URL s pridaným "`m`"alebo "`mobile`"na začiatku, napr.:

- `http://m.google.com/mail` (Gmail)
- `http://mobile.twitter.com`



Obrázek 4.1: Prístup na webové stránky pomocou služieb tretích strán.

- `http://m.facebook.com` alebo `http://touch.facebook.com`
- `http://m.flickr.com`

4.1.3 Využitie webových stránok tretích strán

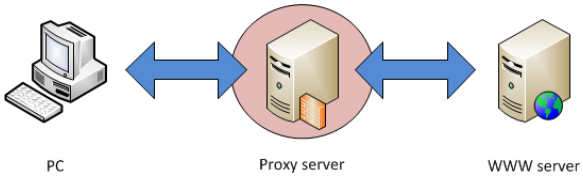
Existuje viacero možností ako pristúpiť k obsahu nejakej webovej stránky pomocou webovej stránky tretej strany. Takto je možné sa vyhnúť priamemu spojeniu s webovou stránkou a obísť blokovaný prístup. Príkladom sú služby ako *Google Cache*, *RSS agregátory*, prekladače webových stránok (*Google Translate*, *Bing Translator*, ...) alebo webové archívy (*Wayback Engine* - `http://www.archive.org/web/web.php`). V týchto prípadoch je často potrebné počítať s faktom, že zobrazené dáta nemusia byť vždy aktuálne.

4.1.4 Využitie e-mailových služieb

E-mailové a webmailové služby je možné využiť na zdieľanie dokumentov alebo na prehliadanie webových stránok. Je možné si obsah webovej stránky poslať pomocou e-mailovej pošty či už ako obsah v tele správy alebo ako prílohu. Takýto prístup na web môže byť pomerne zložitý, zdĺhavý a nepraktický, pretože je potrebné žiadať o prístup na stránky postupne a následne čakať na odpoveď. Tento spôsob môže byť ale účinný v prípade, kedy sa využíva zabezpečené spojenie pri prístupe na webmailovú službu.

4.1.5 Výhody a riziká

Tieto jednoduché techniky sú rýchle a jednoduché na použitie a je možné skúsiť ich účinnosť vynaložením minimálneho úsilia. Mnohé z nich budú fungovať prinajmenšom v niektorých časovo obmedzených situáciách. Avšak je jednoduché ich detekovať a zablokovať. Keďže nie všetky uvedené metódy využívajú šifrované spojenie alebo nijako inak neskrývajú obsah komunikácie, sú zraniteľné voči blokovaniu na základe kľúčových slov.



Obrázek 4.2: Pripojenie pomocou proxy serveru v Číne.

4.2 Proxy servery

Proxy servery umožňujú získať obsah webovej stránky alebo vyžiť iné Internetové služby aj v prípade, že je ich zdroj blokovaný v danej lokalite. Existuje viacero typov proxy serverov, zahŕňajúc napr.:

- Webové proxy, ktoré vyžadujú len znalosť URL adresy webovej stránky. URL webového proxy môže vyzerať nasledovne: `http://example.com/cgi-bin/nph-proxy.cgi`.
- HTTP proxy, ktoré vyžadujú externý software na modifikáciu nastavení webového prehliadača. Fungujú len na webový obsah. Informácie o proxy serveri môžu mať formát `proxy.example.com:3128` alebo `192.168.0.1:8080`.
- Proxy servery typu *SOCKS*, ktoré taktiež vyžadujú externý software na modifikáciu nastavení webového prehliadača, SOCKS proxy fungujú s mnohými Internetovými aplikáciami, vrátane e-mailov a instant messagingu. Informácie o SOCKS proxy vyzerajú podobne ako tie o HTTP proxy.

Webové proxy je možné prirovnať k webovému prehliadaču, ktorý je zabudovaný do webovej stránky. Typicky obsahuje formulár, do ktorého je možné zadať URL stránky, na ktorú chceme pristúpiť. Následne zobrazí obsah stránky bez nutnosti priameho spojenia.

Využitie webového proxy serveru nevyžaduje inštaláciu prídavného softwaru alebo zmenu nastavení, čo znamená, že je možné využiť proxy server z ktoréhokoľvek počítača, vrátane tých v internetových kaviarňach, a zvýšiť tak úroveň anonymity. Využitie proxy serveru neobmedzuje prácu so stránkou a je možné využiť odkazy a navigačné prvky bez strany "anonymného"spojenia. Proxy server prepíše všetky odkazy a obsahy formulárov tak, aby sa predišlo priamemu spojeniu. To môže byť v dnešnej dobe pomerne komplikovaná úloha z dôvodu narastajúcej zložitosti jednotlivých webových služieb. Preto je niekedy možné stratiť proxy spojenie kvôli prílišnej komplexnosti prehliadanej stránky (proxy server si neporadí s prebiehajúcou komunikáciou). Jednoduchým indikátorom tohto stavu je fakt, že na stránke vo webovom prehliadači nie je viditeľný panel s formulárom na zadávania URL cez proxy server.

4.3 VPN

VPN (*Virtual Private Network*) šifruje a tuneluje všetku internetovú komunikáciu medzi dvoma počítačmi [19]. Vzdialený počítač môže patriť poskytovateľovi komerčnej VPN služby, vašej organizácii alebo dôveryhodnému obchodnému partnerovi.

Keďže VPN tuneluje všetku internetovú komunikáciu, vrátane e-mailov, instant mesagingu, VoIP a všetkých ostatných internetových služieb dopĺňajúcich webové služby, stáva

sa tak všetko, čo prechádza tunelom nečitateľné pre každého, kto by komunikáciu mohol potenciálne odchytávať (ak nepredpokladáme útoky na známe šifrovacie algoritmy apod.).

Ak druhý koniec tunela končí mimo lokalitu, kde dochádza k filtrovaniu Internetu, je VPN vhodným nástrojom na obídenie cenzúry, keďže dáta sú šifrované a nie je možné nahliadať na obsah komunikácie. Všetka komunikácia vyzerá navonok rovnako a bez jasného významu.

Mnohé spoločnosti využívajú VPN ako prostriedok pre zamestnancov, ktorí potrebujú pristupovať k citlivým firemným informáciám z domu alebo zo vzdialených lokalít počas služobných ciest. Preto je veľmi nepravdepodobné, že by bolo v krajinách so záujmom o cenzúru blokované pripojenie pomocou VPN.

Dáta sú šifrované počas prenosu cez tunel. Od okamihu doručenia na druhú stranu putujú ďalej nešifrované až k zdroju požadovanej internetovej služby. To znamená, že tieto dáta môžu podliehať cenzúre v cieľovej lokalite VPN tunela.

4.4 TOR

TOR (*The Onion Router*) je veľmi sofistikovaná sieť proxy serverov. Pri využití TORu na prístup k webovým službám, je komunikácia náhodne smerovaná cez sieť nezávislých proxy serverov. Komunikácia medzi TOR servermi je šifrovaná a každý článok spojenia pozná len IP adresu dvoch uzlov - IP adresu predchádzajúceho uzlu a IP adresu nasledujúceho uzlu.

Cieľom je dosiahnuť nemožnosť spájať obsah komunikácie s jej účelom, zdrojovým a cieľovým počítačom. TOR výrazne sťažuje situáciu pre ISP pri zisťovaní cieľovej webovej stránky alebo pri zisťovaní obsahu komunikácie (prinajmenšom zistiť vašu IP adresu).

4.5 SSH Tunnelling

SSH (*Secure Shell*) je štandardný protokol na šifrovanie spojenia medzi počítačom a webovým serverom. Šifrovanie zabraňuje nahliadnutiu na prenášané dáta alebo ich modifikáciu. SSH je možné využiť v kombinácii s mnohými aplikáciami, kde je typické bezpečné prihlásenia sa na server [8].

SSH je výhodné v prípade cenzúry na Internete, pretože poskytuje šifrovaný tunel a funguje ako všeobecný proxy klient. Opäť platí, že je SSH pomerne široko využívané (napr. systémovými administrátormi na správu služieb) a je preto nepravdepodobné, že by ktorákoľvek krajina so snahou o Internetovú cenzúru kompletne blokovala komunikáciu pomocou SSH.

Na využitie SSH je potrebné mať účet na fyzickom stroji, kde beží server, vo všeobecnosti je to Unixový alebo Linuxový systém. Pre potreby obídenia cenzúry je potrebné, aby tento server mal nefiltrovaný prístup k Internetu.

Kapitola 5

Prístup do Číny

Pri overovaní cenzúry je potrebné mať fyzický prístup do krajiny, v ktorej chceme analýzu vykonať. Táto kapitola sa venuje analýze možností prístupu do Číny. Prezentuje jednotlivé varianty, ktoré boli vyskúšané pri pokusoch rozbehnúť testovacie prostredie na overovanie cenzúry priamo z čínskeho prostredia. Mnohé cesty sa ukázali ako nepoužiteľné, či už z technického alebo administratívneho hľadiska. V nasledujúcom texte je možné dohľadať popis problémov, s ktorými bolo potrebné sa vysporiadať.

5.1 Web hosting

Prevádzkovanie webovej stránky v Číne je komplikované. Webové stránky sú často zablokované a odblokované bez rozpoznateľného dôvodu. V prípade, že je stránka v Číne zablokovaná, je jej prevádzkovanie výrazne zložité, ak nie nemožné. Vybrať správny webhosting je často krát veľmi náročné.

Pri výbere plánu pre webhosting je potrebné zvažovať fyzické umiestnenie prevádzkovateľa. Hlavnou výhodou umiestnenia priamo v Číne sa môže zdať rýchlejší prístup pre kohokoľvek priamo z Číny, bez ohľadu na to, že existuje pravdepodobnosť zablokovania stránky kvôli porušeniu pravidiel súvisiacich s aplikovaním Internetovej cenzúry.

Prevádzkovať webovú stránku v Číne znamená riešiť niekoľko problémov. Po prvé je potrebné brať na vedomie, že obsah webovej stránky podlieha cenzúre. Sloboda vyjadrovania sa na webových stránkach je obmedzená. Ignorovanie pravidiel môže znamenať stratu webhostingu. Po druhé, vybaviť webhosting v Číne je náročné z administratívneho hľadiska. Väčšina komunikácie, či už v papierovej podobe alebo online je v čínštine. Po tretie, prístup na webové stránky z územia mimo Čínu môže byť veľmi pomalý.

Možným kompromisom medzi webhostingom fyzicky lokalizovaným v Číne a webhostingom v zahraničí, napr. v USA je prevádzkovať webovú stránku v Hong Kongu. Webhosting v Hong Kongu je typicky oveľa drahší ako v iných lokalitách.

