Vysoká škola ekonomická v Praze

Fakulta informatiky a statistiky

Katedra systémové analýzy

Studijní program: Aplikovaná informatika

Studijní obor: Podniková informatika

Státní cenzura Internetu a metody jejího obcházení

Autor diplomové práce : Bc. Martin Kolář

Vedoucí diplomové práce : prof. Ing. Petr Doucek, CSc.

Oponent diplomové práce : Ing. Jiří Hološka, Ph.D.

Rok odevzdání: 2016

Prohlášení

Prohlašuji, že jsem diplomovou práci zpracoval samostatně a že jsem uvedl všechny použité prameny a literaturu, ze kterých jsem čerpal.

V Praze dne 27. června 2016 …….

Bc. Martin Kolář

Poděkování

Na tomto místě bych rád poděkoval panu prof. Ing. Petru Douckovi, CSc., za jeho vstřícnost, hodnotné připomínky a odborné rady, kterými mi pomohl k vypracování této diplomové práce.

Abstrakt

Diplomová práce se zabývá státní cenzurou Internetu. Cílem práce je zmapovat techniky státní cenzury Internetu a metody pro její obcházení. V první části jsou představeny organizace a projekty monitorující státní cenzuru Internetu ve světě, dále autor popisuje metody odhalení nežádoucí komunikace a způsoby její filtrace. Následující část se zabývá základními principy obcházení Internetové cenzury a přiblížením jednotlivých technik. Dalším cílem je ověření metod obcházení Internetové cenzury v praxi. Pro měření si autor zvolil Čínskou lidovou republiku, kde zjišťoval stav cenzury a následně testoval úspěšnost metod pro obcházení cenzury. Prezentované výstupy slouží jako přehled technik státní cenzury Internetu a obcházecích metod a mohou být využity jako základ pro další práce.

Klíčová slova

státní cenzura Internetu, obcházení Internetové cenzury, Internetová cenzura v ČLR, svoboda Internetu

Abstract

This Master thesis deals with the subject of state censorship of the Internet. The objective of the thesis is to map the techniques of state censorship of the Internet and the methods of circumvention. In the first part, the author introduces organisations and projects that investigate state Internet censorship in the world. Subsequently, the author depicts inspection methods of communication and the Internet filtering techniques. The following part focuses on the principles of circumventing Internet censorship. It also describes various techniques of the circumvention of the censorship. Another objective of the thesis is to test the practical methods of circumventing Internet censorship. As a destination for the measurement the author has chosen People's Republic of China where he scrutinised the state of Internet censorship and tested the success rate of circumvention methods. The presented outputs serve as an overview of the techniques of the state censorship of the Internet and the methods of its circumvention, and can be used as a foundation for further work.

Keywords

state Internet censorship, Internet censorship circumvention, Internet censorship in China, freedom on the Internet

Obsah

Úvod **1**

Cíle práce **1**

Struktura práce **2**

Rešerše prací **2**

1 Státní cenzura Internetu **5**

1.1 Pojem cenzura **5**

1.2 Důvody státní cenzury **6**

1.3 Organizace mapující cenzuru **6**

1.3.1 OpenNet Initiative 7

1.3.2 Freedom House 7

1.3.3 Reportéři bez hranic 8

1.3.4 Ministerstvo zahraničí Spojených států amerických 9

1.4 Projekty mapující cenzuru **9**

1.4.1 OONI 9

1.4.2 Herdict 10

2 Techniky cenzury Internetu **11**

2.1 Místa aplikování cenzury **11**

2.1.1 Cenzura na straně klienta 11

2.1.2 Cenzura na straně serveru 12

2.1.3 Cenzura komunikace 12

2.2 Struktura komunikace **13**

2.2.1 Přenos dat 13

2.2.2 Vrstvy modelu TCP/IP 13

2.3 Způsob odhalení nežádoucí komunikace **18**

2.3.1 Stavová inspekce paketů 18

2.3.2 Hloubková inspekce paketů 19

2.3.3 Probabilistické metody 20

2.3.4 Aktivní metody 20

2.4 Manipulace ze strany cenzora **21**

2.4.1 Degradace kvality spojení 21

2.4.2 Podvržení obsahu 21

2.4.3 Blokování komunikace 23

3 Techniky obcházení cenzury Internetu ... 25
3.1 Klasifikace metod ... 25
3.1.1 Skrytí cíle ... 25
3.1.2 Skrytí obsahu ... 26
3.1.3 Skrytí odlišností ... 26
3.1.4 Autonomie ... 26
3.1.5 Využití specifik cenzurovacích systémů ... 26
3.2 Specifikace technik ... 27
3.2.1 Šifrování ... 27
3.2.2 DNS ... 29
3.2.3 Proxy servery ... 32
3.2.4 Anonymní peer-to-peer sítě ... 36
3.2.5 Transportní protokoly ... 37
3.3 Distribuce programů ... 41
3.3.1 Přímý kontakt ... 41
3.3.2 Zrcadlící servery ... 42
3.3.3 Email/IM ... 42
3.3.4 Sociální sítě ... 43
3.3.5 Ověření pravosti ... 43
4 Aplikování technik obcházení cenzury ... 45
4.1 Výběr země ... 45
4.1.1 Míra cenzury ... 45
4.1.2 Počet postižených obyvatel ... 45
4.1.3 Bezpečnost pohybu po cílové zemi ... 46
4.1.4 Výsledek výběru země ... 46
4.2 Metodika sběru dat ... 46
4.2.1 Zjištění stavu cenzury ... 47
4.2.2 Testování technik obcházení cenzury ... 47
4.2.3 Nástroje pro měření ... 50
4.3 Praktický test a výsledky ... 52
4.3.1 Stav cenzury ... 52
4.3.2 Obcházení cenzury ... 55
4.3.3 Shrnutí ... 58
Závěr ... 60

Zdroje **62**

Terminologický slovník **72**

Seznam obrázků **74**

Seznam tabulek **74**

***Příloha* A** **74**

***Příloha* B** **74**

***Příloha* C** **74**

Příloha A: Výběr testované země **75**

Příloha B: Metodika praktického test obcházení cenzury **76**

Příloha C: Výsledky praktického testu obcházení cenzury **79**

Příloha D: Soubory s výsledky praktického testu **82**

Úvod

Internet je médium umožňující předávání informací napříč světem v reálném čase. Se zvyšujícím se množstvím dat na něm dostupných rostou také tendence jeho kontroly, a tak je čím dál častěji předmětem státní cenzury.

Kybernetická bezpečnost, bující Internetová kriminalita a hrozba terorizmu. To vše jsou argumenty k prosazování kroků umožňujících důkladnější monitorování a také blokování Internetového obsahu. Reakcí je nevole ze strany obyvatel, kteří žádají svobodný přístup k informacím a právo na soukromí. Za těmito účely tak vznikají stále nové nástroje umožňující lépe blokovat a naopak zpřístupňovat webový obsah.

Rychlost vývoje technik pro obcházení cenzury by měla jít ruku v ruce se zvyšováním povědomí o těchto technikách, jelikož právě lidem jsou určeny a každý jedinec by měl mít právo volby svobodného přístupu k informacím. Toto je hlavní popud, který vidím v publikování práce na téma státní cenzury Internetu a jejího obcházení.

Cíle práce

Cíle této diplomové práce jsou celkem tři: Prvním cílem je zmapovat mechanismy určené k provádění státní Internetové cenzury. Druhým je popsat techniky, které slouží k obcházení státní Internetové cenzury. Posledním cílem je ověřit funkčnost technik pro obcházení v praxi v cenzurou postižené zemi.

Důraz je kladen na technickou stránku cenzury a jejího obcházení, zejména na metody aplikované na komunikační kanál tj. blokování komunikace, nikoliv cenzury na straně klienta či serveru. Ačkoliv je práce zaměřena primárně na státní cenzuru, popisované metody mohou být využity též v korporátní sféře a dalších oblastech, kde k cenzuře Internetu dochází.

Popisované techniky a mechanismy chci kategorizovat tak, aby práce mohla být využita nejen jako seznámení s technickou stránkou státní cenzury Internetu, ale i jako zdroj pro další práce, které budou moci na zvolené struktuře stavět a dále ji rozvíjet. Klíčem pro kategorizaci bude tedy představit obecný rámec zahrnující co nejvíce technik.

Struktura práce

Práce je rozdělena do čtyř částí, které postupně seznamují s problematikou státní cenzury Internetu. Nejprve se zaměřím na pojem cenzury Internetu, důvody jejího využívání a zmapuji organizace a projekty, které monitorují státní cenzuru Internetu ve světě.

Další část se zabývá jednotlivými technikami pro aplikování cenzury Internetu. V této části budu zkoumat metodiku blokování Internetové komunikace, zejména kde je možné ji zachytit a jakými metodami se detekuje. Následně popíši způsoby, jak lze nežádoucí komunikaci filtrovat.

Třetí sekce je věnována metodám obcházení státní cenzury. Nejdříve kategorizuji techniky obcházení, a pak se zaměřím na jednotlivé metody určené pro obcházení blokace ze strany cenzora. Závěrem sekce se podívám i na možnosti distribuce nástrojů v cenzurou postižených zemích.

Poslední část diplomové práce testuje využitelnost nástrojů pro obcházení cenzury v praxi. Nejprve provedu výběr vhodné země, ve které později bude probíhat měření jak současného stavu Internetové cenzury, tak nástrojů k jejímu obcházení. Výstup měření by měl odpovědět na otázky, zdali je v dané zemi možné se dostat k cenzurovanému obsahu a popřípadě jaký způsob obcházení je účinný.

Rešerše prací

Teoretický základ, na kterém bylo v této práci stavěno, se převážně skládá z vědeckých článků, technických specifikací vydaných respektovanými autoritami a dalších zdrojů od organizací zabývajících se svobodou Internetu a volným přístupem na něj. Jelikož práce čerpá z mnoha různých zdrojů, byl by výčet prací v této kapitole velmi rozsáhlý, proto jsem se rozhodl uvést pouze několik zdrojů, u kterých jsem shledal zásadní vliv na směřování či obsah práce. Pořadí prací odpovídá tematickému sledu jednotlivých částí práce.

Systemization of Pluggable Transports for Censorship Resistance

Khattak, Sheharbano, Simon, Laurent a Murdoch, J. Steven. Systemization of Pluggable Transports for Censorship Resistance. [Online] 1, 23. 12 2014. [Citace: 20. 04 2016.] http://arxiv.org/abs/1412.7448.

Tento článek poskytuje přehlednou kategorizaci technik Internetové cenzury a transportních protokolů určených k jejich obcházení. V článku jsou proti sobě postaveny jednotlivé mechanismy cenzorů proti existujícím transportním protokolům, které se snaží cenzuru obcházet. Autor tak názorně prezentuje slabosti jednotlivých protokolů a výhody, kterých by dosáhly, kdyby byly postaveny na modulárním základě a navzájem propojeny.

RFC

V částech této práce, kde představuji technické detaily Internetové komunikace či protokolů, jsem čerpal především z RFC (Request for Comments) specifikací vydávaných pod hlavičkou organizace Internet Engineering Task Force (IETF). Tyto dokumenty jsou sepsány odborníky z oboru a obsahují technické detaily ohledně fungování a organizace Internetu. V současnosti jsou uznávány jako *de facto* standardy, ačkoliv jsou vydávány pouze jako doporučení.

The Tor Project

Tor Project [online]. 2016 [cit. 2016-05-01]. Dostupné z: https://www.torproject.org

Primárním účelem The Tor Project je vývoj software k zachování anonymity na Internetu (Tor). Vedlejším produktem této činnosti jsou vědecké práce, dokumentace využívaných technik, měření a přidružené projekty, které se zabývají svobodou Internetu a pomáhají s mapováním prostředí Internetu a vymýšlením nových metod k zefektivnění softwaru a udržení Interentu anonymního, necenzurovaného a otevřeného pro všechny.

Zvláště užitečným zdrojem informacím je část webu zabývajících se výzkumem (research.torproject.org), která nabízí technické reporty ve formátu článků dokumentující fungování Toru a technik anonymizace. Dalším zdrojem informací je sekce dokumentace popisující využívané protokoly a blog (blog.torproject.org), která informuje o nových zjištěních z oblasti cenzury, Internetové svobody a vývoje Tor.

Understanding cryptography

Paar, Christof a Pelzl, Jan. *Understanding cryptography : a textbook for students and practitioners*. Berlin : Springer, c2010. ISBN 978-3-642-04100-6.

Kniha Understanding cryptography popisuje přehledným způsobem techniky symetrické a asymetrické kryptografie, šifry a jejich využití. Její obsah měl vliv zejména na porozumění šiforvacím mechanismům, které jsou základem obcházecích technik.

Internet Censorship in China: Where Does the Filtering Occur?

Xu, Xueyang , Mao, Z. Morley a Halderman, J. Alex. Internet Censorship in China: Where Does the Filtering Occur? *Proceedings of the 12th International Conference on Passive and Active Measurement.* 2011, stránky 133-142.

Autoři práce se v článku zabývají strukturou firewallu v ČLR a monitorují zásahy do komunikace v souvislosti s poskytovateli Internetového připojení. Tento článek byl jedním z těch, který mi sloužil k zorientování se v problematicke cenzury v ČLR a zvolení vhodné metody pro praktické měření.

1 Státní cenzura Internetu

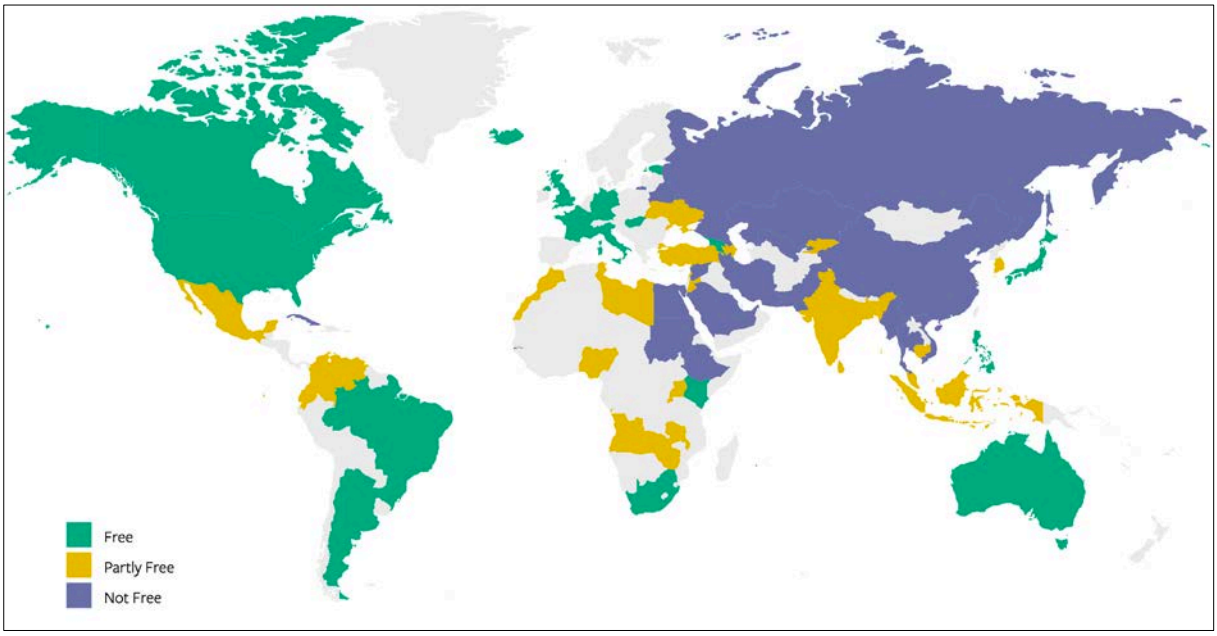
Tato kapitola slouží jako úvod do problematiky státní cenzury Internetu. Nejprve vysvětlím pojem cenzura a příčiny, které k zavedení cenzury vedou, a poté se zaměřím na organizace a projekty mapující cenzuru ve světě.

1.1 Pojem cenzura

S termínem *censor* se můžeme poprvé setkat již ve starém Říme roku 443 př. n. l., který byl ustanoven pro vysoké představitele státního úřadu. Ti měli za úkol odhadovat počet obyvatel (*census populi*) a na základě jejich majetku jim vyměřit práva a povinnosti (1) (2).

V moderním pojetí *cenzuru* chápeme jako systém či praktiky kontrolující zdroje informací (knihy, dopisy, fotografie atd.) za účelem odstranění obsahu, který je shledán urážlivý, nemorální či závadný pro společnost (3). Cenzura bývá prováděna *cenzorem*, což je subjekt zmocněný autoritou k provádění cenzury.

V kontextu této práce nás bude zajímat především *cenzura Internetu*, čili praktiky cenzury aplikovány na prostředí celosvětové elektronické informační a komunikační sítě Internet. Ta se běžně objevuje napříč světem, nejvýznamněji v Číně, Sýrii a Iránu, ale v mírnější formě se aplikuje i v Evropských zemích či ve Spojených státech (4).



Obr. 1.1: Mapa cenzury Internetu ve světě
Zdroj: (4)

Pro ilustraci je na Obr. 1.1 k nahlédnutí mapa cenzury Internetu v 65zemích světa za rok 2014, kterou vydala organizace Freedom House (viz 1.3.2). Škála od fialové (nesvobodné) přes žlutou (částečně svobodné) až po zelenou (svobodné) zobrazuje míru svobody Internetu v jednotlivých zemích.