Najlacnejšou možnosťou pre webhosting je využiť nejakú z amerických spoločností. Rýchlosť pripojenia je dostatočná, doplnkové služby nadštandardné.

Z hľadiska overovania cenzúry je ale nemožné prevádzkovať webhosting mimo Čínu. Cieľom diplomovej práce je overovať sieťové pripojenie z Číny smerom von a nie z lokalít mimo Čínu smerom do krajiny. Preto je potrebné nájsť webhosting lokalizovaný v Číne. Hongkong, ako jedna z dvoch osobitných administratívnych oblastí Číny a od roku 1997 od Číny nezávislý štát spravovaný spojeným kráľovstvom, nie je vhodným miestom pre analýzu čínskej cenzúry, pretože Internet tam nepodlieha rovnakým restrikciám[18]. V rámci

testovania cenzúry je potrebné pristupovať na webové stránky, ktoré sú v Číne blokované. Spúšťaním rôznych testov z územia mimo Čínu nebude možné sledovať komunikáciu, ktorá by prechádzala cez čínskej firewally a výsledky nebudú použiteľné.

Možnosti webhostingu výrazne záležia na dostupnosti jednotlivých technológií. Základnou požiadavkou pri výbere bola dostupnosť PHP, pomocou ktorého by bolo možné vykonávať základné testy pomocou špecializovaných knižníc. Jednou z možností je využitie technológie *BSD sockets* na komunikáciu so službami bežiacimi v Číne (HTTP, SMTP, ...) a sledovať tak chovanie sieťového pripojenia pri rôznych testoch. Ďalšou požiadavkou bola dostupnosť špecializovaných knižníc na prácu s aplikačnými protokolmi. Konkrétne knižnica *libcurl* umožňuje jednoduchú prácu s protokolom HTTP. Takto je možné overovať prístup na webové stránky a sledovať priebeh HTTP komunikácie pri aplikovaní cenzúry jednoducho pomocou PHP skriptu.

V rámci komplexnejších testov sieťového pripojenia bolo potrebné získať prístup k sofistikovanejším nástrojom ako je `ping` a `traceroute`. Prostredníctvom týchto utilít je možné priamo sledovať dostupnosť jednotlivých serverov a prípadne identifikovať umiestnenie čínskych sieťových firewallov. V prípade tohto typu testov je zároveň potrebný prístup ku konzole, resp. k obdobnému prostrediu, z ktorého je možné a tieto príkazy s vhodnými parametrami spúšťať. Webhosting má často len webové rozhranie bez priameho prístupu ku konzole, čo komplikuje vykonávanie pokročilejších testov. Jednou z možností je teda využiť podporované knižnice na strane serveru, tak ako to bolo diskutované vyššie.

Nasledujúce podkapitoly zhrnú možnosti voľne dostupných a platených služieb v Číne, ktoré boli vyskúšané pri pokusoch rozbehnúť testovacie prostredie.

5.1.1 Free web hosting

Voľne dostupný webhosting predstavuje spôsob ako získať webový priestor bez investovania finančných zdrojov. Zároveň má veľa obmedzení, medzi ktoré patria veľkosť úložného priestoru, šírka prenosového pásma limitovaná na mesiac a iné. Poskytovatelia často pridávajú rôzny druh reklamy priamo do webových stránok. Najvýznamnejšia nevýhoda je chýbajúca garancia dostupnosti. Niektoré alternatívy ponúkajú prístup pomocou FTP na prenos súborov, resp. podporu skriptovacích jazykov. Práve tieto služby sú z hľadiska možností overovania cenzúry najzaujímavejšie.

Webhosting s okamžitou aktiváciou poskytuje veľmi malú diskovú kapacitu. Čím rýchlejšie a bezproblémovejšie je získať prístup, tým slabšie služby sa s aktivovaných účtom spájajú. S tým súvisí obmedzené množstvo obrázkov, ktoré je možné na webovú stránku umiestniť.

Počas riešenia diplomovej práce boli vyskúšané viaceré voľne dostupné alternatívy v Číne. Žiadna z nich sa v konečnom dôsledku neosvedčila.

Prvý využitý voľne dostupný webhosting bol na adrese `http://www.5944.net/`. Garantovaná podpora PHP a rýchle pripojenie predstavovali vhodné predpoklady na testovacie prostredie. Práve garancia rôznych služieb je najväčší problém ponúkaných možností. Pripojenie často končilo vypršaním času spojenia a podpora PHP chýbala. V konečnom dôsledku nebolo možné uvedený webhosting vôbec využívať k zamýšľaným zámerom. Ďalším problémom spoločným pre mnohé čínske webové služby je fakt, že stránka poskytovateľa je dostupná iba v čínštine. Služby ako *Google Translate* (`http://translate.google.com`) je možné využiť pri základnej orientácii sa po stránkach, ale v mnohých prípadoch sú texty v obrázkoch, čo prakticky znemožňuje využívať tento typ prekladačov pri registrácii webového účtu.

Okrem prvého pokusu sa následne nepodarilo zohnať vhodnú ďalšiu voľne dostupnú variantu umiestnenú v Číne. Ďalším krokom bolo skúsiť webhosting v Hong Kongu, ktorý by mohol využitím proxy serverov v Číne poslúžiť ako krajná alternatíva. Na adrese `http://host-hongkong.net/` je prevádzkovaný webhosting, ktorý je možné registrovať a zadarmo využívať za určitých podmienok. Podmienky sú ale striktne dané a účelom aplikácie na overovanie cenzúry ich nebolo možné splniť. Webové stránky by museli mať náboženský, resp. charitatívny charakter. Ďalšou možnosťou bolo koncipovať webovú stránku ako osobnú stránku alebo tzv. *fan page*. V kombinácii s nutnosťou využívania proxy serverov som túto možnosť ďalej neuvažoval a hľadal rozdielny prístup.

5.1.2 Platený web hosting

Ďalším krokom bolo investovať finančné prostriedky a využiť platený webhosting, kde je väčšia pravdepodobnosť poskytnutia garantovaných služieb. S pridanými službami sa patrične zvyšuje cena. Okrem cenového zaťaženia je výrazným obmedzením Licencia ICP (*Internet Control Provider*). Ide o povolenie čínskeho Ministerstva priemyslu a informačných technológií prevádzkovať webové stránky v rámci Číny. Podľa zákona, všetky webové stránky prevádzkované na území Číny musia byť licencované a ISP majú právo blokovať stránku v prípade, ak pre uvedenú doménu nie je evidovaná licencia ICP.

Získať licenciu ICP neprináša len zdĺhavé administratívne úkony s čínskymi úradmi. Na získanie licencie je potrebné mať obchodnú licenciu na podnikanie v Číne a v mene spoločnosti musí s poskytovateľom webhostingu komunikovať čínsky občan. Tieto podmienky nebolo možné splniť a preto ani táto cesta neviedla k získaniu testovacieho prostredie priamo v Číne[1].

Viacero platených webhostingov je cenovo nedostupných. Z finančne rozumných alternatív najlákavejšie vyzeral webhosting na adrese `http://www.sinohosting.net/`. Spoločnosť ponúkala dva druhy plánov. Rozdiely boli v umiestnení serverov, na ktorých by stránky bežali. Prvá možnosť zahŕňala čínsky webhosting lokalizovaný v mestách Shanghai, Jiangsu a Guangdong. Druhá možnosť predstavovala medzinárodnú variantu s umiestnením serverov v Hong Kongu, Singapure alebo v USA. Medzinárodný webhosting ale nie je z dôvodov uvedených vyššie zaujímavý. Čínska alternatíva má obmedzenie, ktoré nebolo možné prekonať. Na stránkach spoločnosti je uvedená poznámka: "Z dôvodu nových regulácii v Číne je využitie tejto služby obmedzené len na klientov s platnou ICP licenciou, resp. so spoločnosťou registrovanou v Číne, ktorá sa môže u licenciu uchádzať."[2].

5.2 Voľne dostupné e-mailové služby

Ďalšou možnosťou overovania internetovej cenzúry v Číne je zamerať sa na aplikačné protokoly, konkrétne e-mailové protokoly ako POP3, IMAP, SMTP. Mnohé webhostingové služby poskytujú k webovému priestoru rovnako aj e-mailové účty s podporou jednotlivých protokolov. Alternatívne je dostupný webmail. Vzhľadom na to, že webhosting sa v Číne nepodarilo vybaviť, ostáva alternatívou využitie produktov poskytovateľov voľne dostupných e-mailových služieb. Na adrese `http://www.fepg.net/asiachina.html` je uvedený dlhý zoznam poskytovateľov e-mailových služieb. Problematické sú ale podmienky získania potrebného prístupu. Služby sú dostupné čínskym občanom, čo je rovnako neprekonateľná podmienka ako získanie ICP licencie v prípade plateného webového priestoru.

[1] Viac informácií o licencii ICP je možné dohľadať na `http://www.sinohosting.net/icp.php`
[2] Viac informácií na `http://www.sinohosting.net/china-hosting.php`

5.3 Alternatívne spôsoby prístupu

Po neúspechu so zriadením webhostingu, resp. samostatného e-mailového účtu bolo potrebné získať prístup alternatívnym spôsobom. Analýzou rôznych poskytovaných komerčných produktov, ako sú dedikované servery apod., sa podarilo nájsť najrozumnejšie riešenie z hľadiska pomeru cena/výkon. Konkrétne ide o službu, v rámci ktorej je možné získať prístup k vzdialenej linuxovej stanici lokalizovanej v Číne pomocou SSH. Marketingovo je produkt prezentovaný ako vhodný na overovanie dostupnosti vlastnej webovej stránky, testovanie odozvy služieb z čínskeho prostredia a všeobecne na uskutočnenie prieskumu trhu v Číne.

Využívanie daného produktu je podmienené mesačnými platbami poskytovateľovi, ktorým je kanadská spoločnosť Compevo (`http://compevo.com/`). Čiastka, ktorú je potrebné mesačne uhradiť je $14,95. Po zaplatení prvej mesačnej periódy a niekoľkodňovom schvaľovacom procese bolo možné začať zisťovať možnosti testovacieho prostredia. Stanica je však výrazne obmedzená a neposkytuje celú radu potrebných nástrojov a utilít. Celkovo obsahuje 136 príkazov, z nich je na praktické testovanie možné použiť príkaz `host` na DNS preklady a príkaz `wget` na prístup na webové stránky. Z hľadiska informácií dostupných k danému produktu na stránkach spoločnosti je to nevyhnutné minimum a na rozsiahlejšie testy bolo potrebné stanicu rozšíriť o ďalšie pridané možnosti.