To, že je státní cenzura Internetu aktuální téma i v České republice, dokazuje nedávno schválený Zákon č. 186/2016 Sb., o hazardních hrách, který s účinností od 1. ledna 2017 zavádí restrikce na přístup k Internetu vyžadované přímo od poskytovatelů připojení k Internetu. Tento zákon dle § 82 poskytovatelům na území ČR ukládá povinnost zamezit v přístupu k Internetovým stránkám uvedeným na seznamu stránek s nepovolenými Internetovými hrami vedeným Ministerstvem financí.

V České republice je cenzura v rozporu s legislativou na základě odst. 3 čl. 17 zákona č. 2/1993 Sb. - Listiny základních práv a svobod, která udává, že je cenzura nepřípustná. Obdobnou úpravu deklarující práva svobody projevu a přístupu k informacím můžeme nalézt například ve Všeobecné deklaraci lidských práv OSN (čl. 19).

1.2 Důvody státní cenzury

Míra státní cenzury se v jednotlivých státech liší stejně jako důvody pro její zavedení. Politický režim, vládnoucí garnitura, kultura, momentální ohrožení státu a nálada ve společnosti. To vše jsou faktory, které mají na cenzuru vliv. Mezi hlavní příčiny patří aspekty:

- sociální – cenzura zaměřená na gambling, drogy, alkohol, pornografii, projevy rasismu a jiných forem společensky nepřijatelného chování;
- politické – propaganda, politická manipulace, upevnění politické moci, omezení svobody slova;
- bezpečnostní – cenzura utajovaných informací, záležitosti ohrožující bezpečnost státu;
- ekonomické – ochrana duševního vlastnictví, boj proti Internetovému pirátství.

1.3 Organizace mapující cenzuru

V souvislosti s rostoucí mírou cenzury v jednotlivých zemích se začaly objevovat i organizace, které ji monitorují. Tyto organizace pravidelně vydávají reporty mající za úkol upozornit na stav cenzury obyvatele žijící v postižené zemi i zbytek světa.

1.3.1 OpenNet Initiative

OpenNet Initiative (ONI) vznikla spojením sil *Citizen Lab* (Munk School) z Torontské univerzity (angl. University of Toronto), *Berkman Center for Internet & Society* z Harvardovy univerzity (angl. Harvard University) a *SecDev Group* (Ottawa). Jejím hlavním cílem je analyzovat a odhalovat praktiky filtrování obsahu Internetu a dohledu (angl. surveillance). Svou činnost ukončila po více než dekádě svého působení na konci roku 2014, kvůli klesající aktivitě jednotlivých členů pod hlavičkou této iniciativy.

K získání věrohodných informací využívala kombinaci zpráv od státních představitelů týkajících se uplatňovaných pravidel cenzury a průzkumu prováděného experty a dobrovolníky v jednotlivých zemích. Pomocí programu vyvinutého speciálně pro tyto účely sbírali informace o blokování URL, IP adres, DNS, specifických HTTP odpovědích a hlavičkách, obsahu stránek a v určitých případech i cestě paketů. Tyto data následně analyzovali a snažili se vyčíslit celkový dopad na konkrétní zemi (5).

Iniciativa vydala edici tří knih, které dokumentují poznatky získané během svého aktivního působení, dále spravovala profily 74 vybraných zemí a informovala o stavu omezování Internetu v nich jak textovou formou, tak i v podobě přehledné mapy monitorující jednotlivé aspekty cenzury. Všechny získané informace jsou nadále k dispozici na webových stránkách.

Pro svoji činnost se ONI rozhodla popsat 4 hlavní témata Internetové cenzury (politické, sociální, konflikty/bezpečnost, Internetové nástroje), pro které v každé zemi určila míru filtrace pomocí následující stupnice (6):

- všudypřítomná (angl. pervasive) – filtrování charakteristické *hloubkou* (blokování velkého množství obsahu v dané kategorii), tak i *rozsahem* (blokování ve více kategoriích v rámci daného tématu);
- značná (angl. substantial) – filtrování mající *hloubku*, nebo *rozsah* (viz předchozí bod);
- selektivní (angl. selective) – úzce cílené filtrování blokující nízký počet specifických stránek;
- domnělá (angl. suspective) – zjištěny abnormality, které mohou být způsobeny filtrováním, ale nepodařilo se ho uspokojivě prokázat;
- nezjištěná (angl. no evidence) – ONI testování nezaznamenalo žádné známky blokování stránek.

1.3.2 Freedom House

Freedom House (FH) byl založen roku 1941 skupinou amerických novinářů, podnikatelů, akademiků a dřívějších vládních představitelů, mezi kterými byla například Eleanor

Roosvetlová či Wendell Wilkie, jako odpověď na hrozby totalitních režimů. Jedná se o organizaci, jejíž hlavní úkol spočívá v šíření svobody a demokracie po světě.

FH vydává pravidelně reporty mapující stav svobody tisku, demokracie, politických a občanských práv a od roku 2009 také svobodu na Internetu, kterou monitoruje již v 65 zemích světa. Pro výzkum využívá jak zaměstnance z FH, tak převážně výzkumníky a poradce z řad akademiků, bloggerů, novinářů a IT profesionálů, kteří působí v cílové zemi (7).

Systém hodnocení funguje na principu dotazníku, který se skládá z 21 otázek a téměř 100 podotázek, které jsou rozděleny do následujících kategorií (7):

- překážky v přístupu – infrastrukturní a ekonomické bariéry v přístupu, právní a vlastnická kontrola nad poskytovateli Internetového připojení a nezávislost regulátorů;
- omezování obsahu – právní regulace obsahu, filtrování a blokování webových stránek, auto-cenzura, aktivita a diverzita online médií a míra využití Internetových nástrojů k občanským protestům;
- porušení uživatelských práv – dohled, soukromí a omezování svobody slova na Internetu a aktivity jako jsou zatýkání, přílišné právní obtěžování a kybernetické útoky.

Jednotlivé odpovědi jsou bodově ohodnoceny a celkové výsledky dodatečně konzultovány na regionálních mítincích, popřípadě rozebírány po telefonu s jednotlivými výzkumníky. Zemi je nakonec přiřazeno finální skóre na škále od 0 do 100 a přístupu k Internetu uděleno následující hodnocení (7):

- svobodný (angl. Free) – 0-30 bodů;
- částečně svobodný (angl. Partly Free) – 31-60 bodů;
- nesvobodný (angl. Not Free) – 61-100 bodů.

Jednotlivé státy jsou zobrazeny na mapě podle jejich skóre nebo je možné se podívat na jejich podrobný profil vysvětlující dosažené výsledky.

1.3.3 Reportéři bez hranic

Reportéři bez hranic (franc. Reporters sans frontières, RSF) je nevládní organizace založena roku 1985 v Montpellier (Francie). Mezi její cíle patří: monitorování útoků na svobodu informací a veřejné odsuzování těchto útoků; spolupráce s vládami na odstranění cenzury a

zavedení zákonů posilující svobodu informací; morálně a finančně pomáhat stíhaným novinářům a jejich rodinám a nabízet materiální pomoc válečným obětem (8).

RSF se zvlášť zaměřuje na svobodu Internetu. Konkrétně pod hlavičkou Nepřátelé Internetu (angl. Ennemies of the Internet) označuje ty země, kde v souvislosti s Internetem dochází k cenzuře, dohledu, věznění či k manipulaci s informacemi. Svoje označení země dostane na základě zjištění novinářů a jejich článků.

Každý rok se mění zaměření sekce Nepřátelé Internetu a přibývají nebo ubývají sledované kategorie. V posledním roce (2015) mapa zemí nebyla aktualizována, namísto toho se RSF zaměřila na zpřístupnění blokovaných webových stránek v jedenácti postižených zemích (9).

1.3.4 Ministerstvo zahraničí Spojených států amerických

Úřad pro demokracii, lidská práva a práci pod Ministerstvem zahraničí Spojených států amerických každoročně vydává hlášení o stavu lidských práv ve státech mimo USA (Country Reports). Součástí tohoto hlášení je i pasáž zabývající se přímo svobodou Internetu (Internet Freedom) v jednotlivých státech.

1.4 Projekty mapující cenzuru

Projekty pro mapování cenzury se snaží sbírat aktuální informace o stavu cenzury Internetu přímo od obyvatel. Uživatelé Internetu testují přístup na webové stránky či služby a tyto údaje jsou zaznamenávány na serveru projektu a zpřístupněny veřejnosti. V následujících kapitolách popíši projekty OONI a Herdict.

1.4.1 OONI

Open Observatory of Network Interference (OONI) je volně šiřitelný open-source program pod záštitou společnosti The Tor Project, Inc, který každému, kdo si program nainstaluje, umožňuje anonymně přispět do celosvětové statistiky cenzury Internetu. Posbíraná data jsou volně přístupná ve strojově čitelném formátu pro další analýzu.

OONI provádí testy TCP, DNS, HTTP a TLS spojení a porovnává odpovědi se známými očekávanými hodnotami. Ke svému běhu využívá program Tor (The Onion Routing), který slouží k obcházení Internetové cenzury a anonymizaci spojení (viz 3.2.3.3), programu OONI tak umožňuje získávat necenzurované verze odpovědí (10).

1.4.2 Herdict

Herdict je webová aplikace vytvořená v The Berkman Center for Internet and Society (Harvardova Univerzita), která umožňuje na základě uživatelských příspěvků reportovat blokované stránky. Dle těchto hlášení je následně tvořena mapa, která zobrazuje míru blokování v jednotlivých zemích. Data jsou volně přístupná přes Herdict API a v podobě CSV souborů.

Princip Herdict funguje tak, že každý uživatel může navrhnout webovou stránku, kterou shledá za důležitou a měla by proto být přístupná z celého světa. Uživatelé, kteří chtějí přispět svými informacemi, poté stránku společně s dalšími navrženými otestují a následně reportují do Herdict (přes email, Twitter, webový formulář či doplněk v prohlížeči), zdali se stránka zobrazuje v pořádku. Herdict tyto data zpracuje a zpřístupní ostatním výzkumníkům k analýze.

2 Techniky cenzury Internetu

Tato kapitola se zabývá technikami aplikovanými při provádění cenzury Internetu. Nejprve popíši místa, kde je cenzura nejčastěji realizována, následně vysvětlím principy síťové komunikace a představím metody sloužící k odhalení nežádoucí komunikace. Ke konci kapitoly pak shrnu způsoby, jakými je cenzor schopen komunikaci omezit.

2.1 Místa aplikování cenzury

V následujících odstavcích jsou popsána místa, kde je možné odhalit a filtrovat nežádoucí Internetovou komunikaci. Od bodu aplikování cenzury se pak dále odvíjí způsoby jejího možného blokování. Tato práce se zaměřuje primárně na cenzuru komunikace (viz 2.1.3), avšak pro celkovou představu je důležité si představit i ostatní způsoby.

2.1.1 Cenzura na straně klienta

Klientem je myšlena klientská stanice a cenzura probíhá pomocí aplikace pod správou cenzora, která je na ni nainstalována. Instalace může probíhat vědomě, například na základě státního nařízení, či nevědomě ve formě *malware*. Jakmile cenzor dostane kontrolu nad systémem, může pak (11):

- upravovat příchozí data předtím než dorazí k cílové aplikaci,
- vyhledávat zakázaná klíčová slova v dokumentech či komunikaci a reportovat je,
- blokovat komunikaci s nežádoucími servery/doménami,
- mazat cenzuře nevyhovující soubory.

V minulosti jsme se mohli s tímto přístupem setkat v ČLR v podobě programu Green Dam, který byl instalován na běžné domácí počítače pod záminkou ochrany dětí před závadným obsahem. Kromě proklamovaného ovšem dále blokoval i politicky a nábožensky nežádoucí obsah (12). Dalším příkladem je Children's Internet Protection Act (CIPA) vyžadující ve vybraných školách a knihovnách ve Spojených státech instalovat software filtrující potenciálně nebezpečný obsah.

2.1.2 Cenzura na straně serveru

Odlišný přístup je zaměřit se na zdroj informací, kde jsou data uložena, čili na samotné servery. Serverů je zákonitě menší počet než klientů a zásah do jejich fungování nemusí být pro klienty patrný, což činí tuto formu pro cenzora snazší, méně nápadnou a tedy lákavější. Specializovaný software nainstalovaný na server poté může (11):

- upravit a mazat data poskytovaná serverem,
- zakázat přístup na server,
- blokovat komunikaci obsahující zakázaná klíčová slova.

Podobné tendence je možné pozorovat například na čínském mikro-blogovacím portálu Sina Weibo, kde je patrné systematické mazání příspěvků a blokování vyhledávání obsahující nežádoucí klíčová slova (13).

2.1.3 Cenzura komunikace

Dosud jsem se zabýval stranou klienta žádající obsah a serveru, který ho poskytuje. Mezi těmito subjekty se nachází komunikační kanál, který bývá často pod správou cenzora. Tato forma kontroly má výhodu v tom, že ovládáme-li páteřní sítě, máme pak přehled nad většinou komunikace, která na Internetu v daném regionu probíhá. Tímto způsobem je také možné kontrolovat síťový provoz překračující hranice státu a určit tak pravidla, co smí a co naopak nesmí být přeneseno.

Ovládáme-li komunikační kanály, můžeme (11):

- kontrolovat obsah procházejících dat[1],
- blokovat komunikaci s nežádoucími servery/doménami,
- blokovat komunikaci zakázaných protokolů.

Tato forma cenzury je využívaná zejména díky její možnosti centralizované správy a vysoké efektivitě. Příkladem může být tzv. Velký čínský firewall (angl. Great Firewall of China), který se stará o filtrování čínského webového provozu (14), či Cleanfeed provozovaný na straně

[1] Nešifrovaná komunikace

poskytovatelů Internetu ve Velké Británii (15) a Kanadě (16), který má za úkol blokovat přístup na stránky s dětskou pornografií.

2.2 Struktura komunikace

Pro lepší porozumění technikám cenzury je nutné si nejprve představit principy síťové komunikace. V prostředí Internetu dominuje vrstevnatý model TCP/IP, který nám udává strukturu přenášených dat a využívané protokoly. Nejprve popíši systém přenosu dat a poté charakterizuji jednotlivé vrstvy protokolu TCP/IP.

2.2.1 Přenos dat

Komunikace na Internetu je zajišťována pomocí zpráv, které si mezi sebou jednotlivé uzly vyměňují. Jelikož celá zpráva může být příliš velká a při její ztrátě a opakovaném vysílání by došlo k přílišné zátěži sítě, bývají zprávy rozděleny do menších datových bloků. Tyto bloky jsou dále obaleny hlavičkami, které obsahují informace pro doručení dat cílové stanici napříč Internetem.

Pří pohybu dat sítí pak síťové prvky, které zajišťují doručení dat po síti (směrovače, přepínače atd.), nemusí číst veškerý obsah, ale postačí, aby si přečetli příslušné hlavičky a na základě nich vyhodnotil, kam paket poslat dál. Tento princip je výhodný jak z hlediska paměťové a výpočetní náročnosti, kdy si síťový prvek nemusí zpracovávat celý paket, ale také z hlediska programového vybavení, jelikož síťovému prvku postačí mít nástroje pro zpracování pouze určitých hlaviček a lze tak optimalizovat jeho výpočetní výkon a zaměření prvku.

Obsah jednotlivých hlaviček závisí na vrstvě, ve které byly vytvořeny, a následně na zvoleném protokolu. Mluvíme-li o Internetu, tak nás bude zajímat především vrstevnatý model TCP/IP, který logicky člení protokoly v závislosti na jejich funkci do jednotlivých vrstev (viz dále).

2.2.2 Vrstvy modelu TCP/IP

Model TCP/IP definuje celkem čtyři vrstvy. Každá z nich slouží k jinému účelu a využívá rozdílné protokoly. Jedná se o vrstvu síťového rozhraní, síťovou, transportní a aplikační. Na

Obr. 2.1 je patrné jejich pořadí a jakým způsobem dochází k postupnému obalování přenášených aplikačních dat.

Pořadí vrstev je fixní, stejně tak jako velikost hlaviček a polí v nich uložených. Při přijetí datového segmentu dochází k sekvenčnímu čtení a postupnému rozbalování vrstev. Jednotlivé vrstvy jsou blíže popsány níže.

Obr. 2.1: Zapouzdření aplikačních dat na vrstvách TCP/IP

Zdroj: autor

2.2.2.1 Vrstva síťového rozhraní

Vrstva síťového rozhraní (angl. link layer) je nejnižší vrstva v rámci modelu TCP/IP a zajišťuje přístup k fyzickému přenosovému médiu, adresování v rámci *přímo připojené sítě* a bezchybný přenos bloků dat. Datové bloky přenášené mezi vrstvami síťového rozhraní se nazývají *rámce* (17). Jejich struktura závisí na zvoleném protokolu.

Na Obr. 2.2 je vidět formát rámce u protokolu Ethernet. Pole preambule a SFD (Start Frame Delimiter) slouží k synchronizaci síťových zařízení a k označení začátku rámce. Následují pole s cílovou a zdrojovou adresou rámce, která je v rámci Etnernetu určena 64bitovým identifikátorem nazývaným MAC adresa. Další pole obsahuje doplňující informace o přenášených datech a za ním jsou odvysílána samotná data. Celý rámec je ukončen sekvencí FCS (Frame Check Sequence), která slouží k ověření správnosti přenesených dat (18).

Preambule	SFD	Cílová adresa	Zdrojová adresa	Délka / Typ	*DATA*	FCS

Obr. 2.2: Formát Ethernetového rámce

Zdroj: (18)

MAC adresa je jednoznačný identifikátor, který je pro každé síťové zařízení unikátní, čili by teoreticky mohla být využita pro účely cenzury a to při blokování určitých zařízení či hledání původce komunikace. Hlavička rámce je ovšem během cesty Internetem mazána a znovu generována při každém průchodu zařízením propojující dvě rozdílné logické sítě (nejčastěji směrovače), kterých je po cestě rámce mnoho. To a také fakt, že uživatel je schopný svou MAC adresu maskovat, činí vrstvu síťového rozhraní pro účely cenzury nezajímavou.

2.2.2.2 Síťová vrstva

Síťová vrstva (angl. Internet layer) umožňuje posílat data napříč sítěmi. Datové bloky opatřené hlavičkou síťové vrstvy se nazývají *pakety*. Pro adresování se využívá identifikátor cílového hosta zvaný *IP adresa*, jejíž formát závisí na zvoleném protokolu (19). V současnosti se využívá převážně protokol IPv4, který je postupně nahrazován novou verzí 6 (IPv6), jejíž hlavní výhoda je ve větším množství potencionálních adres.

Na Obr. 2.3 je vidět základní hlavička protokolu IPv6, která může být doplněna o další rozšiřující hlavičky. Ty umožňují například určit cestu paketu sítí či rozdělit data do více paketů (fragmentování) atd. (20). Pro účely cenzury nás bude zajímat převážně *zdrojová a cílová adresa* popřípadě pole *maximum skoků* (angl. Hop Limit).

<table>
<tr><td>Verze</td><td>Třída provozu</td><td colspan="3">Značka toku</td></tr>
<tr><td colspan="3">Délka dat</td><td>Další hlavička</td><td>Max. skoků</td></tr>
<tr><td colspan="5">Zdrojová adresa</td></tr>
<tr><td colspan="5">Cílová adresa</td></tr>
</table>

Obr. 2.3: Hlavička protokolu IPv6

Zdroj: (20)

IP adresa v prostředí Internetu (někdy též nazývaná *veřejná adresa*) je jednoznačný identifikátor cílového uzlu napříč sítí Internet. Chceme-li komunikovat s určitým uzlem (např. webový server), musíme nejprve znát jeho IP adresu, která se uvede jako cílová adresa v hlavičce IP paketu, a až poté je možné navázat s ním komunikaci. O doručení paketu se starají směrovače, které mají informace o okolních sítích (tzv. směrovací tabulky), na základě nichž vybírají cestu paketu a snaží se ho přiblížit k cíli.

Skutečnosti, že napříč sítí Internet existuje jednoznačný identifikátor uzlu, využívá i cenzura. Díky tomu dokáže identifikovat poskytovatele nežádoucích služeb (viz 2.3.1.1) a následně s těmito informacemi naložit dle nastavených politik (viz 2.4).

Dále bych se chtěl zmínit o poli *maximum skoků*, které je obdobou TTL (Time to live) definovaného v IPv4. Toto pole obsahuje číslo, které je při každém přeposlání paketu sníženo právě o jedna, a v případě, kdy dosáhne nuly, tak je paket zahozen (20). Tento mechanismus slouží k ochraně proti zahlcení sítě při vzniku smyčky. Zároveň nám také udává přibližnou vzdálenost původce zprávy, čehož lze využít při zkoumání metod blokování komunikace. Pomocí hodnoty maximum skoků v paketu jsme schopni odlišit, kdy komunikujeme skutečně s cílovým serverem (hodnoty maxima skoků budou obdobné), a kdy naopak je do naší

komunikace zasahováno po cestě, což může být právě firewall pod správou cenzora. Této techniky bylo využito například při zkoumání technik cenzury v ČLR (21).

2.2.2.3 Transportní vrstva

Transportní vrstva (angl. Transport layer) zajišťuje komunikaci pro koncové aplikace. Od jednotlivých aplikací přijímá data, které předává síťové vrstvě, nebo naopak bloky dat skládá a předává aplikacím na aplikační vrstvě.

Pro komunikaci na transportní vrstvě se využívají převážně dva protokoly: UDP a TCP. UDP (User Datagram Protocol) se vyznačuje vyšší rychlostí díky nižší režii, nevýhodou je, že nezaručuje v jakém stavu a jestli vůbec budou data doručena. Oproti tomu TCP (Transmission Control Protocol) obsahuje mechanismy, které zajišťují spolehlivé doručení ve správném pořadí (19). Výběr protokolu vždy závisí na charakteru komunikace: pro zobrazování webových stránek se využívá protokol TCP, který zaručuje jejich kompletní načtení, naopak pro přenos dat v reálném čase (např. VoIP) se použije rychlejší UDP, jelikož plynulost a rychlost odezvy převáží nevýhody v podobě možné občasné ztráty dat.

Porovnáme-li hlavičku UDP (viz Obr. 2.4) a TCP protokolu (Obr. 2.5), je patrný rozdíl ve velikosti, což je daň za režii, která je s TCP zprávami spojena. Z hlediska cenzury je důležitý hlavně *cílový port*, podle kterého lze odhadnout účel komunikace. Prvních 49 152 portů (0-49 151) spravuje IANA (Interent Asigned Number Authority), která přiřazuje službám porty a protokoly, na kterých by měly poslouchat (22). Standardně je například port 80 (tcp i udp) určen pro http komunikaci a lze tedy předpokládat již na této vrstvě, že účelem komunikace směřující na port 80 je zobrazení webových stránek.

Zdrojový port	Cílový port
Délka	Kontrolní součet

Obr. 2.4: Hlavička protokolu UDP

Zdroj: (23)

Zdrojový port								Cílový port
Číslo sekvence								
Potvrzený bajt								
Offset dat	Rezervováno	URG	ACK	PSH	RST	SYN	FIN	Okénko
Kontrolní součet								Urgent Pointer
Volby								Výplň

Obr. 2.5: Hlavička protokolu TCP

Zdroj: (24)

Znalost TCP protokolu lze využít cenzorem nejen k identifikování nežádoucí komunikace, ale i k jejímu blokování. K tomuto účelu slouží mechanismy ukončení komunikace, více v kapitole 2.4.3.2.

2.2.2.4 Aplikační vrstva

Nejvyšší vrstva v rámci modelu TCP/IP je aplikační vrstva, která předává aplikační zprávy koncovým programům, čili lze říci, že v této vrstvě už pracujeme se samotným obsahem datových bloků. Protokoly využívané na této úrovni jsou například HTTP, FTP, SMTP, ale i TLS/SSL, které zajišťují šifrování komunikace. Formát a obsah zpráv se výrazně liší v závislosti na charakteru aplikačního protokolu.

Na Obr. 2.6 je příklad HTTP GET žádosti zobrazené pomocí programu Wireshark. Z výstupu je patrné jakou webovou stránku jsem se pokoušel zobrazit, zároveň je vidět použitý webový prohlížeč, nastavený jazyk i operační systém a jeho verze. Stejné údaje si mohou zobrazit všechny síťové prvky na Internetu, přes které je moje komunikace směrována.

```
▽ Hypertext Transfer Protocol
  ▽ GET / HTTP/1.1\r\n
    ▽ [Expert Info (Chat/Sequence): GET / HTTP/1.1\r\n]
        [GET / HTTP/1.1\r\n]
        [Severity level: Chat]
        [Group: Sequence]
      Request Method: GET
      Request URI: /
      Request Version: HTTP/1.1
    Host: www.vse.cz\r\n
    Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8\r\n
  ▽ Cookie: _ga=GA1.2.1600516749.1409311648\r\n
      Cookie pair: _ga=GA1.2.1600516749.1409311648
    User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_11_5) AppleWebKit/601.6.17 (KHTML, like Gecko) Version/9.1.1 Safari/601.6.17\r\n
    Accept-Language: en-us\r\n
    Accept-Encoding: gzip, deflate\r\n
    Connection: keep-alive\r\n
    \r\n
```

Obr. 2.6: HTTP GET žádost o stránku www.vse.cz

Zdroj: autor

Z příkladu výše je jednoduché si odvodit využitelnost těchto dat pro cenzuru. Analýza na této vrstvě je prováděna pomocí metod *hloubkové inspekce paketů* (viz 2.3.2). Přenášené aplikační

zprávy je možné skrýt pomocí šifrování (viz 3.2.1.1), ale ani tento způsob není stoprocentní. Například v případě HTTPS (HTTP přes TLS/SSL) je možné v odesílaných žádostech (Client Hello) stále nalézt název webových stránek (pole Server Name) a podle toho komunikaci identifikovat.

2.3 Způsob odhalení nežádoucí komunikace

Pokud se cenzor nerozhodne pro odstřižení celého Internetu, musí mít metody, jak identifikovat nežádoucí komunikaci či pokusy o obcházení cenzury, na které pak následně reaguje pomocí technik Internetové cenzury (viz 2.4). Hlavní faktory, které je třeba zohledňovat je jejich *náročnost* a *spolehlivost*.

Náročnost je dána množstvím systémových prostředků nutných k vykonání algoritmu. Musíme mít na paměti, že kontrolovaných dat je obrovské množství a nechceme-li dopustit snadnou odhalitelnost cenzury či zhoršený přístupu k Internetu pro běžné obyvatele, musíme se snažit o udržení co nejnižšího zpoždění (angl. latency) na lince. Druhým faktorem je *spolehlivost*, která naopak určuje přesnost odhalování nežádoucí komunikace. Spolehlivost se vyznačuje co nejnižším počtem falešných detekcí, ať už pozitivních či negativních, při zachování co nejvyššího počtu správných detekcí. Náročnost a spolehlivost často stojí proti sobě a je tak nutné určit priority nebo najít efektivní algoritmus zohledňující oba faktory.

V následujících odstavcích představím metody odhalování nežádoucí komunikace. Začnu od té nenáročné, kterou reprezentuje metoda stavové inspekce paketů, po ty náročnější jako je hloubková inspekce paketů a probabilistické a aktivní metody.

2.3.1 Stavová inspekce paketů

Stavová inspekce paketů (angl. Stateful Packet Inspection, SPI) pracuje s hlavičkami paketů na první až třetí vrstvě TCP/IP. SPI se běžně využívá na firewallech, které jsou zvyklé na velký průtok dat.

Z hlediska cenzury nás budou zajímat především vrstvy síťové a transportní. Struktura a význam paketů byl popsán v předchozí kapitole (viz 2.2.2), nyní popíši, jak se tyto informace využívají pro odhalení nežádoucí komunikace.

2.3.1.1 Síťová vrstva

Dle hlavičky paketu síťové vrstvy TCP/IP je možné zjistit především zdrojovou a cílovou IP adresu, protokol a velikost paketu. Na této úrovni je tedy možné odhalit pokus o spojení s konkrétním serverem a jeho přibližnou polohu (25). Jsou tvořeny tzv. blacklisty IP adres, na které se zapisují adresy serverů obsahující nežádoucí obsah či umožňující oslabit účinky cenzury.

Odhalování na základě IP adres je nenáročná metoda běžně využívaná *paketovými filtry*, ale je možné ji využít až po prvotním identifikování závadných serverů. Rizikem je také spolehlivost, jelikož za jednou IP adresou se může skrývat více webových stránek či služeb.

2.3.1.2 Transportní vrstva

Na úrovni transportní vrstvy je možné od sebe odlišit komunikace jednotlivých služeb a dokonce podle portu odhadnout o jakou službu se jedná (viz 2.2.2.3). Ačkoliv podvržení portu není nikterak složité, tak u veřejně přístupných serverů bývají zachovány standardní porty, kvůli jejich snadné dostupnosti. Dále máme přehled o stavu této komunikace, čili jsme schopni rozeznat, zda se jedná o začínající komunikaci, nebo již navázanou.

Je-li cenzorem kupříkladu zakázána šifrovaná komunikace, lze už na této úrovni vytipovat porty (např. 443, 465, 993, 995) či protokoly (např. tcpcrypt (26)) sloužící výhradně k šifrované komunikaci a s nimi pak zacházet odpovídajícím způsobem.

2.3.2 Hloubková inspekce paketů

Hloubková inspekce paketů (angl. Deep Packet Inspection, DPI) je metoda, která se zabývá zkoumáním datové části paketu, nejen jeho hlaviček, a operuje tedy na nejvyšší *aplikační vrstvě* TCP/IP (viz 2.2.2.4). Oproti předchozím SPI metodě, dokáže blíže analyzovat význam přenášené zprávy.

Každý paket podrobený DPI prochází sadou pravidel hledajících shodu s předem určenými *klíčovými slovy* či *signaturami protokolů*. Příkladem může být vyhledávání názvů nežádoucích domén v HTTP GET žádosti (viz Obr. 2.6) či identifikace protokolu pomocí specifických řetězců pro *handshake*. S množstvím pravidel narůstá také náročnost celého procesu, čili při porovnání s předchozími metodami se jedná o výpočetně náročnou záležitost (27). Proti sobě vždy stojí snaha dosáhnout co nejnižšího počtu pravidel při co nejvyšší spolehlivosti.

DPI je běžně využíváno při ochraně před viry, spamem či ohrožením počítačové sítě, ale slouží také pro odposlechy či cenzuru Internetu. Technické provedení bývá realizováno pomocí zrcadlení portů či optického spliterru tak, aby nedocházelo ke zpomalení síťového provozu (28). Samotná analýza probíhá na výkonných počítačích určených výhradně pro tento účel.

2.3.3 Probabilistické metody

Probabilistické metody nabízejí alternativu k využívání manuálně definovaných sad pravidel a rozšiřují schopnosti DPI. Pomocí analýzy statistické distribuce klíčových vlastností jako je délka, množství, intervaly mezi doručením a entropie sady paketů, je možné s určitou přesností klasifikovat daný protokol (29).

V učícím režimu nejprve nasloucháme dostatečně dlouhé konverzaci protokolu, který chceme do budoucna identifikovat, a zaznamenáváme si vlastnosti spojení. Poté na reálných datech za pomocí metod probabilistické klasifikace (Bayesovská klasifikace, neuronové sítě atd.) jsme schopni s určitou pravděpodobností zjistit, zdali sledovaná komunikace náleží našemu hledanému protokolu.

2.3.4 Aktivní metody

Předchozí metody jsou zaměřeny na pasivní detekci aplikovanou na tok dat procházející přes kontrolovaný síťový prvek bez jakéhokoliv zásahu do komunikace. Kromě pasivních metod existují metody aktivní, které umožní získat další informace o spojení pomocí zasílání dodatečných paketů vytvořených cenzorem.

2.3.4.1 Zásahy do komunikace

Jednou z aktivních metod je zasahování do již probíhající komunikace, kdy cenzor dokáže vytěžit znalosti sémantiky protokolu a narušením komunikace vyvolat sled zpráv, které odhalí program či protokol vyšší vrstvy skrývající se za komunikací. Zastavíme-li po cestě například pakety z probíhající TCP komunikace, očekávaným chováním od TCP protokolu je pokus o znovu zaslání dat.

2.3.4.2 Active probing

Monitorovaná komunikace je často šifrovaná, a tak cenzor bez znalosti klíče či slabiny šifrovacího protokolu nemůže s určitostí využít než zdrojové a cílové IP adresy, porty,

popřípadě části nešifrované komunikace sloužící k navázání komunikace. V takové situaci je možné aplikovat *active probing* a zachovat se jako klient, který se pokouší připojit k cílovému serveru. Tímto způsobem je například možné zjistit obsah cílové webové stránky či odhalit servery sloužící k obcházení cenzury (30).

2.4 Manipulace ze strany cenzora

Reakce cenzorů na odhalení nežádoucí komunikace se různí, a tak se můžeme setkat s pouhým pasivním sledováním a sbíráním dat až po aktivní blokaci. Záleží především na nastavené politice cenzury, která určuje, jak bude cenzura patrná a efektivní. V této kapitole jsou popsány metody degradace kvality spojení, podvržení obsahu a blokování komunikace, kterými se cenzoři nejčastěji snaží zabránit přístupu k nepovoleným informacím (11).

2.4.1 Degradace kvality spojení

Degradace kvality spojení patří mezi obtížně odhalitelné metody, obzvlášť jeli používána sporadicky. Ze strany obyvatel se může jevit jako chyba na straně poskytovatele služeb či lokální přetížení sítě a nemusí si ji vůbec spojovat s cenzurou, přesto působí demotivačně a může návštěvníky serverů odradit od jejich navštěvování.

Pro degradaci kvality spojení mohou být využívány různé formy snižování priority komunikace, zahazování či pozdržování paketů. Tato metoda byla v minulosti zaznamenána v Iránu v období politických nepokojů a cílila především na zahraniční média (31).

2.4.2 Podvržení obsahu

Podvržení obsahu je invazivní metoda mající za cíl zobrazit uživatelům jiné data než ty, které pocházejí z legitimního serveru. Jednou z cest je přesměrování komunikace na servery pod správou cenzora, kde se může nacházet například varovná zpráva o přistupování na nepovolený server, nebo může docházet k *hijackingu* cílového serveru, kdy cenzor provozuje server na první pohled nerozeznatelný od legitimního. Při hijackingu může docházet jak k podvržení nežádoucího obsahu na serveru, tak může být využit i k jiným účelům než je cenzura například k pasivnímu sběru informací o návštěvnících.