Jedným zo spôsobov ako rozšíriť stanicu bolo preložiť potrebné nástroje pre cieľovú architektúru a prekopírovať výsledné binárne súbory. Na to bolo nevyhnutné poznať cieľovú architektúru. V tomto smere je stanica obmedzená natoľko, že neobsahuje žiadny nástroj, pomocou ktorého by to bolo možné zistiť. Nenachádza sa tam ani príkaz `uname`, `file`, resp. v rámci prideleného užívateľského účtu nebolo možné pristupovať k súboru `\proc\cpuinfo`. Podarilo sa ale skopírovať existujúce binárne súbory z čínskej stanice pomocou `scp` a následne ich analyzovať pomocou príkazu `file`, ktorý slúži na analýzu súborov. Na nasledujúcom výpise je možné vidieť výsledok analýzy binárneho súboru aplikácie `echo`.

```
[jakub@localhost ~]$ file /bin/echo
/bin/echo: ELF 32-bit LSB executable, Intel 80386, version 1 (SYSV),
dynamically linked (uses shared libs), for GNU/Linux 2.6.18, stripped
```

Z uvedeného výpisu je možné zistiť cieľovú architektúru a jadro, pre ktoré bola aplikácia `echo` preložená. V tomto smere dopadla analýza pozitívne, pretože informáciám odpovedala konfigurácia súkromného notebooku, čo umožnilo jednoducho preložiť potrebné nástroje a staticky ich zlinkovať kvôli nedostupnosti potrebných knižníc na cieľovej stanici.

Na rozšírenie analýzy boli preložené utility ako `dig`, `nslookup`, `telnet`, `ping` a `traceroute`. Problematické sú posledné dve utility. Obe vyžadujú administrátorské práva z dôvodu manipulácie s paketmi priamo na sieti. Namiesto `traceroute` je možné použiť utilitu `tracepath`, ktorá administrátorské práva nevyžaduje, pretože využíva len protokol UDP a sleduje dostupnosť bez modifikácie ICMP komunikácie.

Utility `ping` alebo `traceroute` by bolo spúšťať aj bez administrátorských práv v prípade, že by sa poradilo nastaviť SUID bit. SUID, Set User-ID, predstavuje koncept systému UNIX, ktorý umožňuje nastaviť pokročilé práca súborom, ktoré dovoľujú užívateľovi spúšťať skript alebo binárny súbor ako ich vlastník (zvyčajne `root`)[10]. To je veľmi výhodné napríklad v situácii, kedy potrebujeme otvoriť ICMP socket pod iným ako `root` užívateľom. Pomocou príkazu `chmod u+s` je možné nastaviť SUID bit. Nastavenie je potrebné vykonať pod účtom vlastníka súboru. Vzhľadom na to, že potrebujeme, aby vlastníkom súboru bol užívateľ `root`, je táto operácia opäť nerealizovateľná na stanici v Číne. Možným riešením

je využitie nejakého známeho exploitu (software, skupina dát, sekvencia príkazov, ktoré sa snažia využiť chybu alebo zraniteľnosť za účelom spôsobenia nechceného, resp. neočakávaného správania sa softwaru [19]) na získanie administrátorských práv. Vzhľadom na to, že spomínaný SSH účet je jediným funkčným testovacím prostredím, ktoré sa podarilo získať, bol by takýto krok príliš riskantný. Preto na overovanie cenzúry nebude možné využiť `ping` a `traceroute`.

Po príprave prostredia je potrebné zabezpečiť prenos získaných výsledkov. Medzi dostupnými príkazmi priamo na stanici v Číne je príkaz `scp`. Aby bolo možné prenášať súbory z Číny automaticky pomocou skriptu (bez manuálneho zadávania hesla), musíme využiť autentizáciu pomocou verejného SSH kľúča a prekopírovať ho na vzdialenú stanicu. Je potrebné vykonať nasledujúce kroky [10]:

1. Overiť dostupnosť `openSSH` na oboch staniciach pomocou `ssh -V`
2. Vygenerovať kľúčový pár na lokálnej stanici využitím `ssh-keygen`
3. V prípade, že na vzdialenej stanici neexistuje zložka `~/.ssh/`, vytvoriť ju.
4. Nainštalovať (preniesť) verejný kľúč na vzdialenú stanicu (napr. pomocou `scp`) do súboru `~/.ssh/authorized_keys`
5. Nastaviť potrebné práva na uvedenú zložku a súbor:

   ```
   $ chmod 755 ~/.ssh
   $ chmod 644 ~/.ssh/authorized_keys
   ```

6. Prihlásiť sa bez nutnosti zadávať heslo

Na stanici v Číne sa nepodarilo rozbehnúť prenos súborov cez `scp` bez nutnosti zadávať heslo. Pravdepodobným dôvodom je konfigurácia `ssh`, v rámci ktorej je možné vypnúť podporu autentizácie pomocou verejného kľúča. Konfiguračný súbor nie je na stanici prístupný pod prideleným užívateľským účtom, preto túto hypotézu nie je možné prakticky overiť. Zvolený spôsob komunikácie a prenášania získaných výsledkov je prezentovaný v kapitole zaoberajúcej sa implementáciou.

Kapitola 6

Implementácia

V tejto kapitole je opísaná implementácia celého systému na overovanie cenzúry v Číne. V úvodnej časti sú predstavené nástroje pre vývoj výslednej aplikácie. Každej časti je venovaná samostatná podkapitola, ktorá predstaví implementačné riešenie.

6.1 Zvolené nástroje a technológie na vývoj

Bash

Bash je unixový (Linux/Unix/BSD) príkazový *shell* interpret naprogramovaný v rámci projektu GNU. Názov je skratkou názvu *Bourne again shell* - je založený na *Bourne Shell* (bsh), čo bol najpoužívanejší unixový shell. Bash bol taktiež portovaný na operačný systém Microsoft Windows projektom Cygwin. Bash sa snaží o širokú kompatibilitu. Zaujímavou vlastnosťou je automatické rozpoznanie, pod ktorým menom bol spustený a prispôsobený danému typu shellu. Cieľom tvorcov je dosiahnuť 100% kompatibilitu s implementáciou IEEE POSIX shellu a špecifikáciou nástrojov (tools specification) (IEEE Working Group 1003.2). Bash však nie je len veľmi výkonným shellom, ale taktiež silným skriptovacím jazykom. Podporuje prácu s premennými, cyklami `while`, `for`, `do`, funkciami a s mnohým ďalším.

Z dôvodu širokej kompatibility, rozšírenosti a dostupnosti na vzdialenej stanici v Číne je Bash vhodným kandidátom na implementáciu skriptov, pomocou ktorých budú vykonávané testy na overovanie internetovej cenzúry v Číne.

C++

Hlavnou výhodou jazyka C++ oproti ostatným rozšíreným objektovo-orientovaným jazykom ako Java alebo C# je jeho nezávislosť na konkrétnej programovej platforme, je štandardizovaný medzinárodným štandardom a poskytuje vysokú úroveň abstrakcie. Pre výsledný vygenerovaný kód je charakteristická vysoká rýchlosť behu[16][11].

PHP

PHP (*Hypertext Preprocessor*) je populárny open source skriptovací programovací jazyk, ktorý sa používa najmä na programovanie klient-server aplikácií (na strane servera) a pre vývoj dynamických webových stránok.

PHP bolo inšpirované jazykmi podporujúcimi štruktúrované programovanie. Najviac vlastností prebralo od jazyka C a jazyka Perl. V neskorších verziách bolo rozšírené o možnosť používať objekty.

Jedna zo zaujímavých vlastností PHP je, že umožňuje viac ako bežný skriptovací jazyk. Vďaka modulárnemu návrhu možno PHP používať aj na vývoj aplikácií s užívateľským rozhraním (GUI) [12].

Práve z hľadiska popularity a rozšíriteľnosti je tento skriptovací jazyk zvolený na implementáciu webovej služby umožňujúcej zobrazovanie výsledkov internetovej cenzúry.

HTTP

Protokol HTTP je jedným zo stavebných kameňov celej služby *World-Wide-Web*. Používa sa pre komunikáciu medzi prehliadačom a serverom. Protokol HTTP vychádza z architektúry klient/server. Klient, v našom prípade prehliadač, sa spojí so serverom a pošle mu požiadavku. Server ako reakciu na jeho požiadavku zasiela odpoveď. HTTP beží nad transportným protokolom TCP a štandardne sa klient pripojuje k WWW serveru na port 80, aj keď je možné v konfigurácii serveru vybrať ľubovoľný iný port. Protokol umožňuje vytvoriť si vlastné hlavičky, ktoré sú využité aj v tejto diplomovej práci.

Knižnica libcurl

Libcurl je voľne dostupná knižnica (*open source*) na prenos dokumentov a súborov zo serverov využitím niektorého z podporovaných protokolov. Knižnica je navrhnutá tak, aby naprogramované aplikácie nevyžadovali interakciu s užívateľom a umožnili tak jednoducho vykonávať automatické úlohy.

Táto knižnica umožňuje jednoducho vytvoriť aplikácie, ktoré sa dokážu pripojiť a komunikovať s mnohým rôznymi typmi serverov využitím rôznych typov protokolov. Aktuálne podporované protokoly sú:

- HTTP
- HTTPS
- FTP
- Telnet
- Dict
- LDAP

Libcurl je možné využiť spolu so skriptovacím jazykom PHP, čo dovoľuje webovým vývojárom písať aplikácie, ktoré komunikujú bezpečne pomocou SSL (*Secure Sockets Layer*). Z hľadiska diplomovej práce je podstatná podpora protokolu http, pomocou ktorého bude implementovaná komunikácia medzi lokálnou webovou službou a vzdialeným čínskym serverom. Samotnú knižnicu, rovnako ako aj bližšie informácie o nej je možné nájsť na adrese `http://curl.haxx.se/`.

AJAX

AJAX (*Asynchronous JavaScript* + XML) je súhrnné označenie pre technológie vývoja interaktívnych webových aplikácií, ktoré umožňujú meniť obsah stránok bez potreby ich kompletného znovunačítania zo serveru. V porovnaní s klasickými webovými aplikáciami môžu aplikácie využívajúce AJAX pri vhodnom návrhu poskytovať používateľsky komfortnejšie prostredie, vyžadujú však použitie moderných webových prehliadačov [17].

6.2 Doplňujúce balíčky

Prvou časťou systému na overovanie cenzúry v Číne je upravené prostredie na vzdialenej stanici. Vzhľadom na to, že jej možnosti sú výrazne obmedzené (tak ako bolo spomínané v predchádzajúcej kapitole) je potrebné rozšíriť ju o základné utility, pomocou ktorých je možné vykonávať sofistikovanejšie testy. Prostredie o jednotlivé utility doplníme ich preložením a následným prekopírovaním z pôvodných balíčkov, ktoré na stanici v Číne chýbajú.

Prvým prekladaným a staticky zostaveným balíčkom je balíček `bind`, ktorý obsahuje utility `nslookup` a `dig`. V zdrojových kódoch balíčku je potrebné upraviť Makefile, kvôli statickému zostaveniu výsledného programu. Konkrétne boli pridané dva nasledujúce riadky:

```
dig_LDFLAGS = -static-libgcc -static
nslookup_LDFLAGS = -static-libgcc -static
```

Uvedené riadky zabezpečia statické zostavenie jednotlivých utilít spolu s potrebnými súčasťami. Výsledné binárne súbory nie sú závislé na externých knižniciach, čo umožní ich bezproblémové spúšťanie na obmedzenom testovacom prostredí v Číne. Rovnakým spôsobom boli upravené a preložené balíčky `inetutils` a `iputils`. Vo všetkých prípadoch bol doplnený príslušný Makefile o potrebný prefix pre konkrétnu sieťovú utilitu. Zo sieťových balíčkov boli preložené tieto utility: `dig`, `nslookup`, `ping`, `telnet`, `tracepath` (utilita `ping` nebola v konečnom dôsledku využitá z dôvodov uvedených v predchádzajúcej kapitole).

6.3 Testovacie skripty

Základom vykonávania testov na overovanie cenzúry je sada skriptov implementovaná v skriptovacom jazyku Bash. Celkovo ide o štyri skripty, ktoré umožňujú overovať DNS preklady, možnosť pristúpiť na danú webovú stránku, resp. IP adresu a porovnanie výsledkov jednotlivých webových vyhľadávačov. Všetky skripty využívajú prostriedky dostupné na stanici v Číne, resp. doplnené utility. Všetky skripty zobrazujú získané výsledky vo formáte XML.

Skript na DNS preklady

Prvý skript s názvom dns.sh slúži na základné DNS preklady. Na vstupe očakáva URL domény, pre ktorú chceme získať príslušné IP adresy. Pre každú vrátenú IP adresu následne pomocou protokolu whois získa informácie o tom, komu je konkrétna IP adresa pridelená. Algoritmicky funguje skript nasledovne:

1. Pomocou utility `dig` získa pridelené IP adresy pre zadané URL.
2. Pre každú IP adresu kontaktuje server `whois.arin.net` príkazu `telnet` na porte 43, aby získal informácie o registrátorovi IP adresy.
3. Skript ako prvé hľadá vo výstupe protokolu `whois` položku `ReferralServer:`, ktorá informuje o tom, že informáciu pre daný rozsah IP adries spravuje iný server.
4. Ak skript položku `ReferralServer:` nenájde, pokúsi sa získať obsah položiek `OrgName:` a `Country:`.
5. Ak neuspeje, hľadá obsah položiek `descr:` a `country:`.
6. Ak vo výpise nájde vyplnenú položku `ReferralServer:`, presmeruje sa na uvedený `whois` server a vyhľadávanie opakuje na ňom.
7. Skript vypíše nájdené údaje vo formáte XML.

Nasledujúce XML reprezentuje príklad výstupu skriptu `dns.sh` pre `www.facebook.com`:

```
<?xml version="1.0"?>
<lookup>
<item>
<value>69.171.242.70</value>
<org>Facebook, Inc.</org>
<country>US</country>
<item>
</lookup>
```

Z výstupu je viditeľné, že URL www.facebook.com odpovedá IP adresa 69.171.242.70, ktorá je pridelená spoločnosti Facebook, Inc. so sídlom v Spojených štátoch (skript bol spúšťaný z českého prostredia).

Skript na testovanie prístupu na webové stránky

Druhý skript overuje možnosť na stránku priamo pristúpiť. To znamená, že bez ohľadu na DNS preklad sa pokúsi na zadané URL pripojiť a stiahnuť obsah stránky. Využíva utilitu `wget` a ako výstup vracia vo formáte XML výpis získaný cez `wget`, z ktorého je možné usúdiť, či pri pokuse o prístup na webovú stránku došlo k chybe alebo spojenie prebehlo v poriadku a obsah stránky sa podarilo bez problémov stiahnuť.

Skript na porovnanie výsledkov vyhľadávačov

Ďalší skript (`search.sh`) testuje vyhľadávače a ich výsledky. Testované vyhľadávače sú `www.google.com`, `search.yahoo.com`, `www.baidu.com` a `s.weibo.com`. Prvé dva vyhľadávače sú vyhľadávače mimo Čínu a druhé dva sú priamo čínske. Výsledky sú doplnené o overenie prístupu na anglickú verziu Wikipedie (`en.wikipedia.org/wiki`).

V tabuľke 6.1 sú uvedené vyhľadávacie parametre pre jednotlivé vyhľadávače.

Skript vracia výsledok vyhľadávania kľúčových slov pre všetky uvažované vyhľadávače. Výsledok je opäť zobrazený v podobe XML. Výstup obsahuje tak ako v prípade predošlého skriptu výpis z aplikácie `wget`.

URL	Parameter
www.google.com	/search?q=keyword
search.yahoo.com	/search?p=keyword
www.baidu.com	/?wd=keyword
s.weibo.com	/weibo/keyword
en.wikipedia.org	/wiki/keyword

Tabulka 6.1: Testované vyhľadávače a ich parametre

Skript na priamy prístup na IP adresu

Posledný skript (`ip.sh`) sa pripája na konkrétnu IP adresu a posiela HTTP požiadavku na zobrazenie odpovedajúcej webovej stránky. Následne vypíše výsledok. Prácu skriptu je možné zhrnúť do nasledujúcich krokov:

- Vytvorenie HTTP požiadavky:

  ```
  GET / HTTP/1.1 CRLF
  Host: <URL> CRLF
  CRLF
  CRLF
  ```

- Pripojenie sa na zadanú IP adresu pomocou utility `telnet`.
- Poslanie vytvorenej HTTP požiadavky.
- Zobrazenie výsledku.

Skript umožňuje sledovať reakciu čínskeho firewallu na priamy prístup na IP adresu, ktorá je blokovaná (beží na nej blokovaná webová stránka).

6.4 Server v čínskom a českom prostredí

Ďalšou časťou systému je server implementovaný v jazyku C++ za využitia *BSD socketov*. Server prijíma požiadavky a na ich základe vykoná požadovaný skript s predloženými parametrami. Požiadavky, ktoré prijíma sú požiadavkami protokolu HTTP. Z prijatej požiadavky vyparsuje obsah proprietárnej hlavičky `X-1234`. Táto hlavička obsahuje zakódovaný názov skriptu spolu s parametrami v Base64 formáte. Po dekódovaní obsahu hlavičky, spustí uvedený skript a pomocou rovnakej hlavičky vráti výsledok klientovi.

Server je implementovaný pomocou triedy `Server`, ktorá obsahuje metódy na prácu so sieťovým spojením. Samostatným modulom serveru je modul `Exec`, ktorý sa stará o spúšťanie predaného skriptu. (De)kódovanie formátu Base64 vykonáva rovnomenný modul `Base64`.

V nasledujúcom výpise je možné vidieť ukážku komunikačného protokolu medzi serverom a klientom:

- Požiadavka protokolu HTTP od klienta:

```
GET / HTTP/1.1
Host: 221.12.162.30:7777
Accept: */*
X-1234: YXZhaWxhYmlsaXR5LnNoIHd3dy5mYWNlYm9vay5jb20=
```

Uvedená požiadavka je interpretovaná ako žiadosť o overenie dostupnosti stránky www.facebook.com (zakódované pomocou Base64)

- Vrátenie odpovede klientovi:

```
HTTP/1.0 200 OK
Connection: close
X-1234: PD94bWwgdmVyc2lvbj0iMS4wIj8+PGxvb2t1cD48aXRlbT48dmFsdWU+MjA
zLjk4LjcuNjU8L3ZhbHVlPjxvcmc+IFRlbHN0cmFDbGVhciBMdGQ8L29yZz48Y291bn
RyeT4gTlo8L2NvdW50cnk+PC9pdGVtPjwvbG9va3VwPgo=
```

Interpretáciou výsledku dostaneme informáciu o tom, že URL `www.facebook.com` odpovedá IP adresa `203.98.7.65`, ktorú má registrovanú spoločnosť TestraClear Ltd. sídliaca na Novom Zélande.

6.5 Webová aplikácia

Viditeľným výstupom diplomovej práce je aplikácia na analýzu cenzúry. Jej základom je webová aplikácia, pomocou ktorej je možné zobraziť fakty týkajúce sa aktuálneho stavu internetovej cenzúry v Číne. Aplikácia sa bude pripájať na vzdialený počítač pomocou HTTP komunikácie na konfigurovateľnom porte (komunikačný protokol je uvedený vyššie). Medzi jednotlivými funkcionalitami je možné sa prepínať pomocou menu. Všetky možnosti webovej aplikácie súvisia so skriptami opísanými v úvode tejto kapitoly. Každá položka menu umožňuje požiadať o výsledok niektorého zo skriptov.

Základom komunikácie medzi webovou aplikáciou a vzdialeným serverom je PHP skript `client_curl.php`, ktorý využitím knižnice *libcurl* vytvorí HTTP požiadavku, ktorú pošle serveru na spracovanie. Skript zároveň prijíma odpoveď, vykonáva kódovanie/dekódovanie formátu Base64 a výsledok predáva hlavnej aplikácii (dekódovaný výsledok predstavuje XML s výsledkom overovania cenzúry).

Klientská časť webovej aplikácie je implementovaná pomocou technológie AJAX. Na získavanie dát využíva objekt `XMLHttpRequest()`, resp. `ActiveXObject ("Microsoft-.XMLHTTP")` v závislosti na použitom webovom prehliadači. V prípade overovanie DNS prekladu sú vytvorené objekty dva, aby bolo možné sa pripojiť na dve rôzne inštancie serveru (jedná beží v českom prostredí, druhá v čínskom). Po prijatí dát vo formáte XML dôjde k ich zobrazeniu koncovému užíateľovi.

Kapitola 7

Dosiahnuté výsledky

V tejto kapitole sú zhrnuté výsledky analýzy cenzúry v Čínskej ľudovodemokratickej republike. V úvodnej časti je predstavený postup ako bola analýza overovaná pomocou jednotlivých nástrojov. Na jej základe bolo možné vyvodiť viaceré závery, ktoré je možné overiť pomocou aplikácie, ktorej návrh a implementácia boli opísané v predchádzajúcich kapitolách.