2.4.2.1 Podvržení DNS

DNS informace slouží k překladu *doménových jmen* na IP adresy. V prostředí Internetu nám umožňují využívat pro lidi snáze zapamatovatelné textové řetězce (např. www.vse.cz), které jsou pak převáděny na číselné kódy (např. 146.102.16.4) zajišťující správné doručení v počítačové síti. Podvržení DNS je běžně označováno jako *DNS hijacking* a jeho princip spočívá v záměně DNS informací, které mají za cíl nabídnout uživateli podvrženou IP adresu namísto té patřící legitimnímu serveru. Při úspěšném podvržení se pak uživatel dostane na falešný server, aniž by to bylo z *doménového jména* patrné.

K podvržení může docházet přímo na DNS serverech nebo zasahováním do UDP datagramů, které slouží k předávání DNS záznamů, na kanálech pod správou cenzora. Podvržení DNS bylo zaznamenáno například v ČLR, kde v případech, kdy *doménové jméno* obsahovalo zakázané klíčové slovo, došlo k záměně DNS informací (32).

2.4.2.2 Podvržení směrovacích informací

Jednotlivá data přenášená přes Internet musí nést informace o jejich zdroji a destinaci tak, aby mohla být doručena napříč sítí. Síťové prvky zpracovávají tyto informace a podle přednastavených pravidel se snaží data přiblížit k cíli. Narušením těchto elementů má za následek komunikaci s jiným než zamýšleným cílovým serverem, či úplnou ztrátu dat, která během stanoveného limitu (TTL) nebudou schopna dorazit k cíli.

Cenzor je díky těmto technikám schopen provést podvržení obdobné tomu popsanému v předchozí sekci o DNS (viz 2.4.2.1), kdy uživatel má pocit, že komunikuje s legitimním serverem, ale přitom je jeho komunikace po cestě přesměrována na falešné servery. Tato metoda (masquerade) je o to zákeřnější, že ani přesná znalost IP adres (nejen *doménového jména*), neumožní uživateli toto podvržení odhalit, jelikož IP adresy se budou shodovat s těmi, které patří legitimnímu serveru.

Další možností, kterou cenzor má, je zaměřit se plošně na určité poskytovatele služeb (např. YouTube, Facebook či ISP (Internet Service Provider) atd.), které své adresy sdružují pod *autonomními systémy* (AS). Bloky veřejně dostupných IP adres jsou rozděleny mezi jednotlivé AS, jejichž administrátoři je spravují a provádějí v rámci nich vlastní směrování. Nejčastěji jsou ve vlastnictví ISP, ale s rozmachem nadnárodních Internetových společností stoupá jejich využití i mezi jednotlivými Internetovými službami, které si tak zajišťují lepší

dostupnost v rámci Internetu. Hraniční směrovače AS si s ostatním vyměňují informace o jim známým AS a tímto způsobem je možné dostat dynamicky informace o adresách ve zbytku Internetu a také zajistit nalezení alternativní trasy pro doručení dat v případě výpadku části sítě. Toto uskupení ovšem cenzorovi nabízí možnost:

- hromadně podvrhnout destinace nežádoucích AS v rámci jím spravovaných AS,
- avizovat vlastnictví bloků adres, které nemají pod správou, ostatním AS a tím narušit směrování napříč Internetem,
- skrýt adresy patřící do jejich AS v distribuovaných směrovacích tabulkách a tím zamezit komunikaci s těmito adresami zbytku světa.

Podobné způsoby cenzury se vyskytly například v Pákistánu v roce 2008, kdy došlo k neautorizovanému vysílání adres náležících do bloku AS36561 patřící YouTube, což mělo za následek dočasnou nedostupnost serverů YouTube nejen v Pákistánu, ale i v dalších zemích (33). Ke skrytí adres a tedy zamezení komunikace se zbytkem světa docházelo například v Egyptě (34) či Libyi (35).

2.4.2.3 Podvržení toku dat

Při zpracovávaní toku dat cenzor nemusí být závislý na znalosti protokolu, aby do datové části mohl zasahovat. Obzvlášť pokud se jedná o nešifrovanou komunikaci, pak je možné při zvolení správné znakové sady obsah v čitelné podobě nahlížet a tím pádem i měnit.

V případě šifrované komunikace je podvržením toku dat možné zasahovat například do vyjednávání (angl. negotiating) pravidel komunikace a tím ovlivnit sílu šifrování budoucí komunikace nebo neustálými zásahy otrávit uživatele natolik, že šifrování vypne a bude využívat nešifrovaný protokol. Možností je skutečně mnoho a záleží pouze na zabezpečení protokolů a konkrétních cílech a kreativitě cenzora.

2.4.3 Blokování komunikace

Nejradikálnější a z pohledu cenzora nejjednodušší zásah je úplné blokování nežádoucí komunikace. Odhalí-li cenzor závadnou komunikaci (viz 2.3), je schopen ji okamžitě přerušit, dále pak může zařadit server či doménu na seznam blokovaných adres (angl. blacklist), který pak rozdistribuuje na další síťové prvky a zamezí této komunikaci do budoucna i ostatním uživatelům.

Seznamy blokovaných adres nejčastěji obsahují názvy domén, IP adresy a cílové porty, čili data běžně zpracovávaná při směrování paketů sítí, které umožňují rychlé odhalení (viz 2.3.1) a blokování bez dodatečné výpočetně náročné analýzy typu DPI (viz 2.3.2).

Z pohledu uživatele je blokování možné dosáhnout i pomocí metod popsaných v sekci *podvržení obsahu* (viz 2.4.2), ovšem z čistě technického hlediska jsou to odlišné způsoby. Blokování má za následek okamžité ukončení komunikace, bez dalšího pokračování v konverzaci či obelhávání uživatele.

2.4.3.1 Zahazování paketů

K zahazování (angl. drop) paketů dochází nejčastěji na směrovačích. Směrovače podle přednastavených pravidel prochází každý paket a najdou-li v sadě pravidel shodu (nejčastěji IP adresa či port) s právě zkoumaným paketem, provedou příslušnou akci – při zahazování DROP. Tímto způsobem dojde k okamžité ztrátě paketu a tudíž zamezení komunikace.

2.4.3.2 Ukončení komunikace

Při ukončení komunikace je snahou dosáhnout blokování pomocí znalosti komunikačního protokolu a jeho standardních syntaxí pro završení či přerušení konverzace. Ukončením komunikace tedy simulujeme snahu protistrany o završení konverzace. Respektujeme-li pravidla protokolu, můžeme se vyhnout jeho reakcím na ztrátu paketů po cestě, ke které dochází při zahazování paketů. Tato reakce je charakteristická pro *spolehlivé protokoly*, které garantují doručení dat, čili vyžadují potvrzení o přijetí od protistrany a v případě jeho absence se snaží data doručit znovu.

V praxi je možné se setkat s ukončováním komunikace například v ČLR u protokolu TCP. Je-li detekováno nepřípustné slovo v HTTP GET žádosti, dojde k okamžité odpovědi v podobě TCP RST paketů na obě strany komunikace a tím se ukončí spojení (21).

3 Techniky obcházení cenzury Internetu

Zavedení cenzury Internetu vedlo k přirozené obraně, která s sebou nesla využití již známých, ale i vývoj nových technologií, k jejímu prolomení. Sofistikovanost a efektivita jednotlivých nástrojů se liší v závislosti na technologické úrovni nástrojů využívaných cenzory a ocitáme se tak v závodech ve zbrojení, kdy obě strany se snaží vymyslet nové způsoby cenzury/prolomení dříve než oponent a dosáhnout tak co nejdelší dominance.

V této kapitole nejprve popíši jednotlivé cesty obcházení cenzury a následně blíže charakterizuji jednotlivé techniky. Závěr kapitoly bude věnován způsobům distribuce nástrojů k obcházení Internetové cenzury.

3.1 Klasifikace metod

Metody obcházení cenzury lze kategorizovat dle základních způsobů, kterými se snaží vyhnout detekci a blokování. Některé z komplexnějších metod mohou spadat do více skupin zároveň a potenciálně tak nabízet lepší schopnosti v obcházení cenzury. Kategorie popsané v této kapitole jsou: skrytí cíle, skrytí obsahu, skrytí odlišností, udržení co nejvyšší míry autonomie a využití specifik cenzurovacích systémů.

3.1.1 Skrytí cíle

Jelikož cenzor po detekci serveru, jenž poskytuje nežádoucí obsah, zvládne poměrně efektivně blokovat komunikaci s ním na základě jeho IP adresy (viz 2.4.2 a 2.4.3), tak existují metody, které se snaží tomuto filtrování zamezit. Možným způsobem je využití prostředníka (viz 3.2.3), jehož IP adresa je zobrazena v odchozích paketech namísto té blokované, a doufat, že danou adresu cenzor neoznačí jako závadnou. Tuto metodu je možné zefektivnit znásobením cílových IP adres zahrnutých do komunikace. Blokováním všech adres by cenzor omezil přístup k značnému množství nezávadného obsahu, čímž by mohl zvednout vlnu nevole (viz 3.2.5.1).

Alternativou je vyhnout se konvenčním metodám přístupu k serverům, které bývají nejvíce monitorovány cenzory, a zapojit se do sítě postavené na odlišných principech jako jsou peer-to-peer sítě (viz 3.2.4). Ty umožňují svým uživatelům poskytovat si navzájem obsah a sdílet služby mezi sebou. Tyto sítě přináší odlišné způsoby adresování a směrování a jejich obsah je přístupný pouze zapojeným členům.

3.1.2 Skrytí obsahu

Znepřístupnění obsahu pro nahlížení třetí stranou je cesta, jak se vyhnout detekci nežádoucí komunikace ze strany cenzora (viz 2.3.2), a tak předcházet možnému blokování. Pro tyto účely se využívá šifrování (viz 3.2.1), které pomocí kryptografických algoritmů zajišťuje, že obsah komunikace je možný nahlížet pouze účastníky, kteří vlastní klíč(e) k rozšifrování. Šifrování bývá standardní součástí většiny technik určených k obcházení cenzury.

3.1.3 Skrytí odlišností

Pokročilejší metody odhalování nežádoucího obsahu (viz 2.3.2, 2.3.3 a 2.3.4) umožňují identifikovat jednotlivé protokoly sloužící k obcházení cenzury na základě specifik, kterými se vyznačují. Možnou obranou proti odhalení jsou mimikry, čili zabalení konverzace do zvláštního protokolu (viz 3.2.5.2), který skryje veškeré odlišnosti od běžných povolených protokolů a znesnadní tak možnou identifikaci cenzorem.

3.1.4 Autonomie

Cenzor mívá kontrolu nad klíčovou infrastrukturou Internetu v dané zemí, pomocí které provádí detekci (viz 2.3) a zásahy do komunikace (viz 2.4). Logickým krokem při obcházení cenzury je vyhýbání se jakýmkoliv prvkům, které jsou pod správou cenzora. Příkladem cest k vyšší autonomii je využití alternativních ISP, změna způsobu připojení k Internetu (např. satelitní Internet), vytváření důvěryhodných sítí (viz 3.2.4) či spoléhat se jen na nezávislé servery (viz 3.2.2).

3.1.5 Využití specifik cenzurovacích systémů

Systémy zaměřené na provádění cenzury mají jako jakékoliv jiné systémy své silné a slabé stránky, kterých lze využít při jejich obcházení. Podaří-li se nám odhalit metodologii detekce a blokování je možné se zaměřit na jejich slabiny a alespoň dočasně tak prolomit cenzuru. Jednotlivé systémy cenzury jsou unikátní a v průběhu času se vyvíjejí, proto nelze vyjmenovat seznam univerzálně platných technik a i jejich povaha se značně liší. Může se jednat o speciálně vyvinuté *transportní protokoly* (viz 3.2.5.2), drobné úpravy ve firewallu (21) či využívání specifického *internetového protokolu* (např. IPv6 (36)).

3.2 Specifikace technik

V následující sekci se podrobně věnuji jednotlivým technikám, které mohou být využity při obcházení cenzury. Účelem není poskytnout výčet programů a protokolů odolávající cenzuře, ale spíše popsat principy jejich fungování. Postupně se podívám na techniky založené na šifrování, DNS, proxy serverech, anonymních peer-to-peer sítích a transportních protokolech.

3.2.1 Šifrování

Šifrováním se rozumí metoda záznamu informace (kódování), která slouží k utajení obsahu zprávy (37). Za pomocí kryptografických algoritmů jsme schopni zamezit nahlížení obsahu třetí straně. Bezpečnost šifrování vždy závisí na algoritmu, jeho implementaci, zvolených parametrech a míře dodržování bezpečnostních zásad nakládání s klíči.

3.2.1.1 Šifrování kanálu

Šifrování kanálu je cesta, jak zamezit komukoliv po cestě nahlížet na obsah, který je přenášen jednotlivými pakety. Na začátku konverzace dojde mezi serverem a klientem k navázání šifrovaného spojení a zbytek konverzace je pak šifrován pomocí klíče, který je znám pouze serveru a klientu.

Ve spojení s cenzurou Internetu se nejčastěji setkáváme s blokováním webových stránek, kde je pro zabezpečení využíván protokol HTTPS. Jedná se o běžné HTTP, které využívá šifrovací protokol SSL/TLS (38) k zabezpečení komunikace. Přistupujeme-li na server pomocí HTTPS protokolu, před samotným načtením obsahu dojde k *TLS handshake,* při kterém se server s klientem dohodnout na využívaných šifrovacích a hashovacích algoritmech, klient zkontroluje certifikát, jímž se server prezentuje, a poté dojde k předání *sdíleného klíče* (v šifrované podobě pomocí *veřejného klíče* serveru či *Diffie-Hellman algoritmu*) pomocí kterého je šifrován zbytek komunikace. Následně už komunikace probíhá standardním způsobem jako HTTP, akorát před odesláním požadavků či odpovědí je obsah zašifrován a na druhé straně opět rozšifrován.

SSL/TLS je hojně využívaný protokol pro zabezpečení komunikace v prostředí Internetu, ale mimo něj se můžeme setkat například s SSH (39), které slouží k šifrování komunikace především mezi Unix systémy, DTLS (40) pro datagramy, IPsec (41) či OTR (42) určený pro *instant messaging*.

V následující podkapitole se blíže podívám na VPN, které spadá do technik šifrování kanálu, ale díky svým vlastnostem lze využít i jako proxy server (viz 4.2.2.2). To činí z VPN populární techniku pro obcházení Internetové cenzury (43).

3.2.1.1.1 VPN

Virtuální privátní síť (angl. Virtual Private Network, VPN) je specifický případ využívání šifrování kanálu, kdy pomocí šifrovacích protokolů vytvoříme zabezpečené spojení mezi dvěma uživateli či sítěmi, které nám umožní přenášet data přes nedůvěryhodnou síť, jakou je například Internet (44).

K obcházení cenzury se využívá VPN pro spojení se zprostředkujícími servery, které jsou mimo dosah cenzora. Tímto způsobem je možné směrovat přes šifrovaný kanál buď veškerý síťový provoz (server pak slouží jako *výchozí brána*), nebo pouze zvolené požadavky (viz HTTP/SOCKS Proxy).

3.2.1.2 Šifrování obsahu

K šifrování nemusí docházet na úrovni kanálu, ale můžeme se rozhodnout pro zabezpečení jednotlivých zpráv. Tento způsob spoléhá na to, že příjemce a odesílatel se navzájem znají, a tak autor zprávy zabezpečí zprávu takovým způsobem, aby ji mohl otevřít pouze cílový příjemce.

Šifrování obsahu se často využívá u emailové korespondence a při předávání citlivých dokumentů. Na tomto poli dominují především OpenPGP (45) standardy založené na kombinaci asymetrické a symetrické kryptografie, kde odesílatel pomocí vygenerovaného *sdíleného klíče* zašifruje zprávu, tento klíč pak zašifruje s volně dostupným *veřejným klíčem* adresáta a společně odešle. Algoritmy využívané k asymetrické kryptografii (nejčastěji RSA, DSA, ElGamal, ECDH) mají zajistit, že jediný, kdo bude schopen *sdílený klíč* a potažmo celou zprávu rozšifrovat, je vlastník *soukromého klíče*. Na obdobném principu funguje například i protokol S/MIME (46) zaměřený primárně na mailovou komunikaci.

Šifrováním obsahu jsme schopni zamezit třetí straně přečtení obsahu zprávy, ale je třeba mít na paměti, že stále mohou být viditelna tzv. meta data (odesílatel, příjemce), které slouží k doručení zpráv napříč sítí.

3.2.2 DNS

Jediná cesta, jak se bránit před DNS servery poskytující podvržené adresy, je nespoléhat se na ně. Při volbě, jak se vyhnout cenzuře DNS, stojíme před vícero možnostmi, které se liší ve své spolehlivosti a pohodlnosti pro použití. Jejich výběr vždy závisí na míře DNS cenzury v dané zemi.

3.2.2.1 Lokální DNS databáze

Nejbezpečnější obranou proti podvržení DNS záznamů je vytvoření lokální databáze *doménových jmen* a jim přiřazeným IP adres. Většina operačních systémů (OS) pracuje s tzv. hosts soubory, které obsahují IP adresu a odpovídající *doménové jméno*. OS ve výchozím stavu přednostně nahlížejí na tento soubor a až potom se dotazují vzdáleného DNS serveru. Alternativou je nakonfigurování plnohodnotného lokálního DNS serveru a nahrání známých záznamů do něj, což nám přináší přehlednější správu, pohodlnější sdílení záznamů s ostatními a kontrolu nad zabezpečením v případě napojení na externí DNS servery (viz DNSSEC).

Provozování lokální DNS databáze je závislé na věrohodném zdroji IP adres, které do databáze zadáváme, a její pravidelné aktualizaci. Pro běžné procházení Internetu je *výhradní* využívání lokální DNS databáze skoro nemožné, jelikož bychom pro každou navštívenou doménu museli mít odpovídající záznam. Lokální databáze se ovšem jeví jako vhodný doplněk pro přístup k důležitým serverům při nestabilitě či nedůvěryhodnosti vzdálených DNS serverů.

3.2.2.2 Důvěryhodné DNS servery

Uživatelé si mohou na svém zařízení nastavit DNS servery, které pro ně budou překládat *doménová jména*. Základem je zvolit si takový server, který není pod kontrolou cenzora, čili při volbě záleží na důvěryhodnosti provozovatele. Pokud není známý žádný důvěryhodný provozovatel DNS, pak alternativou jsou otevřené DNS resolvery (OpenDNS, Google DNS, CZ.NIC ODVR), což jsou po Internetu volně dostupné DNS servery, které umožňují každému využít jejich služby.

Využívání alternativních DNS serverů bývá atakováno cenzory (47), jelikož samotné předávání záznamů probíhá formou UDP datagramů v nezašifrované podobě, což je velmi zranitelná metoda (viz 2.4.2.1). Prioritou je tedy zabezpečit spojení s DNS serverem a zajistit ověření záznamů. K tomu nám mohou pomoci buď běžné techniky obcházení cenzury, které za pomocí

šifrování, proxy, popřípadě transportních protokolů dostanou pakety za hranici ovládanou cenzurou, nebo protokoly vyvinuty přímo pro DNS jako jsou DNSSEC, DNSCurve či DNSCrypt (viz dále).

3.2.2.2.1 DNSSEC

DNSSEC je mechanismus *ověření původu* a *integrity* DNS záznamů založen na principu asynchronní kryptografie. Zúčastnění správci autoritativních DNS serverů své záznamy podepíší pomocí soukromých klíčů a následně předají hlavní veřejný klíč (KSK) nadřazené autoritě. Stejným způsobem podepíše své záznamy i nadřazená autorita a veřejný klíč předá zas o úroveň výš. Touto cestou vznikne řetězec důvěry, který zajistí, že nám stačí důvěřovat pouze jednomu prvku a poté si jsme schopni ověřit důvěryhodnost veškerých prvků podřízených (48).

V běžném provozu provádí validaci vzdálený DNS resolver, čili je důležité si dávat pozor při jeho volbě. I v případě, kdy jsme si jisti, že námi zvolený DNS server provádí validaci záznamů, je stále nutné zabezpečit kanál mezi námi a tímto DNS resolverem (48).

Alternativou je konfigurace vlastního lokálního DNS resolveru, který bude kontrolovat důvěryhodnost jednotlivých záznamů. Při provozování vlastního serveru je nezbytné bezpečnou cestou získat veřejný klíč kořenové zóny tzv. Root Key Signing Key (49), od kterého se odvíjí bezpečnost celého systému DNSSEC.

Přestože DNSSEC zažívá prudký nárůst a již jej podporuje naprostá většina TLDs (Top Level Domains), jeho implementace na běžných doménách SLDs (Second Level Domains) je stále pomalá, a proto u mnoha navštívených domén není možné ověřit důvěryhodnost DNS záznamů (50). Česká republika s TLD .cz je v tomto výjimka, kromě toho že v roce 2010 patřila mezi průkopníky DNSSEC, tak v současnosti s počtem 485 tis. podepsaných domén patří mezi nejlepších 5 TLD vůbec (50; 51).

Je třeba si uvědomit, že ani řetězec důvěry nezabrání tomu, že některá z vyšších autorit může podvrhnout údaje u jí podřízených zón, jelikož právě ona je nositelem veřejného klíče k ověřování. V praxi tak může dojít k tomu, že správce TLD domény (např. CZ.NIC pro .cz) se rozhodne změnit DNS údaje pro jemu (nebo jednomu z registrátorů) podřízenou doménu (např. www.vse.cz) a změnit adresu serverů, ze kterých se stránky budou načítat. Pomocí nově vygenerovaného soukromého klíče opatří DNSSEC záznamy a z hlediska validace se bude jevit vše v pořádku.

3.2.2.2.2 DNSCurve

Alternativou k DNSSEC je DNSCurve, které namísto ověřování záznamů volí cestu zabezpečení komunikace s autoritativními DNS servery, které poskytují zdrojové DNS záznamy. Veřejný klíč DNS serveru je uložen v NS záznamech, čili je jednoduše dostupný přes DNS dotazy (52).

Nevýhodou DNSCurve pro provozovatele DNS serverů rozmístěných po celém světě kvůli zachování dostupnosti je, že vyžaduje distribuci soukromých klíčů na všechny DNS servery. To by znamenalo umožnit autentizovat komunikaci i z politicky nestabilních zemí, což přináší potencionální riziko. Z tohoto důvodu správce *kořenové zóny* (angl. Root Zone) ICANN odmítá do budoucna nasadit DNSCurve (53).

Bezpečnost DNSCurve je stejně jako u DNSSEC závislá na důvěře autoritě, která je dotazované DNS zóně nadřazena, a vzhledem k tomu, že v současnosti na svých serverech DNSCurve nepodporuje žádný správce TLD (zdroj: autor) ani ICANN (53), tak možnosti ověření autenticity veřejného klíče ke koncovým DNS zónám se komplikují. Nemáme-li alternativní zdroj veřejných klíčů, jeho nasazení nám umožní pouze šifrovat komunikaci mezi DNS resolverem a autoritativním DNS serverem, čili ochranu před úpravou paketů po cestě, nikoliv ověřit autenticitu autoritativního DNS serveru.

3.2.2.2.3 DNSCrypt

Pro ochranu přenosového kanálu mezi koncovým uživatelem a DNS resolverem byl vytvořen specializovaný protokol s názvem DNSCrypt. Ten pomocí šifrování komunikace zajišťuje ochranu proti podvržení záznamů mezi koncovým uživatelem a resolverem.

DNSCrypt má svého klienta (dnscrypt-proxy), který je nainstalován na uživatelském zařízení, a nejčastěji na portu 443 (HTTPS) dochází ke komunikaci s DNS resolverem. Serverová část je pak zajišťována běžným DNS programem (BIND, Unbound atd.) a DNSCrypt aplikací (DNSCrypt-Wrapper), která se stará o komunikaci s DNSCrypt klienty. Seznam DNSCrypt serverů společně s jejich IP adresami, porty a veřejnými klíči se stahuje společně s klientskou částí DNSCrypt (dnscrypt-proxy) nebo je možné je ručně aktualizovat změnou příslušného CSV souboru. Mezi otevřené DNS resolvery, které v současnosti poskytují DNSCrypt, patří například OpenDNS či OpenNIC (54; 55).

3.2.3 Proxy servery

Proxy server je služba, která funguje jako prostředník mezi uživatelem a cílovým serverem a přináší přidanou funkcionalitu do komunikace. Jeho přínosy se mohou týkat například bezpečnosti, výkonosti či logování. Přesné vlastnosti záleží na konkrétním typu proxy serveru.

V oblasti obcházení cenzury je hlavním důvodem pro využití proxy serverů zamaskování cílové adresy, popřípadě i přenášeného obsahu. Jednotlivé typy se liší v efektivnosti, obtížnosti jejich odhalení a složitosti nasazení pro uživatele. Postupně představím jednoduché webové proxy, HTTP/SOCKS proxy a nakonec specializované programy pro obcházení cenzury, které se chovají jako lokální proxy servery.

3.2.3.1 Jednoduché webové proxy

Webové proxy přístupné přes webový prohlížeč, někdy též nazývané web-based proxy, nejčastěji fungují na principu webové stránky, kde se nachází formulář, kam stačí zadat URL stránky, kterou chceme zobrazit, a ta se nám poté načte bez toho, abychom my museli komunikovat s cílovým serverem. Po celou dobu co brouzdáme po Internetu, komunikujeme pouze s webovým proxy serverem a tato cílová adresa je také na všech odchozích paketech.

Jedná se o nástroj jednoduchý na použití, stačí pouze zadat webovou adresu proxy serveru do adresního řádku prohlížeče, nicméně cenzor je ho schopný po odhalení blokovat jako kteroukoliv jinou stránku, navíc volně přístupné proxy jsou často velmi pomalé a obsahují velké množství reklam. Dalším charakteristickým neduhem pro tento typ proxy je problém se zobrazováním dynamického obsahu a prvků Adobe Flash (43).

Na Internetu se můžeme setkat i s webovými proxy, které se snaží obejít snadné blokování ze strany cenzora pomocí rychlé změny doménových jmen a IP adres, na kterých je proxy provozovaná. Distribuce přístupových údajů je pak realizována alternativními kanály, jako jsou IM zprávy, emaily či osobním předání. Příkladem takové služby je DynaWeb specializující se primárně na ČLR (43; 56).

Datová část přenášených paketů obsahuje texty a obrázky z cílových stránek, čili je vhodné vybírat takové proxy, na které se přistupuje přes HTTPS a zkontrolovat si zda je certifikát stránky důvěryhodný. Přestože některé proxy slibují anonymitu a prohlašují, že neukládají

žádné údaje o uživatelích a navštívených stránkách, tak plnění těchto prohlášení nelze snadno ověřit a je třeba počítat s tím, že informace o provozu proxy serveru jsou provozovateli známi.

Specifickým typem proxy jsou některé portály, které primárně nemají za cíl sloužit jako webová proxy pro cenzurou postižené uživatele či anonymizační nástroj, ale charakter jejich služeb to umožňuje. Příkladem může být Google a jeho služby Cache, Reader a Translate, který byl před šesti lety velmi populární k obcházení cenzury například v Barmě či Ázerbájdžánu (43).

3.2.3.2 HTTP/SOCKS Proxy

HTTP/SOCKS Proxy jsou servery, které slouží k přístupu na Internet a běžně se využívají ve firemních sítích. Pro účely obcházení cenzury se tyto servery obvykle nachází mimo sféru kontroly cenzora a uživatel se k nim připojuje zadáním adresy a portu serveru v nastavení Internetového prohlížeče, popřípadě v systémovém nastavení. Veškerá HTTP(S) komunikace pak proudí přes tyto proxy namísto kontaktování cílových webových serverů napřímo.

Konfigurace na uživatelské stanici je jednoduchá a podporu pro proxy je samozřejmostí u všech majoritních operačních systému i webových prohlížečů (zdroj: autor). K blokování ze strany cenzora může docházet na základě cílové IP, při identifikování proxy, nebo pomocí DPI, pokud komunikaci s proxy nešifrujeme.

Na Internetu je možné najít volně přístupné proxy, které se ale často vyznačují vysokou latencí a nízkou spolehlivostí. Stejně jako u předchozího typu proxy, je nutno počítat s tím, že provozovatel proxy serveru má přehled nad navštěvovanými stránkami, ale oproti předchozímu typu má HTTP/SOCKS varianta výhodu ve chvíli, kdy přistupujeme na stránky zabezpečené pomocí SSL certifikátu (https). V takovém případě, za předpokladu, že je certifikát stránky ověřen důvěryhodnou autoritou, jsme schopni před provozovateli proxy serverů skrýt obsah navštěvovaných stránek (nikoliv servery samotné).

V následující podkapitole provedu porovnání HTTP a SOCKS proxy tak, aby bylo jasné jejich určení, přednosti a naopak nevýhody.

3.2.3.2.1 HTTP versus SOCKS

Při výběru proxy serverů se můžeme setkat primárně s dvěma typy: HTTP Proxy a SOCKS Proxy. Rozdíl je zejména ve vrstvě, na které jednotlivé proxy operují a z toho vyplývajících konsekvencí. Zatímco HTTP Proxy pracuje na sedmé, nejvyšší vrstvě ISO/OSI a čte jednotlivé

HTTP požadavky, které upraví a předá cílovému serveru, tak SOCKS funguje na páté vrstvě a s obsahem určeným pro cílový server nepracuje, pouze zpracovává příkazy instruující kam ho má předat. HTTP Proxy je primárně určena pro prohlížení webových stránek, ačkoliv pomocí příkazu CONNECT dokáže zpracovat i další TCP spojení (nikoliv UDP) (57). Oproti tomu SOCKS není vázaný na žádný protokol a od verze 5 podporuje přenos UDP paketů, IPv6 i autentizaci (58).

Při využívání SOCKS Proxy je nutno mít na vědomí, že ačkoliv obsah webových stránek se nám načítá přes proxy, ještě se musíme ujistit, že i DNS dotazy jsou směrovány přes SOCKS server. Jelikož DNS dotazy jsou vysílány pomocí UDP protokolu je potřeba využít protokol SOCKS v5. Alternativou je rozšíření protokolu v4a, které přidává k SOCKS v4 podporu pro překlad doménových jmen na straně serveru (59).

3.2.3.3 Specializované lokální proxy servery

Specializované proxy servery zahrnují rozmanité druhy proxy serverů, které jsou programovány jako odpověď na specifickou potřebu klientů. Pro naše účely nás budou zajímat zejména proxy servery pro obcházení cenzury a anonymizační servery, které jsou s nimi často spjaty.

Oproti dříve popsaným variantám proxy serverů mají specializované nástroje klientskou část v podobě programu, který je potřeba spustit na uživatelském počítači, a ta se pak spojí s cílovým proxy serverem pomocí definovaných protokolů. Klientská část vytváří lokální HTTP/SOCKS Proxy, kterou je pak možné využít pomocí běžných programů, obdobně jako v případě vzdálené HTTP/SOCKS Proxy. Klientská aplikace často vyřeší i problém s hledáním adres zprostředkujících serverů, jelikož často obsahuje mechanismy k získání těchto adres.

Jednotlivé programy se liší v ceně, otevřenosti kódu, možnostech zabezpečení, množství zprostředkujících serverů a nabízené anonymitě. Nejčastěji se setkáme se softwarem, který má uzavřený kód a vyznává přístup *bezpečnost skrze utajení* (angl. security through obscurity), čili znepřístupněním vlastního kódu se snaží zabránit odhalení možných zranitelností, a tak znesnadnit práci cenzora. Zastánci tohoto přístupu jsou programy Freegate (60), Ultrasurf (61) či Hotspot Shield (62). U těchto programů není znám přesný princip komunikace a způsob nakládání s nasbíranými informacemi o klientech.

Protagonisti opačného přístupu, tzv. *bezpečnost díky návrhu* (angl. secure by design), věří, že kód by měl být bezpečný sám o sobě a jeho zveřejněním by nemělo vzniknout riziko prolomení, naopak využívá jeho přístupnosti k nalezení chyb a jejich rychlé opravě (viz Linusův zákon (63)). Příkladem otevřeného kódu je například Tor (64), Psiphon (65) či Lantern (66) (viz dále).

3.2.3.3.1 Tor

V případě Toru probíhá přístup na Internet alespoň přes tři proxy servery a data jsou přeposílána pomocí metody *onion routing*, která jednotlivé zprávy zabaluje do vrstev, z nichž každou zvládne rozšifrovat pouze jeden ze zúčastněných proxy serverů. Tento přístup zajišťuje, že přenášená data dokáže přečíst pouze poslední server tzv. *exit relay*, ale ten ovšem neví, kdo je původcem této komunikace. Databáze jednotlivých proxy serverů se nachází na předem určených serverech tzv. *directory authorities* a v cache jednotlivých proxy serverů, odtud si je klientské programy stahují pomocí protokolu HTTP (67; 68).

3.2.3.3.2 Psiphon

Psiphon je aplikace určená pro systémy Windows a mobilní zařízení se systémem Android, která za pomocí VPN, SSH tunelu, meek transportního protokolu (viz 3.2.5.2) a lokální HTTP/SOCKS Proxy zjišťuje šifrované spojení s proxy servery Psiphonu. Na rozdíl od Toru je primárním účelem Psiphonu obcházení cenzury, nikoliv zajištění anonymity. Základní seznam proxy serverů je stahován spolu s klientem a tento seznam je postupně obnovován tak, aby reagoval na postupné blokování serverů cenzorem (69; 70).

3.2.3.3.3 Lantern

Lantern klient má seznam blokovaných stránek a využívá svoji síť k jejich zpřístupnění. Tento princip umožní rychlé prohlížení běžných stránek a ke směrování přes Lantern síť dojde, pouze když je potřeba obejít cenzuru. Lantern síť je tvořena jak proxy servery Lanternu, tak jednotlivými klienty, kteří umožní využívat část jejich Internetového pásma k přístupu na Internet pro uživatele z cenzurovaných zemí. Informace o jednotlivých proxy serverech jsou distribuovány za využití algoritmu *Kaleidoscope*, který má zajistit, aby žádný klient nebyl schopen získat adresy všech proxy serverů (66; 71).

3.2.4 Anonymní peer-to-peer sítě

Peer-to-peer sítě jsou postaveny na principu komunikace rovný s rovným, čili jednotliví klienti jsou mezi sebou propojeni a navzájem si poskytují data, bez potřeby centrálního řídícího prvku. Tento decentralizovaný princip je využíván i v případě obcházení cenzury a to na základě vytváření uzavřených sítí, které jsou přístupné pouze uživatelům do těchto sítí zapojených, nejčastěji pomocí specializovaného programu, ve kterých si jednotliví uživatelé sdílejí data či poskytují jiné služby.