7.1 DNS cache poisoning

Prvou vypozorovanou technikou internetovej cenzúry v Číne je *DNS cache poisoning*. V nasledujúcich riadkoch je možné nájsť vysvetlený základný princíp DNS cache poisoningu v súvislosti s fungovaním systému DNS. Následne je opísaný prístup k podvrhávaniu IP adries v Číne.

7.1.1 Základy DNS

Pri pokuse o prístup na webovú stránku je potrebné poznať presnú adresu, tzn. preklad URL na IP adresu. Pri zisťovaní IP adresy sa systém ako prvé pokúsi skontrolovať, či nie je potrebná informácia dostupná v lokálnej DNS cache a ak nie, tak pokračuje dotazovaním DNS serverov nakonfigurovaných pre konkrétnu stanicu (PC). Tieto servery následne začnú posielať sériu požiadaviek, najprv na root servery, potom na rodičovské servery pre cieľovú doménu (napr. pre doménu `.com`) a nakoniec na autoritatívne servery pre špecifickú doménu.

V momente, keď DNS server získa IP adresu z autoritatívneho serveru, s najväčšou pravdepodobnosťou si ju uloží do lokálnej cache pamäte (aby predišiel budúcim prekladom v krátkom časovom období) predtým, ako vráti správnu IP adresu. V nasledujúcich krokoch je opísaný popis prekladu pomocou DNS pri prístupe z webového prehliadača[9]:

1. Užívateľ zadá URL adresu do webového prehliadača, napr. `www.google.com`
2. Počítač nemá uloženú IP adresu pre túto doménu v lokálnej cache pamäti, preto sa pripojí na aktuálne nakonfigurovaný DNS server (pravdepodobne na ten, ktorý má nakonfigurovaný ISP) a požiada o IP adresu `www.google.com`
3. DNS server požiada root servery, ktoré presmerujú požiadavku na *Global Top Level Domain* (gTLD) servery pre všetky `.com` domény, ktoré presmerujú požiadavku na menné servery Google.

4. Tieto menné servery vrátia IP adresu pre túto doménu (napr. `74.125.237.80`).

5. DNS server si túto IP adresu uloží do lokálnej cache pamäte a pošle ju počítaču, ktorý o preklad žiadal.

6. Počítač sa následne pripojí na IP adresu (`74.125.237.80`) a požiada o obsah webovej stránky `www.google.com`.

7.1.2 DNS cache poisoning

DNS cache poisoning nastáva v prípade, kedy je IP adrese vrátená z autoritatívneho serveru pre doménu zmenená (podvrhnutá) predtým, ako je doručená koncovému užívateľovi. Keďže táto informácia prechádza cez značné množstvo systémov pred tým, ako ju získa koncový užívateľ, existuje mnoho bodov, v ktorých mohlo dôjsť k jej podvrhnutiu. Podvrhnutú IP adresu môže vrátiť kompromitovaný router v rámci siete ISP alebo transparentný DNS server niekde na ceste paketov ku koncovému užívateľovi. Podvrhovaná IP adresa môže byť získavaná aj lokálne z dôvodu prítomnosti vírusu, malware alebo iného *end-point* softwarového riešenia[14].

Pri DNS cache poisoningu dochádza k nasledujúcim krokom:

1. Užívateľ zadá `www.google.com` do webového prehliadača.

2. Počítač nemá uloženú IP adresu pre túto doménu v lokálnej cache pamäti, preto sa pripojí na aktuálne nakonfigurovaný DNS server (pravdepodobne na ten, ktorý má nakonfigurovaný ISP) a požiada o IP adresu `www.google.com`

3. DNS server požiada root servery, ktoré presmerujú požiadavku na *Global Top Level Domain* (gTLD) servery pre všetky `.com` domény, ktoré presmerujú požiadavku na menné servery Google.

4. Tieto menné servery vrátia IP adresu pre túto doménu (napr. `74.125.237.80`).

5. DNS server si túto IP adresu uloží do lokálnej cache pamäte a pokúsi sa poslať ju počítaču, ktorý o preklad žiadal.

6. Ako je táto informácia poslaná koncovému užívateľovi, musí prejsť cez niekoľko systémov, medzi ktoré patria servery, routre, atď. V tomto príklade uvažujme, že jeden zo systémov bol nakonfigurovaný tak, aby podvrhával všetky DNS dáta, ktoré vyhovujú `www.google.com` a namiesto správnej odpovede poslal informáciu obsahujúcu rozdielnu IP adresu (napr. `123.123.123.123`).

7. Počítač sa následne pripojí na IP adresu (`123.123.123.123`) a požiada o obsah webovej stránky `www.google.com`. Keďže server na tejto IP adrese nie je serverom Google, stránka sa nenačíta alebo dôjde k načítaniu rozdielnej webovej stránky.

DNS cache poisoning môže mať aj (diskutabilné) legitímne využitie, ako napr. vrátenie stránky internetového vyhľadávača v prípade, že užívateľ žiadal o neexistujúcu doménu (detekované na strane ISP).

7.2 DNS cache poisoning v Číne

Fungovanie cenzúry v Číne je možné overiť porovnávaním DNS prekladov, napr. pomocou nástroja dig. DNS požiadavky putujú fyzicky aj mimo Čínu, aby bolo možné kontaktovať DNS server pre danú webovú stránku. Zmenou DNS záznamu je možné usudzovať, že dochádza ku kontrole na úrovni routrov fyzicky lokalizovaných v Číne. Hneď ako je DNS odpoveď identifikovaná ako nevhodná na základe kľúčových slov, jej obsah je modifikovaný predtým, ako ho dostane koncový užívateľ. Jednoduchým príkladom je napr. webová stránka `www.facebook.com` , ktorá je v Číne nedostupná. Pri pokuse o DNS preklad z českého prostredia dostaneme nasledujúce výsledky:

```
$ dig www.facebook.com A +short
69.171.242.70
```

Daný výsledok prezrádza, že k obsahu stránky `www.facebook.com` sa dostaneme pripojením na IP adresu `69.171.242.70`. Pokúsme sa teraz spustiť rovnaký test z čínskeho prostredia:

```
$ ./dig www.facebook.com A +short
59.24.3.173
```

Z výsledku je možné usúdiť, že došlo k podvrhnutiu IP adresy pre doménu `www.facebook.com`. Pri pokuse pripojiť sa na uvedenú stránku z Číny pomocou aplikácie `wget` dostaneme nasledujúci výsledok:

```
$ wget www.facebook.com
--2012-05-01 20:27:28?http://www.facebook.com/
Resolving www.facebook.com ... 59.24.3.173
Connecting to www.facebook.com[59.24.3.173]:80...
Connection Timeout
```

Nepodarilo sa pripojiť na uvedenú IP adresu, pretože došlo k vypršaniu spojenia. Takéto chovanie nie je jednoznačné, preto je potrebné vykonať viacero testov, ktoré pomôžu odhaliť dôvod. Jedným z dôvodov je, že na uvedenej IP adrese nebeží žiadna služba na použitom porte (v tomto prípade podľa výpisu aplikácie `wget` bol použitý HTTP port 80). Rovnako je možné, že na danej IP adrese nebeží vôbec nič, a že je cieľová stanica nedostupná. Na to pomôže `ping` (vykonaný z českého prostredia, v čínskom nie je dostupný):

```
C:\Users\Jakub Tomaga>ping 59.24.3.173
Pinging 59.24.3.173 with 32 bytes of data:
Request timed out.
Request timed out.
Request timed out.
Request timed out.
Ping statistics for 59.24.3.173:
Packets: Sent = 4, Received = 0, Lost = 4 (100% loss)
```

Podľa výstupu utility `ping`, všetky pokusy o spojenia skončili vypršaním časového limitu (*timeout*). To samozrejme môže znamenať, že je smerom k cieľovej stanici blokovaný protokol ICMP. Overiť dostupnosť aplikačného protokolu HTTP môžeme napr. pomocou nástroja `telnet`:

```
$ telnet 59.24.3.173 80
Trying 59.24.3.173?
telnet: connect to address 59.24.3.173: Connection refused
```

Nie je možné sa na uvedenú IP adresu pripojiť na port 80. Keďže sa takto správa spojenie z Číny, tak z uvedenej IP adresy nedostaneme žiadne dáta, pomocou ktorých by sme boli informovaní, že došlo k zablokovaniu prístupu. Namiesto toho sa len užívateľ dozvie, že prístup na webovú stránku nie možný, pretože je spojenie pomalé, alebo že došlo k vypršaniu času, ktorý sa na obsah stránky čakalo. Takýmto spôsobom sa koncový užívateľ nedozvie, že podlieha cenzúre. Záver, ktorý z toho usúdi je napr. ten, že stránka neexistuje a tým pádom nie je možné na ňu pristúpiť.

V snahe lokalizovať čínske firewally sa pokúsime získať viac informácií o uvedenej IP adrese. Využitím protokolu *whois* môžeme odhaliť, kto registroval danú IP adresu. Za IP adresy v rozsahu `59.0.0.0/8` odpovedá server `whois.apnic.net`. Pomocou utility `telnet` je možné získať informácie o IP adrese `59.24.3.173` (výstup je vhodne skrátený):

```
$ telnet whois.apnic.net 43
Trying 202.12.29.220...
Connected to whois.apnic.net.
Escape character is '^]'.
% [whois.apnic.net node-4]
% Whois data copyright terms    http://www.apnic.net/db/dbcopyright.html

59.24.3.173
inetnum:        59.0.0.0 - 59.31.255.255
netname:        KORNET
descr:          KOREA TELECOM
descr:          Network Management Center
country:        KR
admin-c:        IM76-AP
tech-c:         IM76-AP
?
status:         Allocated Portable
mnt-by:         MNT-KRNIC-AP
changed:        hm-changed@apnic.net 20040802
changed:        hm-changed@apnic.net 20041007
source:         APNIC

person:         IP Manager
nic-hdl:        IM76-AP
e-mail:         kornet_ip@kt.com
address:        Seoul
address:        206, Jungja-Dong, Bundang-Gu, Sungnam, Gyunggi-Do
address:        463-711
phone:          +82-2-500-6630
fax-no:         +82-2-3674-5721
country:        KR
changed:        hostmaster@nic.or.kr 20111229
mnt-by:         MNT-KRNIC-AP
```

IP adresa	Umiestnenie	Vlastník
8.7.198.45	Spojené štáty americké	ARIN
37.61.54.158	Európa	European Regional Registry
46.82.174.68	Nemecko	Deutsche Telecom
78.16.49.15	Írsko	Esat Telecommunications Limited
93.46.8.89	Taliansko	Fastweb
159.106.121.75	Spojené štáty americké	DoD Network Information Center
203.98.7.65	Nový Zéland	Telstra Clear

Tabulka 7.1: Podvrhávané IP adresy a ich vlastníci

```
source:         APNIC

inetnum:        59.0.0.0 - 59.31.255.255
netname:        KORNET-KR
descr:          Korea Telecom
country:        KR
admin-c:        IA9-KR
tech-c:         IM9-KR
status:         ALLOCATED PORTABLE
mnt-by:         MNT-KRNIC-AP
mnt-irt:        IRT-KRNIC-KR
remarks:        This information has been partially mirrored by APNIC from
remarks:        KRNIC. To obtain more specific information, please use the
remarks:        KRNIC whois server at whois.krnic.net.
changed:        hostmaster@nic.or.kr
source:         KRNIC

Connection closed by foreign host.
```

Z výpisu je možné vyčítať, že IP adresu `59.24.3.173` má registrovanú Korea Telecom. To znamená, že IP adresa, na ktorú je presmerovaná webová stránka `www.facebook.com` je fyzicky lokalizovaná mimo Čínu. Ďalej vieme z predchádzajúcej analýzy, že na nej nebeží port 80 a že je nedostupná cez protokol ICMP.

Vykonaním ďalších testov sa podarilo zistiť, že IP adresy, na ktoré sú presmerované stránky ako sú `www.facebook.com`, `www.twitter.com` apod. rotujú. To znamená, že IP adresy, ktoré získame DNS prekladom sa v čase menia, v jeden moment majú rôzne webové stránky rozdielne podvrhnuté IP adresy. Všetky sú nedostupné (nie je možné sa pripojiť na žiadny port z platného rozsahu) a nie je možné overiť ich dostupnosť pomocou aplikácie `ping`, pretože ten končí na vypršaní spojenia. V tabuľke 7.1 sú zhrnuté IP adresy spolu s organizáciami, ktorým IP adresy patria. Jednotlivé IP adresy boli vypozorované v čase a informácie o vlastníkovi boli získané pomocou implementovanej aplikácie (skript `dns.sh`).

Z vyššie uvedenej tabuľky vyplýva, že do internetovej cenzúry sú zapletené aj ostatné krajiny. Otázkou je prečo uvedené organizácie pomáhajú Číne pri realizácii cenzúry. Zároveň treba vziať do úvahy fakt, že IP adresy sú nedostupné, podľa analýzy na nich nebežia žiadne

služby, čo môže znamenať, že tieto IP adresy nie sú jednotlivými organizáciami využívané. Dáta získané pomocou protokolu `whois` je možné interpretovať aj alternatívnym spôsobom. Možným vysvetlením, je to, že IP adresy boli vybrané na strane čínskej vlády po tom, ako si overili to, že sú skutočne nevyužité a aby presvedčili domáce obyvateľstvo, že výsledok prichádza zo zahraničia.

Zaujímavý prístup k podvrhávaniu IP adries je možné sledovať pri pokuse o preklad neexistujúcej domény. Nasledujúci výpis ukazuje preklad domény, ktorá obsahuje kľúčové slovo facebook:

```
$ dig www.facebooknonexistantdomain.com A +short
220.250.64.20
```

Získaná IP adresa neodpovedá žiadnej z IP adries, ktorú sa podarilo identifikovať počas testovania internetovej cenzúry. Prístupom na túto IP adresu pomocou cez `wget` získame nasledujúci výsledok:

```
wget 220.250.64.20 -U "Firefox/3.0.15"
--2012-05-16 21:14:46--  http://220.250.64.20/
Connecting to 220.250.64.20:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 467 [text/html]
Saving to: ‘index.html’

100%[======================================>] 467         --.-K/s   in 0s

2012-05-16 21:14:46 (42.3 MB/s) - ‘index.html’ saved [467/467]
```

To znamená, že sa podarilo fyzicky stiahnuť webovú stránku, na ktorú bola neexistujúca doména presmerovaná. Zdrojový kód stránky je možné vidieť v nasledujúcom výpise:

```
$ less index.html
<html>
<script language=javascript type="text/javascript">
    window.location.replace("http://nfdnserror3.wo.com.cn:8080?HOST="
        + location.hostname + "&R="
        + location.pathname + "&" + location.search.substr(location.search.
        indexOf("\?")+1));
</script>

<noscript>
<meta http-equiv="refresh" content="0;URL=http://nfdnserror3.wo.com.cn:8080">
</noscript>
<head>
<title>Redirect</title>
</head>
<body bgcolor="white" text="black">
</body>
</html>
```

Analýzou zdrojového kódu je možné si všimnúť, že dochádza k náhrade URL vo webovom prehliadači, za inú URL adresu pomocou kódu skriptovacieho jazyka JavaScript:

`window.location.replace`. Pôvodná URL sa rozparsuje do parametrov. V prípade, že je JavaScript vypnutý dôjde k presmerovaniu na uvedenú doménu bez predania informácii o pôvodnej požiadavke. V tomto prípade môže ísť o podporu na strane ISP, ktorý informuje, že došlo k pokusu o pripojenie sa na neexistujúcu doménu, tak ako to bolo opísané v úvode tejto kapitoly.

7.3 Blokovanie na základe kľúčových slov

V predchádzajúcej podkapitole bolo možné vidieť ako prebieha cenzúra v Číne na úrovni podvrhávania IP adries na blokované stránky. Nie všetky nedostupné stánky sú blokované týmto spôsobom. Mnohé stránky, ktoré podľa rozličných zdrojov v Číne prístupné nie sú, je možné na základe DNS prekladu označiť za bezproblémové. Viaceré stránky z rôznych kategórií sú blokované na základe kľúčových slov, ktoré sa na stránke nachádzajú. Príkladom je napríklad stránka `www.youtube.com`. Pri pokuse o DNS preklad z českého prostredia dostávame nasledujúci výsledok:

```
$ dig www.youtube.com A +short
youtube-ui.l.google.com.
74.125.232.233
74.125.232.238
74.125.232.224
74.125.232.225
74.125.232.226
74.125.232.227
74.125.232.228
74.125.232.229
74.125.232.230
74.125.232.231
74.125.232.232
```

Na porovnanie si zobrazíme výsledok z čínskeho prostredia:

```
$ ./dig www.youtube.com A +short
youtube-ui.l.google.com.
youtube-ui-china.l.google.com.
72.14.203.102
72.14.203.113
72.14.203.138
72.14.203.139
72.14.203.100
72.14.203.101
```

V oboch prípadoch vidíme legitímne výsledky, ktoré naznačujú, že v Číne je server youtube.com prístupný z rozdielnych IP adries. Pomocou protokolu `whois` si môžeme zobraziť výstup zo serveru `whois.arin.net` (skrátený):

```
$ telnet whois.arin.net 43
Trying 199.71.0.48...
Connected to whois.arin.net.
Escape character is '^]'.
72.14.203.102
...
NetRange:       72.14.192.0 - 72.14.255.255
CIDR:           72.14.192.0/18
OriginAS:
NetName:        GOOGLE
NetHandle:      NET-72-14-192-0-1
Parent:         NET-72-0-0-0-0
NetType:        Direct Allocation
RegDate:        2004-11-10
Updated:        2012-02-24
Ref:            http://whois.arin.net/rest/net/NET-72-14-192-0-1


OrgName:        Google Inc.
OrgId:          GOGL
Address:        1600 Amphitheatre Parkway
City:           Mountain View
StateProv:      CA
PostalCode:     94043
Country:        US
RegDate:        2000-03-30
Updated:        2011-09-24
Ref:            http://whois.arin.net/rest/org/GOGL
```

Výstup naznačuje, že aj výsledné IP adresy získané z Číny má vo vlastníctve Google, prevádzkovateľ serveru `youtube.com`. Problém ale nastáva pri pokuse o prístup na stránku priamo, napr. pomocou aplikácie `wget`:

```
wget www.youtube.com -U "Firefox/3.0.15"
--2012-05-16 22:49:59--  http://www.youtube.com/
Resolving www.youtube.com... 72.14.203.139, 72.14.203.100, 72.14.203.101,
Connecting to www.youtube.com|72.14.203.139|:80... connected.
HTTP request sent, awaiting response... Read error (Connection reset by
peer) in headers.
Retrying.
```

V tomto prípade je možné sledovať nový typ chovania: *Connection reset by peer*. To znamená, že po vyhodnotení stránky ako blokovanej druhá strana okamžite ukončí spojenie. Rovnaké chovanie je možné identifikovať pri prístupe na akúkoľvek stránku z kategórie porno stránok.

7.4 Blokovanie vyhľadávania

Internetová cenzúra sa Číne vzťahuje aj na vyhľadávané slová vo vyhľadávačoch. V rámci diplomovej práce sú analyzované vyhľadávače medzinárodné ako `google.com` a `search.-`

Kľúčové slovo
Great Firewall of China
Free Tibet
Massacre
Porn
Massacre
Jasmine revolution
Minghui
Dharma Chakra
Peacehall
Freedom
Red Terror
Tiananmen incident
Brutal torture
Boxun
Liu Xiaobo
Molihua
Falun

Tabulka 7.2: Príklad blokovaných kľúčových slov

`yahoo.com`, rovnako ako čínske národné služby `baidu.com` a`weibo.com`. Prístupom na webové stránky vyhľadávača s priamo zadaným parametrom s kľúčovým slovom je možné overiť, či sa dajú výsledky zobraziť v Číne. Najviac citlivý na anglické výrazy je vyhľadávač Google. Pomocou neho došlo k chybe typu *Read error: Connection reset by peer* okamžite, resp. pri presmerovaní na `www.google.com.hk` došlo k prijatiu odpovede `403 Forbidden` vo všetkých prípadoch, kedy sa testovali citlivé výrazy. Testované výrazy je možné vidieť v tabuľke 7.2.