Komunikace mezi uživateli bývá šifrovaná a zúčastněné strany navzájem nejsou schopny identifikovat svoji lokaci či IP adresu a mají přehled pouze o klientech přímo sousedících. Tento přístup zajišťuje ochranu jednotlivých klientů i v případě, že se do sítě infiltruje nevyžádaný klient. Mezi projekty založené na tomto principu patří například open source programy Freenet (72) a I2P (73). Freenet si vydobyl velkou popularitu zejména v ČLR, kde byl označen za nejpoužívanější nástroj k obcházení cenzury (43).

V následujících podkapitole se na Freenet blíže zaměřím. Dále představím Dark Web, který bývá součástí anonymních peer-to-peer sítí.

3.2.4.1.1 *Freenet*

Freenet slouží jako *decentralizované datové úložiště* a klienti u sebe uchovávají zašifrovaná data, které sdílí s ostatními klienty, čímž se zajišťuje jejich dostupnost i v případě že se zdrojový klient ze sítě odpojí. Data se distribuují mezi uživatele do jejich cache na základě jejich popularity, což zvyšuje výkonnost a rychlost sítě. Klienti mají přehled pouze o přímo připojených sousedech, přes které proudí jejich data. Freenet nabízí možnost si tyto sousedy vybrat a zajistit tak, aby tok dat proudil pouze přes důvěryhodné uživatele (74).

3.2.4.1.2 Dark Web

Dark Web je termín označující webové stránky, které nejsou přístupné pomocí běžných Internetových vyhledávačů a jsou záměrně skryté a nepřístupné z konvenčních Internetových prohlížečů (75; 76). Bývá často spojován s anonymními peer-to-peer sítěmi, ale ty nejsou jeho podmínkou.

Pro přístup k Dark Web Internetovým stránkám je potřeba využít klientskou aplikaci (např. I2P, Freenet, Tor), která nás propojí se sítí (označována jako Darknet (77)), na které se nachází

skryté webové servery. Jednotlivé sítě mají své vlastní formáty doménových jmen a způsoby jejich distribuce. Často využívají vlastní TLD domény (např. .onion, .i2p) a není výjimkou, že názvy, za účelem zvýšení bezpečnosti, nekladou nároky na snadnou zapamatovatelnost. Příkladem doménového jména na síti Tor je *duskgytldkxiuqc6.onion*, které v sobě obsahuje část hashe veřejného klíče umožňující ověřit vzdálený server (78).

3.2.5 Transportní protokoly

Transportními protokoly jsou v kontextu této práce chápány jako protokoly sloužící k přenosu dat způsobem, který přináší do komunikace přidanou hodnotu, ale zároveň nijak neovlivňuje přenášený obsah. Přínosem může být například zvýšení bezpečnosti či odolnosti proti cenzuře. Protokoly nám pomohou vytvořit mezi námi a cílovým uzlem tunel, přes který pak může proudit běžná komunikace.

V následujících odstavcích charakterizuji transportní protokoly pro obcházení cenzury na základě způsobu, kterým se snaží cenzuru obejít. Tyto základní typy pak budu dále kategorizovat a uvedu jejich zástupce.

3.2.5.1 Odolnost proti filtrování IP adres

Protokoly, které mají za úkol co možná nejvíce znesnadnit blokování cílových IP adres cenzorem, jsou postaveny na principu využívání velkého množství cílových IP adres. Tento systém spoléhá na to, že cenzor by při hromadné filtraci blokoval značné množství uživatelů a serverů, kteří nejsou do obcházení cenzury nikterak zapojeni, a proto se radši tomuto blokování vyvaruje – tzv. minimalizace civilních dopadů (11).

3.2.5.1.1 Browser-Based Proxy

První ze způsobů, jak do komunikace zahrnout stále nové neblokované IP adresy, je označovaný jako *Browser-Based Proxy*. Ten spoléhá na zapojení běžných uživatelů Internetu, kteří se nacházejí v oblasti, která není pod kontrolou cenzora. Tito uživatelé dočasně slouží jako zprostředkující proxy servery. Touto cestou se může z prohlížeče pouhým navštívením webové stránky obsahující příslušné skripty (např. JavaScript skripty či Adobe Flash aplikace) stát mezi-prvek pro obcházení cenzury. Po dobu setrvání na zapojených webových stránkách se IP adresa uživatele zapíše do databáze, ze které je předávána uživatelům snažící se obejít cenzuru

a ti se pak přepojují mezi jednotlivými návštěvníky. Na tomto principu je postaven protokol Flashproxy (79).

3.2.5.1.2 Decoy routing

Další z technik, jak bojovat s blokováním IP adres, je *Decoy routing*, nazývané též *End-to-Middle proxying*, který využívá skutečnost, že pakety při své cestě Internetem prochází přes velké množství směrovačů, a každý z nich může potencionálně být prostředkem k obcházení cenzury. Směrovač zapojený do Decoy routingu (tzv. *decoy router*) při zjištění komunikace pomocí protokolu určeného k obcházení cenzury tuto komunikaci odkloní na předem určenou adresu, kde se nachází proxy server zprostředkovávající neblokovanou komunikaci (tzv. *decoy proxy*), a dále se chová, jako by paket byl doručen na falešnou IP adresu nacházející se v hlavičce paketu. Tímto způsobem je možné využít jakoukoliv z IP adres nacházejících se za daným směrovačem jako cílovou za účelem zmatení cenzora, aniž by majitelé těchto IP adres byli do obcházení cenzury jakkoliv zapojeni (80). Projekty postavené na tomto principu jsou například Telex (81), Cirripede (82) či TapDance (83).

3.2.5.1.3 Domain Fronting

Další metodou je využití IP adres a doménových jmen, které by blokováním zamezily přístup i na nezávadné služby, je tzv. *Domain Fronting.* Ten těží ze skutečnosti, že velcí poskytovatelé cloudových služeb jako Content Delivery Network (CDN) (např. Akamai (84), Amazon CloudFront (85), CloudFlare (86)) či webových aplikačních platforem (např. Google App Engine (87)) často užívají pro své služby, za účelem zvýšení spolehlivosti, dostupnosti a bezpečnosti, velké množství IP adres, doménových jmen a šifrované spojení. Tyto služby jsou dostupné veřejnosti a jejich předplacením či volným užíváním je lze využít pro spojení s cílovými proxy servery, které obchází cenzuru, a tím tak skrýt jejich adresu a doménové jméno. V hlavičce paketu je viditelné pouze doménové jméno (TLS SNI (88)) serveru cloudové služby (front domain), které je využito i pro DNS dotazy, a až uvnitř zabezpečeného paketu je doménové jméno (HTTP Host (89)) cílového proxy serveru (např. Tor bridge), ze kterého jsou získávány odpovědi (90; 91). Domain Fronting je základem projektů GoAgent (92) či Meek (90; 91).

Na principu využití veřejně dostupných webových služeb a steganografie jsou postavené i další projekty, liší se především formou předávání zpráv, rychlostí odezvy, popřípadě odolností proti odhalení. Využity mohou být sdílená úložiště dokumentů a fotografií, blogy, sociální sítě či jiné

služby (např. servery provádějící skenování stránek (93)). Příkladem těchto projektů jsou CloudTransport (94), Collage (95), MIAB (96) a OSS (93).

3.2.5.2 Odolnost proti detekování protokolu

S rozšířením DPI technik (viz 2.3.2), probabilistických metod (viz 2.3.3) a rozličných aktivních metod (viz 2.3.4) odhalování protokolů už přestává stačit pouhé šifrování kanálu a musel být zvýšen tlak na zamaskování komunikace a odstranění odlišností, které umožňují komunikaci obcházející cenzuru odhalit. Pro skrývání komunikace lze zvolit dva hlavní směry: ten první se snaží zamaskovat jakékoliv charakteristiky komunikace, které by ji umožňovaly přiřadit ke konkrétnímu protokolu či programu, ten druhý naopak komunikaci co nejvíce připodobňuje k již existujícímu a cenzorem tolerovanému protokolu. Souhrnně se tyto techniky nazývají *zastírací protokoly* (angl. obfuscating protocols) (97). Kromě těchto dvou směrů se lze setkat s doplňkovými technikami, které se snaží reagovat na nově zjištěné způsoby detekce či využívají slabých stránek systémů detekujících a blokujících nežádoucí komunikaci (11).

3.2.5.2.1 Imitování neexistujícího protokolu

Pokoušíme-li se skrýt charakteristické rysy přenášeného obsahu je potřeba se zaměřit na veškeré anomálie, které by mohly být využity k detekci a následném blokování komunikace cenzorem. Techniky se běžně zaměřují na využívané protokoly, porty, velikost paketů, obsah hlaviček a datové části, míru entropie, časový sled komunikace a dobu mezi jednotlivými odpověďmi (97). Jakékoliv přehlédnutí, jako například užití specifického prvočísla pro šifrování či unikátní seznam používaných šifer při navazování spojení, může vést k detekci a následné blokaci (98; 99).

Techniky se snaží zakódovat veškerý obsah tak, aby vypadal náhodně a nezůstaly v něm žádné statické části. Mezi známé projekty patří například Dust (100), obfuscated-ssh (101), obfs2 (102), obfs3 (103), obfs4 (104) či ScrambleSuit (105).

3.2.5.2.2 Imitování existujícího protokolu

Při napodobování existujících protokolů je třeba najít protokoly, které nebývají cenzory blokované a poté přeměnit stávající tok dat tak, aby co nejvíce připomínal povolený protokol. Tento přístup se spoléhá na to, že *hloubková analýza paketů* bude hlásit velký počet falešně

pozitivních (angl. false positive) nálezů a blokování by způsobovalo omezení běžné pro cenzora nezávadné komunikace.

Technika napodobování existujících protokolů je náchylná na jednotlivá specifika jako reakce protokolů na chyby v komunikaci, entropie nebo charakter komunikace (spolehlivost/nespolehlivost). Zástupci projektů založené na steganografii skrz existující protokoly jsou: FTE (106), který upraví formát komunikace tak, aby odpovídal zadanému *regulárnímu výrazu*, StegoTorus (107) imitující peer-to-peer protokoly či nešifrované HTTP, Code Talker Tunnel (108) (dříve SkypeMorph (109)) napodobující Skype protokol či FreeWave (110), který komunikaci moduluje do akustického signálu.

3.2.5.2.3 Využívání slabin detekce a blokování

Odlišným přístupem oproti skrývání komunikace je analýza detekovacích přístrojů cenzora a reakce na jeho slabiny. Firewally a DPI boxy postupují podle předem daných pravidel, které slouží k odhalování a blokování nežádoucí komunikace, pokud se podaří tyto pravidla a postupy vypozorovat, je možné se blokování vyhnout.

Příkladem této techniky může být ignorování *active probing* ze strany cenzora či zahazování TCP RST paketů, zasílaných cenzorem za účelem blokování komunikace (21). Dalším příkladem je využití znalosti o tom, jak DPI boxy postupují při sledování TCP spojení. Ty totiž ukládají informace o vytváření, shromažďování a zahazování dat pro jednotlivá spojení v tzv. *Transmission Control Blocks* (TCB). Běžně dochází ke smazání TCB, jakmile je obdržen příkaz k ukončení komunikace. Tuto skutečnost ovšem můžeme nasimulovat (např. zasláním TCP RST/FIN s krátkým TTL), a tak se vyhnout dalšímu sledování konkrétního spojení (111). Příkladem projektu zabývajících se vyhýbání *active probingu* je BridgeSPA (112), ignorování RST paketů je popsáno v projektu West Chamber (113).

3.2.5.3 Kombinace technik

Kromě protokolů zaměřujících se na jeden z aspektů blokování cenzurou vznikají i komplexní protokoly, které kombinují více přístupů a snaží se vytvořit jeden obtížně překonatelný systém. Příkladem může být již zmíněný FreeWave, který kromě přenosu dat pomocí akustického signálu zvládá také skrývat IP adresu FreeWave proxy serveru. Akustický signál lze totiž předávat pomocí reálných hovorů (skrze Skype, Google Voice a další) a tímto způsobem skrýt

cílovou IP adresu za jeden z VoIP uzlů, které do obcházení cenzury nejsou vědomě zapojeny a slouží pouze k předávání dat dál po síti (110).

Některé protokoly při cestě za obtížnějším odhalením ztrácí na své rychlosti. Například CensorSpoofer využitím asymetrické komunikace je schopný oddělit příchozí a odchozí kanál a zatímco požadavky zasílá pomocí emailů/IM, odpovědi chodí skrze UDP pakety, které napodobující VoIP/Video Chat komunikaci. Tato technika umožňuje skrýt IP adresu proxy serveru, jelikož UDP pakety nevyžadují navazování spojení a proxy server může paketům přiřadit jakoukoliv zdrojovou IP adresu, nevýhodou je pomalejší čas na zpracování a vyřízení požadavku (při testování trvalo načtení stránky wikipedia.org v průměru 27 sekund) (114).

Transportních protokolů existuje velké množství, ovšem k jejich reálnému využití pro obcházení cenzury je zapotřebí dostatečné aplikační a serverové zázemí. Jedním z nejvýznamnějších popularizátorů a také vývojářů transportních protokolů je The Tor Project, který v aktuální verzi programu Tor Browser (verze 5.5.5) implementuje protokoly: FTE, FTE-ipv6, Meek, obfs3, obfs4 a ScrambleSuit (115; 116). Dalším zástupcem je projekt Psiphon, který využívá protokol Meek (70).

3.3 Distribuce programů

Závažnou překážkou při obcházení cenzury bývá samotné získání programů k tomu sloužících. Domovské stránky těchto projektů bývají mezi prvními blokovanými, a proto je často potřeba využít alternativní metody distribuce.

Kromě programů samotných je k využitích některých metod obcházení cenzury (např. *transportních protokolů*) nutné znát dodatečné informace pro připojení jako jsou IP adresy proxy serverů, jejich veřejný klíč či další ověřující data pro připojení (např. fingerprint). Tyto informace nebývají volně přístupné na Internetu, z důvodu ochrany před blokováním poskytovaných adres, a proto jsou potřeba alternativní způsoby distribuce.

3.3.1 Přímý kontakt

Nejbezpečnější cestou distribuce je využití sítě známých, konkrétně kontaktovat důvěryhodnou osobu, která nám předá program či přihlašovací údaje osobně či bezpečným kanálem. Tento způsob se jeví jako nejhůře napadnutelný ze strany cenzora (98), ovšem je třeba počítat s tím, že tyto sociální vazby mohou být zpětně rozkrývány v případě vyšetřování a specifické sdílené

informace (např. verze programu, přihlašovací údaje, zprostředkující servery) se mohou stát vodítkem spojující jednotlivé osoby.

3.3.2 Zrcadlící servery

Zrcadlení (angl. mirroring) webových stránek se provádí vytvořením kopie existujících stránek na jiném serveru s jinou adresou. Tímto lze docílit, že obsah je dostupný z více míst, což je výhodné z hlediska rychlosti stahování ať už díky rozložení zátěže mezi více serverů nebo z důvodu stahování z geograficky bližšího serveru. V případě obcházení cenzury může každá nová adresa umožnit přístup uživatelům z cenzurovaných sítí, dokud nebude zanesena do seznamu blokovaných adres.

Poslední dobou se jako úložiště pro instalační soubory programů pro obcházení cenzury dočkávají stále větší obliby cloudové služby (např. Dropbox, Google Docs, Github, Amazon Web Services) (117). Tyto služby často nebývají blokovány cenzorem z důvodu dopadu i na běžný nezávadný obsah, který tyto služby poskytují.

3.3.3 Email/IM

K distribuci lze využít také mailové zprávy či komunikaci pomocí IM (např. XMPP). Princip je založen na zaslání zprávy na sběrný účet, ze kterého následně přijde odpověď v podobě odkazů ke stažení programů. Ty většinou směřují na *zrcadlící servery* (viz 3.3.2). Z důvodu velikosti programů a omezení mailových poskytovatelů se program nezasílá jako příloha. Některé projekty umožňují pomocí textu či adresy specifikovat jazyk, verzi programu či platformu, pro kterou má být program určen (117; 118).

Stejný postup lze využít i pro získání informací k připojení na skryté proxy servery. Zde se můžeme setkat i s dalšími restrikcemi, například podmínkou, že email musí být odeslán z domény předem specifikovaných mailových poskytovatelů (např. Google, Yahoo). Touto cestou se snaží správci adres zamezit strojovému odhalení velkého počtu adres proxy serverů cenzorem, a delegují tak ochranu proti zneužití na mailové poskytovatele a jejich obranné prvky proti hromadnému vytváření mailových schránek (36; 119).

3.3.4 Sociální sítě

Internetové sociální sítě (např. Twitter) lze využít pro zasílání linků na úložiště programů obdobným způsobem jako emaily a IM zprávy. Sociální sítě nabízejí možnost posílání soukromých zpráv mezi jednotlivými uživateli a pomocí robotů, kteří obsluhují specifikovaný účet, tak odpovídat na požadavky o zaslání odkazů (117).

3.3.5 Ověření pravosti

Jakýkoliv program stažený z Internetu může být teoreticky podvržený cenzorem a upraven tak, aby zasílal třetí osobě důvěrné informace či jiným způsobem měnil chod programu (viz 2.4.2). Proto je doporučené si před instalací jakéhokoliv programu zkontrolovat integritu a autenticitu instalačních souborů.

3.3.5.1 Integrita

Kontrola integrity instalačního souboru, čili skutečnosti, že soubor byl kompletně stažen a je neporušen, se provádí pomocí kontroly *otisku* (angl. fingerprint). Otisk je krátký textový řetězec o fixní délce, který je výstupem hashovací funkce. Ta přijímá za vstup instalační soubor a pomocí definovaných algoritmů (např. MD5, SHA-1, SHA-256) má za úkol vytvořit otisk, který je pro každý vstup unikátní (120).

Porovnáme-li otisk uvedený na webových stránkách, odkud jsme soubor stahovali, s tím, který jsme si vygenerovali po stažení, zjistíme, zdali byl soubor stažen v pořádku. Kontrola integrity nám ovšem nic neříká o autorovi souboru, a pokud by se povedlo třetí osobě převzít kontrolu nad webovými stránky projektu, nic by jí nebránilo v tom si vygenerovat otisky k modifikovaným instalačním souborům a ty uveřejnit. My bychom poté nic nepoznali ani při kontrole otisku.

3.3.5.2 Autenticita

Pro kontrolu autenticity, čili ověření pravosti původu instalačního souboru, slouží *digitální podpisy*. Digitální podpis je obdobou toho ručně psaného, nabízí metodu ověření, zda určitá osoba daný dokument (soubor) schválila. V digitálním světě se podepisování provádí na základe asymetrické kryptografie, kdy autor programu vlastní soukromý klíč, pomocí kterého podepisuje vydávané programy, a veřejný klíč, sloužící k ověření podpisu, umístí na volně

přístupné místo (webové stránky, databáze veřejných klíčů) či předá bezpečnou cestou ostatním (120).

Ověření pak probíhá tím způsobem, že spolu s instalačním souborem si stáhneme i jeho digitální podpis, pomocí kterého pak zkontrolujeme *integritu* a *autenticitu* instalačního souboru. Tedy jestli osoba, která vygenerovala podpis, je tatáž jako ta, která vlastní soukromý klíč, od kterého my vlastníme jeho veřejnou část (autenticita), a zdali jsou instalační soubory v pořádku (integrita) (121; 122). Účinnost této metody závisí na bezpečnosti kanálu, kterým jsme získali veřejný klíč autora. Tento klíč se většinou nemění a lze tak použít pro všechny nové verze programu.

4 Aplikování technik obcházení cenzury

V poslední kapitole budu aplikovat konkrétní techniky obcházení cenzury v praxi na cenzurou postižené zemi. Cílem měření je prokázat přítomnost cenzury ve vybrané zemi a následně ověřit dostupnost blokovaných webových stránek za použití jednotlivých metod obcházení cenzury a tím zjistit využitelnost obcházecích technik.

Nejprve provedu výběr země vhodné pro testování a popíši přesnou metodiku sběru a konečného vyhodnocení dat. Následně zvolím techniky, kterými budu zkoumat přítomnost cenzury a provádět obcházení cenzury v cílové zemi. Závěrem této kapitoly budu prezentovat výsledky získané z měření ve vybrané zemi.

4.1 Výběr země

Jednotlivé země mají při aplikování cenzury Internetu svá specifika, proto je nutné určit kritéria pro výběr. Při výběru země budu zvažovat především následující tři parametry: míru cenzury, počet obyvatel postižených cenzurou a bezpečnost pohybu po cílové zemi.

4.1.1 Míra cenzury

Při testování se chci zaměřit na zemi, ve které je cenzura Internetu všudypřítomná a techniky obcházení cenzury tedy budou mít největší přínos v podobě zpřístupnění značného množství informací. Vysoká míra cenzury navíc usnadní testování, jelikož blokování a následné obcházení bude průkaznější.

Pro určení míry cenzury vycházím z nejnovějších dat organizace Freedom House (4), kde se na žebříčku pro rok 2015 objevilo 19 zemí s nejhorším hodnocení z hlediska Internetové svobody (Not Free), ze kterých budu vybírat (viz Tab. A1).

4.1.2 Počet postižených obyvatel

Počet obyvatel cenzurou postižené země umocňuje závažnost cenzury. Při porovnávání jednotlivých zemí budu zkoumat dostupná data o sčítání obyvatel. Pro účely výběru postačí pouze rámcová data, čili není nutné dohledávat nejaktuálnější počty.

4.1.3 Bezpečnost pohybu po cílové zemi

Jelikož měření bude probíhat přímo v cílové zemi, je nutné také zohlednit míru rizika při setrvání v dané zemi. Pro hodnocení bezpečnosti budu vycházet primárně z aktuálních doporučení a varování Ministerstva zahraničních věcí ČR pro pohyb v jednotlivých zemích (123).

4.1.4 Výsledek výběru země

Na základě daných kritérií byla jako země pro testování vybrána ČLR. Ta obdržela nejvíce bodů na žebříčku Freedom of the Net 2015, zároveň se jedná o nejlidnatější zemi světa s největším počtem Internetových uživatelů (124) a aktuálně není na seznamu zemí nebezpečných pro vycestování (125).

4.2 Metodika sběru dat

Za účelem sběru dat odletím do ČLR a samotný sběr dat budu provádět osobně na předem připravené počítačové stanici (viz 4.2.3). Pro měření se budu připojovat k veřejně dostupným Wi-Fi sítím, abych co možná nejvíce snížil možnost identifikace mé osoby kvůli případným persekucím ze strany ČLR.

ČLR je charakteristická svým *Velkým čínským firewallem* (angl. the Great Firewall of China). Na základě dostupných materiálů o čínském firewallu (126) lze předpokládat, že míra cenzury se může v ČLR lišit v závislosti na lokalitě a konkrétním poskytovateli Internetového připojení. Proto bude měření probíhat ve více geograficky odlehlých místech, čímž chci docílit vyšší vypovídající hodnoty výsledků měření.

Pro měření byly vybrány následující destinace: zvláštní správní oblast Hong Kong, provincie Če-ťiang a Šen-si a dvě přímo spravovaná města Peking a Šanghaj. Při výběru lokalit byl kladen důraz na pokrytí co největšího území ČLR, dále na lidnatost měst kvůli dopadu cenzury a v neposlední řadě na dopravní dostupnost.

V každé lokalitě bude provedeno právě jedno měření. V případě nápadně vysoké neúspěšnosti při testování, které by mohlo být způsobenou špatnou kvalitou připojení k Internetu, bude pokus opakován po připojení k jinému poskytovateli Internetu.

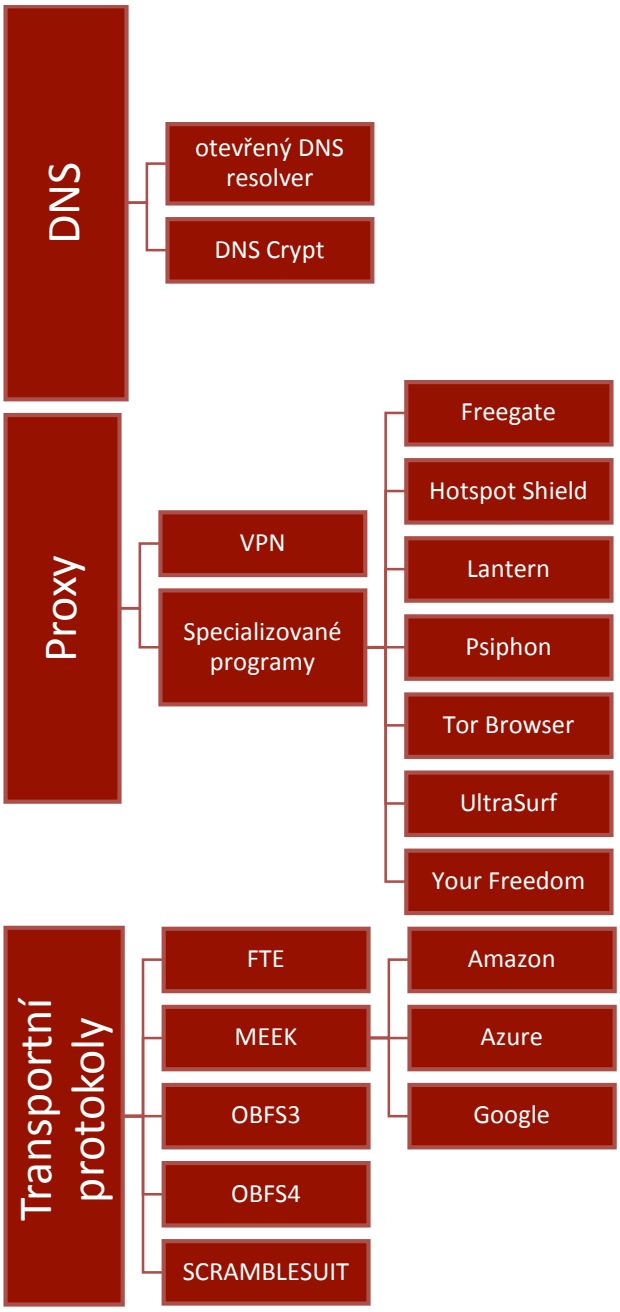
Samotné měření bude probíhat ve dvou hlavních fázích. První bude mít za účel zjistit aktuální stav cenzury a druhá pak za pomoci jednotlivých technik se pokusí cenzurovaný obsah zobrazit. Tyto fáze spolu s nástroji pro měření jsou blíže popsány v následujících podkapitolách.

4.2.1 Zjištění stavu cenzury

Prvním krokem je odhalení, do jaké míry a jestli vůbec k cenzuře dochází. Pro měření bude testován přístup na vybrané webové stránky (viz Tab. B1), které by měly být v ČLR na základě webového portálu Herdict (viz 1.4.2) blokovány. Kromě přístupu na webové stránky bude za pomoci lokální DNS databáze zkoumáno, zdali cenzura spoléhá na DNS manipulaci, nebo jestli jsou stránky filtrovány i při znalosti IP adresy cílového webového serveru. Všechny testy budou probíhat na základě výchozí konfigurace získané od poskytovatele Internetu v dané lokalitě.

4.2.2 Testování technik obcházení cenzury

Jakmile bude potvrzeno, že v dané lokalitě dochází k cenzuře Internetu, další fází bude aplikování jednotlivých metod obcházení cenzury s cílem zobrazit obsah vybraných webových stránek. Testované metody budou rozděleny do tří skupin: DNS, Proxy a transportní protokoly. Schéma testů je možné vidět na následujícím diagramu (viz Obr. 4.1).



Obr. 4.1: Testované metody obcházení cenzury

Zdroj: autor

4.2.2.1 DNS

Pokud by na základě počátečního testu stavu cenzury (viz 4.2.1) byla zjištěna DNS manipulace ze strany cenzora, je šance, že k zpřístupnění cenzurovaných stránek postačí zajistit si důvěryhodný zdroj DNS záznamů. Za tímto účelem se budu nejprve snažit dotazovat na IP adresu *veřejně přístupného DNS resolveru* (konkrétně OpenDNS) a tím ověřit, jestli podvržené

odpovědi neposkytují pouze servery aktuálního poskytovatele Internetu. Další fází bude využití *DNSCrypt* k získání IP adres skrze zabezpečený kanál.

Po získání DNS odpovědí z ČLR bude probíhat následné ověření správnosti získaných IP adres v České republice, kde jsou všechny testované webové stránky (viz Tab. B1) dostupné. Získané IP adresy (v případě více adres v DNS odpovědi bude vybrána právě jedna z nich) zadám do HTML dotazů na cílové webové stránky a obdržím-li jako odpověď žádané webové stránky, bude ověřena správnost adresy.

4.2.2.2 Proxy

Následující fáze testu bude založena na využití zprostředkujících proxy serverů (viz 3.2.3) mimo území ČLR, které zakryjí cílovou adresu webového serveru a případně zajistí šifrování kanálu směrem k proxy serveru. Testovány budou jak klasické *VPN*, tak i specializované programy určené k obcházení cenzury. Test bude považován za úspěšný v případě, kdy se mi podaří připojit ke vzdálenému proxy serveru a zobrazit alespoň jednu z blokovaných webových stránek.

VPN bude testováno oproti předem připravenému serveru v České republice. Využité protokoly a jejich verze je možné dohledat v Tab. B3. Specializované programy pro obcházení cenzury využívají servery a protokoly vlastní. Pro testování budou využívány neplacené verze programu a výchozí konfigurace.

Kvůli velkému počtu programů zaměřených na obcházení cenzury je nutno jejich výběr zúžit. Pro testování budou využity programy, které byly zařazeny organizací FreedomHouse do přehledu využívaného software k obcházení cenzury (43). Vyřazeny budou dále programy, které se již dále nevyvíjí, tj. během posledního roku nevyšla žádná nová verze programu. Po tomto výběru zůstalo k testování sedm programů, jejichž seznam včetně použitých verzí prezentuje Tab. B2.

V testu se neobjeví *jednoduché webové proxy* (viz 3.2.3.1), jelikož jejich úspěšnost závisí primárně na tom, zdali je doména poskytovatele blokována. Výsledek testu by tak závisel pouze na mém výběru stránek, nikoliv na využité technologii, která se při šifrované (HTTPS) komunikaci s proxy serverem shoduje s přístupem na běžné webové stránky a je i stejným způsobem blokována.

4.2.2.3 Transportní protokoly

Poslední sada testů bude zaměřena na transportní protokoly a jejich schopnosti se vyhnout blokování ze strany cenzora. Testovány budou transportní protokoly obsažené v programu Tor Browser (verze 5.5.5, viz Tab. B4). Program Tor Browser byl zvolen, jelikož jeho autor The Tor Project se aktivně stará o implementaci existujících a vývoj nových transportních protokolů a v současnosti nabízí největší výběr transportních protokolů připravených k použití (zdroj: autor). Zároveň je možné instalovat program Tor na vlastním serveru a ten využít jako protistranu komunikace a tím zamezit zkreslení testů, ke kterému by docházelo při komunikaci s již blokovaným serverem.

Pro testování budou využívány jako protistrana komunikace adresy serverů již integrované do programu Tor Browser, dále adresy získané ze stránek The Tor Project (119) a také vlastní server umístěný v České republice, který má naimplementovanou podporu následujících transportních protokolů: fte, obfs3, obfs4, scramblesuit. Cílem testu je zjistit zdali jsou jednotlivé transportní protokoly blokovány cenzurou, čili za využití jednotlivých adres se pokusím o propojení s Tor bridge.

V případě, že se mi podaří připojit k zprostředkujícímu serveru s pomocí konkrétního transportního protokolu a zobrazím alespoň jednu z blokovaných stránek, test bude považován za úspěšný.

4.2.3 Nástroje pro měření

Vzhledem k cílům měření budou využity nástroje pro zobrazení obsahu webových stránek a monitorování DNS dotazů. Další dílčí měření bude zahrnovat trasování cesty paketů, zjištění přidělené veřejné IP adresy a sledování celkové síťové komunikace pro případ upřesnění výsledků a dalšího zkoumání. Nástroje využívané pro měření přímo závisí na platformě, kde bude probíhat testování. Verze jednotlivých programů je možné nalézt v Tab. B5.

4.2.3.1 Zobrazení webových stránek

Pro stažení obsahu webových stránek bude na platformě Linux využit program *wget* s parametry zajišťující 2 pokusy o spojení s dobou čekání 90 sekund. Na operačním systému Windows bude testování probíhat pomocí *Internet Exploreru* s dobou čekání alespoň 2 minut

na kompletní vykreslení stránek. Výstupem měření jsou HTML soubory webových stránek bez obrázků a dalších podpůrných souborů (např. JS, CSS).

4.2.3.2 Monitoring DNS dotazů

Sledování DNS dotazů bude probíhat pouze na platformě Linux pomocí programu *DiG*. Program má za úkol zobrazit IP adresy (tzv. A záznamy) příslušných doménových jmen. Dotazovaný DNS server bude nakonfigurován v závislosti na právě probíhajícím testu (viz 4.2.2.1).

4.2.3.3 Zjištění dostupnosti portů

Dostupnost portů zjistím programem nmap, který bude skenovat otevřené TCP porty na připraveném serveru v ČR. Pro test byly vybrány následující porty: 22, 25, 80, 143, 443, 465, 587 a 993. Ty jsou využívány pro služby jako je mailová komunikace, FTP, SSH či zobrazování webových stránek.

Port 587 je na cílovém serveru filtrován a slouží jako referenční, pro ověření, zdali program nmap vyhodnocuje správně uzavřené porty. Ve výsledku testů by se tedy správně při jeho skenování měla objevit hodnota *filtered.*

4.2.3.4 Trasování cesty paketů

Pro trasování cesty paketů bude využit programe *traceroute*. Trasováno bude pomocí ICMP paketů a TCP SYN paketů. Počet skoků je omezen na 15 a doba čekání je 3 sekundy.

4.2.3.5 Sledování komunikace

Celá komunikace bude sledována pomocí programu *Wireshark* na platformě Windows a *tcpdump* na systému Linux. Výstupy budou ukládány ve formátu *pcap*.

4.2.3.6 Zjištění veřejné IP adresy

Zjištění veřejné IP adresy využívané pro komunikaci s ostatními servery na Internetu bude zajištěno přes Internetovou stránku *iplocation.net*, která bude zařazena do seznamu zobrazovaných webových stránek (viz Tab. B1). Webová stránka je schopna zobrazit IP adresu, ze které se na ni přistupuje, a dále poskytuje základní geografické informace o původu této adresy.

4.2.3.7 Tunelování komunikace přes proxy

Na platformě Linux bude využit program *proxychains* pro zajištění komunikace ostatních nástrojů přes proxy server. Nastavení proxy serveru bude závislé na konkrétně testované metodě obcházení cenzury.