Zaujímavosťou je fakt, že nie všetky uvedené výrazy sú blokované pomocou vyhľadávača `search.yahoo.com`. Spoločnou vlastnosťou oboch vyhľadávačov je to, že sa blokované kľúčové slovo nemusí nachádzať len ako korektný parameter (`search?q=param`, resp. `search?p=param`) vyhľadávania, ale je prakticky kdekoľvek v URL ako podreťazec. Blokované URL je aj `search.yahoo.com/search?p=falun`, rovnako ako `search.yahoo.com/search?p=substringfalunsubstring`. Po pokuse o prístup na stránku vyhľadávača s blokovaným kľúčovým slovom ostáva aj základný prístup bez vyhľadávania blokovaný. Ilustruje to nasledujúca sekvencia príkazov `wget`:

```
$ wget search.yahoo.com/search?p=falun --tries=1
--2012-05-23 07:01:51--  http://search.yahoo.com/search?p=falun
Resolving search.yahoo.com... 125.252.225.150, 125.252.225.152
Connecting to search.yahoo.com|125.252.225.150|:80... connected.
HTTP request sent, awaiting response... Read error (Connection reset
by peer) in headers.
Giving up.
```

Hlavná stránka	Nedostupné verzie
www.facebook.com	zh-cn.facebook.com m.facebook.com fr-fr.facebook.com www.new.facebook.com
www.twitter.com	dev.twitter.com search.twitter.com

Tabulka 7.3: Nedostupné verzie blokovaných stránok

```
$ wget search.yahoo.com --tries=1
--2012-05-23 07:01:59--  http://search.yahoo.com/
Resolving search.yahoo.com... 125.252.225.152, 125.252.225.150
Connecting to search.yahoo.com|125.252.225.152|:80... connected.
Failed writing HTTP request: Connection reset by peer.
Giving up.
```

Táto situácia trvá pomerne krátko. Po pár sekundách je spojenie opäť bezproblémové. Vyhľadávače `baidu.com` a `weibo.com` nereagujú na anglické výrazy v parametroch. Podľa všetkého sú citlivejšie na zakázané výrazy v čínštine. Táto varianta v rámci diplomovej práce nebola testovaná.

7.5 Blokovanie špecializovaných verzií stránok

Pri blokovaní stránok je možné sledovať zaujímavý jav. V úvodných kapitolách boli predstavené základné kroky, ktoré je možné vykonať ako prvé pokusy o obídenie cenzúry. Jednou z uvedených možností bolo skúsiť pristúpiť na špecializované verzie stránok, ktoré sú za normálnych okolností blokované. Príkladom môže byť napríklad mobilná verzia pre Wikipediu bežiaca na adrese `http://mobile.wikipedia.org`. Ak by v krajine, kde je aplikovaná internetová cenzúra nebola Wikipedia dostupná pomocou svojho klasického odkazu, bolo by možné vyskúšať tento prístup. Príkladom toho, že ani špecializované stránky nie sú dostupné z Číny je napríklad Facebook a Twitter. V tabuľke 7.3 je možné vidieť, ktoré špecializované stránky sú nedostupné.

Cenzúra v Číne nielen myslí na podobné špecializované verzie v prípade, že je stránka blokovaná, ale naopak neumožňuje pristúpiť na špecializované verzie stránok ani v prípade, že hlavná stránka je dostupná. Konkrétne uvažujme Wikipediu z čínskej perspektívy. V Číne je Wikipedia dostupná. Ale jej špecializované verzie už dostupné nie sú. Pre niektoré stránky nie sú dostupné lokalizované verzie, resp. doplnkové služby. V nasledujúcej tabuľke 7.4 je možné vidieť, ktoré webové stránky nemajú prístupné špecializované verzie, hoci ich pôvodná verzia je plne dostupná. Informácie boli získané pokusmi pripojiť sa na dané webové stránky.

V mnohých prípadoch ide o mobilné verzie (ne)blokovaných stránok a ich rôzne národne verzie. To môže byť spôsobené napríklad tým, že čínska vláda nesúhlasí s názormi vyjadrovanými v danej krajine, resp. ide o krajiny, ktoré čínsky občania často využívajú na obchádzanie cenzúry tým, že pristupujú na web prostredníctvom stránok v inom jazyku.

Dostupná verzia	Nedostupné verzie
www.google.com	feedproxy.google.com sites.google.com plus.google.com picasaweb.google.com translate.google.com
www.yahoo.com	hk.rd.yahoo.com meme.yahoo.com tw.yahoo.com video.yahoo.com tw.myblog.yahoo.com hk.yahoo.com
en.wikipedia.org	en.mobile.wikipedia.org zh.m.wikipedia.org mobile.wikipedia.org

Tabulka 7.4: Nedostupné verzie neblokovaných stránok

7.6 Blokovanie prístupu na IP adresy

Ďalšou možnosťou ako overiť cenzúru v Číne je zamerať sa na priamy prístup na IP adresy webových stránok, ktorých sa týka DNS cache poisoning. Predpokladom je, že originál IP adresy týchto webových stránok budú blokované. Priamo na overenie chovania v tejto situácii slúži jeden z dostupných skriptov, konkrétne `ip.sh`. Ako prvý krok sa pokúsime získať IP adresu odpovedajúcu `www.facebook.com`:

```
$ dig www.facebook.com A +short
69.171.247.53
```

V tomto momente sa pokúsime pristúpiť na IP adresu priamo z Číny a simulovať HTTP komunikáciu manuálne (rovnako postupuje aj uvedený skript):

```
$ ./telnet 69.171.247.53 80
Trying 69.171.247.53...
Connected to 69.171.247.53.
Escape character is '^]'.
GET / HTTP/1.1
Host: www.facebook.com
Connection closed by foreign host.

$ ./telnet 69.171.247.53 80
Trying 69.171.247.53...
Connected to 69.171.247.53.
Escape character is '^]'.
Connection closed by foreign host.

$ ./telnet 69.171.247.53 80
Trying 69.171.247.53...
Connected to 69.171.247.53.
```

```
Escape character is '^]'.
Connection closed by foreign host.
```

Vidíme, že pôvodná IP adresa blokovaná nie je. Je možné sa na ňu bez problémov pripojiť. Problém nastane, až keď sa cez sieť odošle HTTP hlavička `Host: www.facebook.com`. V tomto prípade dochádza okamžite k ukončeniu spojenia. Zaujímavosťou je to, že následne sa okamžite po pripojení na IP adresu spojenie uzavrie. Z viacerých testov vyplynulo, že takéto chovanie pretrváva približne 10 minút. Následne je možné celý test plnohodnotne zopakovať s rovnakým priebehom. Pri vykonaní rovnakého testu z českého prostredia je možné sledovať rozdielne chovanie komunikácie:

```
[jakub@localhost telnet]$ telnet 69.171.247.53 80
Trying 69.171.247.53...
Connected to 69.171.247.53.
Escape character is '^]'.
GET / HTTP/1.1
Host: www.facebook.com

HTTP/1.1 302 Found
Location: http://www.facebook.com/unsupportedbrowser
P3P: CP="Facebook does not have a P3P policy. Learn why here: http://fb.me/p3p"
X-Content-Type-Options: nosniff
X-Frame-Options: DENY
Set-Cookie: datr=pKG6T9xed8yJHIc0_DA69Ga8; expires=Wed, 21-May-2014 20:12:20 GMT;
path=/; domain=.facebook.com; httponly
Set-Cookie: lsd=AVpc41ek; path=/; domain=.facebook.com
Content-Type: text/html; charset=utf-8
X-FB-Debug: ppZzVoLbPUqHEZDBFpyaMTN47BjQqipOHVb/4OEcyHg=
Date: Mon, 21 May 2012 20:12:20 GMT
Content-Length: 0
Connection: keep-alive
```

Mimo Čínu dôjde správne k prijatiu odpovede zo serveru. V tom prípade je možné sledovať fakt, že `www.facebook.com` nepovoľuje prístup bez uvedenej hlavičky `User-agent`. Komunikácia ale prebehla bez problémov v porovnaní s Čínou.

Na základe predchádzajúcich testov sa podarilo overiť blokovanie na základe kľúčových slov v HTTP požiadavkách. Podobným spôsobom je možné nahliadnuť na to, či čínske firewally reagujú aj na HTTP odpovede. V tomto prípade budeme postupovať nasledovne:

```
$ ./telnet 69.171.247.37 80
Trying 69.171.247.37...
Connected to 69.171.247.37.
Escape character is '^]'.
GET / HTTP/1.1
Host: 69.171.247.37

HTTP/1.1 301 Moved Permanently
Location: http://www.facebook.com/
Content-Type: text/html; charset=utf-8
X-FB-Debug: kML2viDWsJ6ehRNTK4SW7fyeWylvRaUyLdXCa2Llgeg=
Date: Mon, 21 May 2012 21:28:15 GMT
Content-Length: 0
```

```
Connection: keep-alive

Connection closed by foreign host.

$ ./telnet 69.171.247.37 80
Trying 69.171.247.37...
Connected to 69.171.247.37.
Escape character is '^]'.
GET / HTTP/1.1
Connection closed by foreign host.
```

Týmto testom sme nechali protokol HTTP, aby vykonal presmerovanie priamo na `www.facebook.com`, čím sme docielili overenie blokovania aj na základe HTTP odpovede. Opäť je možné si všimnúť, že opätovný pokus o pripojenie sa okamžite končí ukončením spojenia. V tejto časti boli vykonané testy za účelom overenia blokovania pri prístupe priamo na IP adresy, ktoré sú v Číne podvrhávané. Podarilo sa odhaliť, že originálne IP adresy zablokované nie sú, ale komunikácia končí kvôli jej obsahu.

7.7 Analýza pomocou tracepath

Pomocou utility `tracepath` je možné overiť, či je komunikácia ukončená pri pokuse o kontaktovanie zakázaných webových stránok alebo nie. Predpokladané chovanie je, že komunikácia priamo na originálnu IP adresu blokovaných služieb prejde bez problémov k cieľu, pretože z predchádzajúcich testov vyplýva, že priame spojenie na IP adresy je bezproblémové a až kľúčové slová v komunikácii spôsobujú blokovanie. Prvým testom overíme predpoklad možnosti komunikácie so vzdialeným serverom, na ktorom beží blokovaná stránka pomocou IP adresy:

```
$ ./tracepath 69.171.242.74
 1:  221.12.162.30                                      0.275ms pmtu 1500
 1:  122.226.50.129                                     1.564ms
 1:  122.226.50.129                                     1.671ms
 2:  172.16.255.25                                      3.016ms
 3:  221.12.89.49                                       1.927ms
 4:  221.12.80.117                                      9.271ms asymm  5
 5:  219.158.14.209                                     14.747ms asymm  6
 6:  219.158.97.102                                     15.694ms asymm  7
 7:  219.158.97.90                                      15.636ms asymm  8
 8:  219.158.30.190                                     189.130ms asymm  9
 9:  vlan121.edge5.LosAngeles1.Level3.net               204.935ms asymm 16
10:  vlan60.csw1.LosAngeles1.Level3.net                 206.020ms asymm 15
11:  ae-82-82.ebr2.LosAngeles1.Level3.net               203.155ms asymm 16
12:  ae-3-3.ebr3.Dallas1.Level3.net                     212.834ms asymm 18
13:  ae-7-7.ebr3.Atlanta2.Level3.net                    268.467ms asymm 16
14:  ae-63-63.ebr1.Atlanta2.Level3.net                  263.670ms asymm 15
15:  ae-2-52.edge5.Atlanta2.Level3.net                  258.939ms asymm 16
16:  no reply
17:  no reply
18:  no reply
```

```
19:  no reply
20:  no reply
21:  no reply
22:  no reply
23:  no reply
24:  no reply
25:  no reply
26:  no reply
27:  no reply
28:  no reply
29:  no reply
30:  no reply
31:  no reply
     Too many hops: pmtu 1500
     Resume: pmtu 1500
```

Z výstupu vyplýva, že sa komunikácia dostala do Spojených štátov, tzn. mimo územie Číny. Pomocou GeoIP Tool (`http://geoiptool.com`) sa podarilo zistiť, že posledný uzol sa fyzicky nachádza v USA, čo potvrdzuje to, že sa komunikácia dostala bez problémov z Číny. Teraz skúsime spustiť `tracepath` pre `www.facebook.com`:

```
$ ./tracepath www.facebook.com
 1:  221.12.162.30           0.258ms pmtu 1500
 1:  122.226.50.129          0.916ms
 1:  122.226.50.129          1.598ms
 2:  172.16.255.25           2.700ms
 3:  221.12.89.49            1.416ms
 4:  no reply
 5:  no reply
 6:  no reply
 7:  no reply
 8:  no reply
 9:  no reply
10:  no reply
11:  no reply
12:  no reply
13:  no reply
14:  no reply
15:  no reply
16:  no reply
17:  no reply
18:  no reply
19:  no reply
20:  no reply
21:  no reply
22:  no reply
23:  no reply
24:  no reply
25:  no reply
```

```
26:  no reply
27:  no reply
28:  no reply
29:  no reply
30:  no reply
31:  no reply
     Too many hops: pmtu 1500
     Resume: pmtu 1500
```

V tomto prípade došlo k ukončeniu komunikácie v Číne, tzn. skôr ako komunikácia prekročila hranice. Na rovnakej IP adrese končí komunikácia pomocou `tracepath` v prípade všetkých blokovaných stránok, na ktoré sa vzťahuje DNS cache poisoning. Je pravdepodobne, že uvedená stanica funguje ako firewall, ktorý neprepúšťa blokovanú komunikáciu. Čo sa týka lokalizácie, tak uvedená IP adresa je v rovnakom meste ako je čínska stanica, z ktorej sú vykonávané všetky testy. Je možné, že ide o politiku lokálneho charakteru, resp. môže byť použitá čínska stanica náhodou lokalizovaná v blízkosti firewallu.

Kapitola 8

Záver

Diplomová práca sa zaoberala internetovou cenzúrou v Číne. Opisuje mnohé problémy so získavaním prístupu do Číny a prezentuje otestované alternatívy. Diskutuje dôvody, prečo boli jednotlivé varianty nepoužiteľné z technického alebo administratívneho hľadiska. Po získaní prístupu v Číne bolo možné začať naplno testovať vlastnosti internetového pripojenia z pohľadu obmedzených užívateľov. Práva vďaka tomu bolo možné získať zaujímavé informácie o vnútornom fungovaní čínskych firewallov. Počas doby riešenia sa postupne objavovali nové a nové aspekty, ktoré spolu umožnili utvárať ucelenejší pohľad na danú problematiku. Základným princípom používaným v Číne je DNS cache poisoning, teda podvrhávanie IP adries. Mnohé webové stránky sú z Číny nedostupné, pretože nie možné ich priamo kontaktovať. Detaily pozadia DNS cache poisoningu v Číne ponúkajú viac otázok ako odpovedí. Základnou otázkou je prečo žiadna podvrhávaná IP adresa nepatrí čínskej organizácii? Prečo sú medzi vlastníkmi podvrhávaných IP adries veľké telekomunikačné firmy, resp. Ministerstvo obrany Spojených štátov amerických? Sú IP adresy využité pri podvrhávaní vedome alebo ide len o náhodné adresy mimo používaný rozsah? Či už ide o DNS cache poisoning alebo blokovanie na základe kľúčových slov, vždy dochádza k obmedzovaniu užívateľov pri využívaní Internetu. Vo väčšine prípadov sú blokované výrazy spojené s historickými udalosťami, politickými prevratmi, resp. novými potláčanými myšlienkami. Blokované webové stránky sú z oblastí sociálnych sietí, pornografie a politiky. Výstup diplomovej práce umožňuje názorne nahliadnuť na Internetovú cenzúru. Poskytuje oproti dostupným nástrojom komplexnejší pohľad, čím bol naplnený prvotný cieľ, ktorým bolo rozšíriť existujúce nástroje o nové možnosti.

8.1 Budúce rozšírenia

Diplomová práca sa z väčšej časti zaoberá protokolom HTTP a prístupom na webové stránky. Ostatné aplikačné protokoly neboli analyzované. Hlavným dôvodom bola snaha udržať si prístup na stanicu v Číne. Viacero zdrojov uvádza rôzne porušenia pravidiel, ktoré stanovuje čínska vláda a následné dôsledky pre občanov, ktorí tak učinili [18]. Preto bola vynechaná kontrola e-mailových protokolov. Vynášanie nevhodných dát mimo územie Číny bolo v minulosti trestané. Jednoduchou analýzou cenzúry nemuselo dôjsť k žiadnej trestnej činnosti, ale riziko zablokovania a straty jediného testovacieho prostredia bolo vysoké. Viackrát počas riešenia diplomovej práce došlo k strate spojenia a bolo potrebné riešiť vzniknuté problémy s poskytovateľom SSH prístupu. Preto bola snaha vykonávať riskantné operácie len minimálne. Ďalším návrhom na rozšírenie je zvýšiť možnosti získanej stanice v

Číne. Na to sú vyžadované administrátorské práva. Tie by bolo možné získať aplikovaním exploitu na nainštalovanú verziu jadra operačného systému. Ak by sa k takému kroku pristúpilo pred odovzdaním práce, hrozila by opäť možnosť straty prístupu. Existuje viacero projektov na univerzitách, ktoré skúmajú internetovú cenzúru v Číne. Jedným takýmto projektom je projekt *Ignoring the Great Firewall of China*[1], v rámci ktorého sa na University of Cambridge snažia obísť cenzúru založenú na blokovaní na základe kľúčových slov tak, že pri výskyte TCP resetu, ignorujú RST flag na oboch koncoch komunikácie. Na vykonanie týchto úkonov je opäť potrebné manipulovať priamo s dátami a preto je súčasný systém nedostatočný z hľadiska dostupných práv. Administrátorské práva na staniciach v Číne by umožnili detailnejšie nahliadnuť na fungovanie cenzúry. Preto je práve iný typ prostredia základom na budúce pokračovanie v analýze cenzúry.

[1] Bližšie informácie na `http://www.cl.cam.ac.uk/ rnc1/ignoring.pdf`

Literatura

[1] China Channel. [online], [cit. 2012-01-04].
URL chinachannel.fffff.at/

[2] Chinese Firewall Checker. [online], [cit. 2012-01-03].
URL http://www.bestvpnservice.com/tools/chinese-firewall-test.php

[3] Great Firewall of China. [online], [cit. 2012-01-03].
URL http://www.greatfirewallofchina.org/about.php

[4] How To Bypass Internet Censorship. [online], [cit. 2011-12-30].
URL https://www.howtobypassinternetcensorship.org/

[5] OpenNet Initiative - About. [online], [cit. 2012-01-02].
URL http://opennet.net/about-oni

[6] OpenNet Initiative - Reports and Articles. [online], [cit. 2012-01-02].
URL http://opennet.net/reports

[7] WatchMouse. [online], [cit. 2012-01-02].
URL http://www.watchmouse.com/en/about.php

[8] Doseděl, T.: *Počítačová bezpečnost a ochrana dat*. Computer Press, 2004, ISBN 80-251-0106-1.

[9] Dostálek, L.; Kabelová, A.: *Velký průvodce protokoly TCP/IP a systémem DNS*. Computer Press, 2008, ISBN 978-80-251-2236-5.

[10] Jelínek, L.: *Jádro systému Linux - Kompletní průvodce programátora*. Computer Press, 2008, ISBN 978-80-251-2084-2.

[11] Josuttis, N. M.: *C++ Standardní knihovna a STL - Kompletní průvodce*. Computer Press, 2005, ISBN 80-251-0511-1.

[12] Kosek, J.: *PHP - tvorba interaktivních internetových aplikací*. Grada Publishing, 1997, ISBN 80-7169-373-1.

[13] Kurose, J. F.; Ross, K. W.: *Computer Networking A Top-Down Approach Featuring the Internet*. Addison-Wesley Publishing Company, druhé vydání, 2003, ISBN 0-321-17644-8.

[14] Olzak, T.: DNS Cache Poisoning: Definition and Prevention. [online], [cit. 2012-05-18].
URL http://adventuresinsecurity.com/Papers/DNS_Cache_Poisoning.pdf

[15] Porter, T.: *Practical VoIP Security.* Syngress Publishing, 2006, ISBN 1-597-49060-1.

[16] Prada, S.: *Mistrovství v C++.* Computer Press, 2007, ISBN 978-80-251-1749-1.

[17] Resig, J.: *JavaScript a Ajax - Moderní programovaní webových aplikácí.* Computer Press, 2007, ISBN 978-80-251-1824.

[18] Ronald J. Deibert, John G. Palfrey, Rafal Rohozinski, Jonathan Zittrain: *Access Denied - The Practice and Policy of Global Internet Filtering.* The MIT Press, 2008, ISBN 0-262-54196-3.

[19] Stallings, W.: *Cryptography and network security: principles and practice.* Prentice Hall, 1999, ISBN 0-13-869017-0.

Příloha A

Obsah CD

Technická správa

Správa vo formáte PDF sa nachádza v hlavnom adresári pod názvom projekt.pdf. Zdrojové súbory práce sa nachádzajú v adresári `tex/`.

Zdrojové kódy

Zdrojové kódy upravených balíčkov sa nachádzajú v adresári `bin/`. Skripty na testovanie sú adresári `scripts/`. Server implementovaný v jazyku C++ je umiestnený v samostatnom adresári `server/` a kompletné zdrojové kódy webovej aplikácie sú v adresári `web/`.