4.3 Praktický test a výsledky

Praktický test probíhal v období od 12. 5. do 31. 5. 2016. Geograficky bylo měřeno na území Čínské lidové republiky ve městech: Hong Kong, Chang-čou (Če-ťiang), Si-an (Šen-si) Peking a Šanghaj. Následně bylo prováděno dodatečné měření v provincii Kan-su, které mělo ověřit a objasnit případné nesrovnalosti zjištěné předchozím měřením.

Ve všech lokalitách kromě Chang-čou bylo provedeno právě jedno měření. V Chang-čou byl zopakován test prováděný linuxovými nástroji (viz Tab. B5), jelikož došlo v průběhu měření ke zhoršení kvality připojení.

Výsledky jsou prezentovány pomocí tabulek, kde znak odškrtnutí (✓) označuje úspěšnost a křížek (✖) naopak neúspěšnost při konkrétním měření.

Nejprve představím výsledky zjištěného stavu cenzury v jednotlivých lokalitách, poté budu prezentovat efektivnost metod pro obcházení cenzury a nakonec výsledky měření shrnu v samostatné podkapitole.

4.3.1 Stav cenzury

Prvotní měření je zaměřeno na zjištění stavu cenzury. V tabulce níže (viz Tab. 4.1) je možné vidět dostupnost webových stránek v jednotlivých lokacích. Z výsledků je patrné, že v Hong Kongu, který je znám svoji volnější politikou, k blokaci vybraných webových stránek nedochází. Ostatní měřené lokality vykazovaly známky cenzury, tzn. na vybraném vzorku byly dostupné pouze stránky iplocation.net, zařazené do testovaných stránek jako součást testovacích nástrojů (viz 4.2.3.6), a vyhledávač Yahoo.

Tab. 4.1: Dostupnost webových stránek

	Hong Kong	Chang-čou	Si-an	Peking	Šanghaj
cn.nytimes.com	✓	✖	✖	✖	✖
dropbox.com	✓	✖	✖	✖	✖
iplocation.net	✓	✓	✓	✓	✓
nytimes.com	✓	✖	✖	✖	✖
search.yahoo.com	✓	✓	✓	✓	✓
twitter.com	✓	✖	✖	✖	✖

Zdroj: autor (Příloha D:/web/wget-DEFAULT)*

Druhé testování bylo zaměřeno na blokaci portů mezi počítačem v ČLR a serverem v České republice. Výsledek prezentuje Tab. 4.2. Je patrné, že kromě portu 587, který byl na straně serveru uzavřen a byl použit jako referenční pro ověření správného výsledku, jsou všechny ostatní porty dostupné.

Výjimku tvoří TCP port 25 při testování v Hong Kongu, v této lokalitě bylo prováděno pouze jedno měření a ze subjektivního hlediska Internet nebyl znatelně filtrován, proto bych se spíše přikláněl ke skutečnosti, že k blokování docházelo pouze na straně vybraného poskytovatel Internetu, nikoliv na úrovní celé správní oblasti.

Tab. 4.2: Dostupnost portů cílového serveru

	Hong Kong	Chang-čou	Si-an	Peking	Šanghaj
22/tcp	✓	✓	✓	✓	✓
25/tcp	✖	✓	✓	✓	✓
80/tcp	✓	✓	✓	✓	✓
143/tcp	✓	✓	✓	✓	✓
443/tcp	✓	✓	✓	✓	✓
465/tcp	✓	✓	✓	✓	✓
587/tcp	✖	✖	✖	✖	✖
993/tcp	✓	✓	✓	✓	✓

Zdroj: autor

Poslední tabulka (viz Tab. 4.3) prezentuje úspěšnost zobrazení webových stránek při znalosti IP adres cílových serverů, čili eliminujeme-li vliv DNS. Dle výsledků je patrné, že cenzura nespoléhá pouze na techniky DNS manipulace, ale ovlivňuje i samotnou komunikaci s nežádoucími webovými servery. V Hong Kongu jako v jediné lokalitě nedochází k blokování komunikace s vybranými webovými stránkami. Na příkladu stránek *dropbox.com* je vidět, že existují webové stránky, na kterých je vykonávána pouze DNS manipulace, jinak jsou přístupné.

Při bližším zkoumání záznamů komunikace je možné odhalit, že dochází k blokování komunikace. Využíváno je zejména zahazování paketů a resetování spojení pomocí paketů TCP RST.

Tab. 4.3: Výsledky měření za pomocí lokální DNS databáze

	Hong Kong	Chang-čou	Si-an	Peking	Šanghaj
cn.nytimes.com	✓	✖	✖	✖	✖
dropbox.com	✓	✓	✓	✓	✓
nytimes.com	✓	✖	✖	✖	✖
twitter.com	✓	✖	✖	✖	✖

Zdroj: autor

4.3.2 Obcházení cenzury

Na základě výsledků získaných při zjišťování stavu cenzury (viz 4.3.1), byl zúžen výběr webových stránek, jejichž dostupnost se dále testovala pomocí nástrojů pro obcházení cenzury. Následně došlo k testům DNS odpovědí, využitelnosti proxy serverů a transportních protokolů.

4.3.2.1 DNS

Tab. 4.4 prezentuje správnost DNS odpovědí získaných postupně od výchozích DNS serverů Internetového poskytovatele (získáno z DHCP), otevřených DNS serverů (OpenDNS) a záznamů získaných protokolem DNS Crypt. Podrobné tabulky zobrazující získané IP adresy pro jednotlivé webové stránky a jejich úspěšnost jsou v Příloze B (viz Tab. C1, Tab. C2 a Tab. C3).

Výsledky nám ukazují, že kromě Hong Kongu ve všech testovaných lokalitách dochází k manipulaci DNS záznamů nezávisle na použitých DNS serverech. Naopak šifrovaná komunikace s DNS serverem (protokol DNS Crypt) se jeví jako účinná metoda pro získání věrohodných IP adres. V testu bylo za pomocí DNS Crypt dosaženo 100% úspěšnosti ve všech měřených lokalitách (viz Tab. C3).

Tab. 4.4: Výsledky úspěšnosti DNS metod v ČLR

	Hong Kong	Chang-čou	Si-an	Peking	Šanghaj
Výchozí DNS servery	✓	✖	✖	✖	✖
Otevřené DNS servery	✓	✖	✖	✖	✖
DNS Crypt	✓	✓	–	✓	✓

Pozn.: V lokalitě Si-an nebyl proveden test pomocí DNSCrypt, proto není možné prezentovat výsledky.
Zdroj: autor (Tab. B1-3)

4.3.2.2 Proxy servery

Test proxy serverů měl ukázat úspěšnost využití zprostředkujících serverů k obcházení cenzury. Tab. 4.5 prezentuje výsledky jak klasického VPN serveru (OpenVPN) využitého jako výchozí brána, tak i sofistikovanějších variant proxy serverů zaměřených na anonymizaci připojení na Internet či obcházení cenzury.

Z výsledků je patrné, že drtivá většina programů byla v obcházení cenzury úspěšná. Problémy nastaly pouze u programu Tor Browser, který nebyl schopen projít fází iniciace (angl. bootstrapping) a nenačetl si tedy potřebné data pro navázání spojení s ostatními servery. Řešení tohoto problému nabízí transportní protokoly (viz 4.3.2.3).

Obdobné obtíže se vyskytly i u programu Your Freedom, který ve dvou lokalitách (Si-an a Šanghaj) nebyl schopen zkontaktovat přednastavený zprostředkující server a metoda vyhledávání nových serverů nebyla úspěšná.

Tab. 4.5: Výsledky úspěšnosti proxy serverů

	Hong Kong	Chang-čou	Si-an	Peking	Šanghaj
VPN server	✓	✓	✓	✓	✓
Freegate	✓	✓	✓	✓	✓
Hotspot Shield	✓	✓	✓	✓	✓
Lantern	✓	✓	✓	✓	✓
Psiphon	✓	✓	✓	✓	✓
Tor Browser	✖	✖	✖	✖	✖
UltraSurf	✓	✓	✓	✓	✓
Your Freedom	✓	✓	✖	✓	✖

Zdroj: autor (Příloha D:/iexplore/;*/wireskark/;*/tcpdump/)*

4.3.2.3 Transportní protokoly

Posledním testovaným prvkem byly transportní protokoly konkrétně jejich odolnost proti odhalení a blokování cenzurou (viz Tab. 4.6). Stoprocentní úspěšnosti dosáhly pouze protokoly meek-amazon, meek-azure a obfs4. Navazování spojení si také vyžádalo nejméně úsilí ze strany uživatele, jelikož připojení probíhalo rychle a bez nutnosti zkoušet více adres zprostředkujících serverů.

Neúspěch ve všech lokalitách byl zaznamenán u transportního protokolu meek-google, příčina ovšem není způsobena cenzurou, ale dočasnou nedostupností tohoto protokolu po čas provádění testů (127).

Tab. 4.6: Výsledky úspěšnosti transportních protokolů

	Hong Kong	Chang-čou	Si-an	Peking	Šanghaj
FTE	✓	✖	✓	✓	✓
MEEK-AMAZON	✓	✓	✓	✓	✓
MEEK-AZURE	✓	✓	✓	✓	✓
~~MEEK-GOOGLE~~	✖	✖	✖	✖	✖
OBFS3	✓	✓	✓	✓	✖
OBFS4	✓	✓	✓	✓	✓
SCRAMBLESUIT	✓	✓	✓	✓	✖

Zdroj: autor (Příloha D:/prtscr/;*/tcpdump/)*

4.3.3 Shrnutí

Měření, které probíhalo v pěti různých lokalitách na území ČLR, bylo zaměřeno na zjištění stavu cenzury v ČLR a úspěšnost jednotlivých metod při jejím obcházení. Testování na vybraném vzorku Internetových stránek potvrdilo přítomnost cenzury ve všech měřených lokalitách kromě Hong Kongu.

Konkrétně bylo prokázáno aplikování metod *DNS manipulace* jak při dotazování výchozích serverů, tak i vzdálených otevřených DNS resolverů (viz 4.3.2.1) a dále byly identifikovány metody *blokování komunikace* při testech zobrazování stránek, kde byl vliv DNS serverů eliminován (viz 4.2.1).

Jako účinný způsob obrany proti DNS manipulaci se ukázal nástroj DNS Crypt (viz 4.3.2.1), který provádí šifrování komunikace mezi uživatelem a DNS serverem. Pro komunikaci s webovým serverem a tedy úspěšné zobrazení cenzurovaných webových stránek je ovšem nutné využít také proxy servery nacházející se v cenzurou nepostižené lokalitě.

Většina testovaných programů zaměřených na obcházení cenzury, včetně klasického VPN serveru, byla při testu úspěšná. Problémy nastaly pouze u dvou programů a to u aplikace Tor Browser, kterou bez transportních protokolů nebylo možné využít v žádné z testovaných lokalit, a Your Freedom, který měl problémy ve městech Si-an a Šanghaj (viz 4.3.2.2).

Test transportních protokolů prokázal bezproblémový chod protokolů meek-amazon, meek-azure a obfs4. Zbytek protokolů projevil problémy alespoň v jedné z testovaných lokalit (viz 4.3.2.3).

Závěr

Prvním cílem diplomové práce bylo zmapovat mechanismy určené k provádění státní Internetové cenzury. Za účelem splnění tohoto cíle jsem nejprve popsal místa, kde je možné cenzuru provádět. Jelikož primárním zaměřením této práce je cenzura aplikovaná na komunikační kanál, stručně jsem představil Internetovou komunikaci a vyzdvihl aspekty využívané cenzurou. Poté už jsem popisoval přímo mechanismy cenzury, které jsem rozdělil na fázi detekce a následné filtrace nežádoucí komunikace. Každá z fází byla blíže charakterizovaná a dále rozčleněna na jednotlivé metody.

Druhým cílem bylo popsat techniky, které slouží k obcházení státní Internetové cenzury. Pro splnění toho cíle jsem definoval klíčové principy obcházení cenzury: skrytí cíle, obsahu, odlišností, udržení co nejvyšší míry autonomie, popřípadě využití slabin cenzury. Dále jsem se zaměřil na jednotlivé techniky obcházení cenzury Internetu, které jsem popsal a uvedl jejich reprezentanty v podobě programů či protokolů.

Třetím a posledním cílem bylo ověřit funkčnost technik pro obcházení v praxi v cenzurou postižené zemi. Pro výběr vhodné země byly zvoleny tři kritéria: míra cenzury, počet postižených obyvatel a bezpečnost pohybu po cílové zemi. Jako cílová země byla nakonec vybrána Čínská lidová republika. Nejprve bylo potřeba zjistit aktuální stav cenzury a následně byl proveden v jednotlivých lokalitách test obcházecích technik. Cenzura byla detekována ve všech měřených lokalitách kromě Hong Kongu. Následný test nástrojů pro obcházení cenzury vykázal úspěšnost především u technik založených na bázi proxy serverů a transportních protokolů.

Cenzura Internetu je závažné téma a dle vývoje v posledních letech lze očekávat její nárůst i v budoucnu, kdy bude prorůstat do více zemí a jejich legislativ. Metody jejího aplikování jsou účinné a dokáží výrazně znesnadnit přístup obyvatel k informacím. Nástroje pro obcházení cenzury s cenzurou stále drží krok, ale s ubývajícím množstvím zemí se svobodným přístupem k informacím, bude zároveň ubývat i množství míst, odkud budou moci zprostředkující servery čerpat data. Jedinou cestou, jak dlouhodobě zabránit rozšiřování cenzury, je zvýšit tlak na politické subjekty určující pravidla cenzury a apelovat na právo svobodného přístupu k informacím.

Pro další zkoumání by bylo užitečné se zaměřit na motivaci a morální zásady korporací poskytujících nástroje cenzury zemím omezujícím lidská práva, dále zmapovat příčiny politického tlaku na zpřísňování cenzury Internetu. Z technického hlediska by si zasloužilo pozornost téma transportních protokolů zejména prozkoumání dalších možností na poli maskování komunikace a praktického využití.

Zdroje

1. Otto, Jan a Pražák, Josef Miroslav. *Ottův Slovník naučný.* Praha : Otto, Jan, 1892. stránky 290-292. Sv. V.

2. Newth, Mette. The Long History of Censorship. *Beacon for Freedom of Expression.* [Online] 2010. http://www.beaconforfreedom.org/liste.html?tid=415&art_id=475.

3. censorship. *Merriam-Webster.* [Online] http://www.merriam-webster.com/dictionary/censorship.

4. Freedom House. Freedom on the Net 2015. *Freedom House.* [Online] 10 2015. [Citace: 01. 06 2016.] https://freedomhouse.org/sites/default/files/FH_FOTN_2015Report.pdf.

5. OpenNet Initiative. ONI Methodology, Tools, and Data FAQ. *OpenNet Initiative.* [Online] https://opennet.net/oni-faq.

6. —. Country Profiles. *OpenNet Initiative.* [Online] https://opennet.net/country-profiles.

7. Freedom House. Methodology - Freedom on the Net 2015. *Freedom House.* [Online] https://freedomhouse.org/report/freedom-net-2015/methodology-freedom-net-2015.

8. Reporters Without Borders. WHO WE ARE ? *Reporters Without Borders.* [Online] 10. 03 2014. https://en.rsf.org/who-we-are-12-09-2012,32617.html.

9. —. Collateral Freedom. *Ennemies of the Internet.* [Online] https://12mars.rsf.org/2015-en/.

10. Tor Project, Inc. OONI. [Online] https://ooni.torproject.org/.

11. Khattak, Sheharbano, Simon, Laurent a Murdoch, J. Steven. Systemization of Pluggable Transports for Censorship Resistance. [Online] 1, 23. 12 2014. [Citace: 20. 04 2016.] http://arxiv.org/abs/1412.7448.

12. OpenNet Initiative. China's Green Dam: The Implications of Government Control Encroaching on the Home PC. [Online] https://opennet.net/chinas-green-dam-the-implications-government-control-encroaching-home-pc.

13. Bamman, David, O'Conner, Brendan a Smith, Noah A. Censorship and deletion practices in Chinese social media. [Online] 3 2012. http://journals.uic.edu/ojs/index.php/fm/article/view/3943/3169doi:10.5210/fm.v17i3.3943.

14. OpenNet Initiative. Internet Filtering in China in 2004-2005: A Country Study. [Online] 14. 4 2005. https://opennet.net/sites/opennet.net/files/ONI_China_Country_Study.pdf.

15. Bright, Martin. BT puts block on child porn sites. *TheGuardian.* [Online] 06. 06 2004. [Citace: 18. 06 2016.] https://www.theguardian.com/technology/2004/jun/06/childrensservices.childprotection.

16. The Canadian Centre for Child Protection. Cleanfeed Canada. *Cybertip.ca.* [Online] [Citace: 18. 06 2016.] https://www.cybertip.ca/app/en/projects-cleanfeed.

17. Socolofsky, T. a Kale, C. RFC1180: A TCP/IP Tutorial. *IETF.* [Online] 01 1991. [Citace: 18. 06 2016.] https://tools.ietf.org/rfc/rfc1180.txt.

18. The Institute of Electrical and Electronics Engineers, Inc. IEEE Standard for Ethernet. *IEEE Std 802.3-2012.* [Online] 28. 12 2012. [Citace: 18. 06 2016.] http://standards.ieee.org/getieee802/download/802.3-2012_section1.pdf.

19. Braden, R. RFC1122: Requirements for Internet Hosts -- Communication Layers. *IETF.* [Online] 10 1989. [Citace: 18. 06 2016.] https://tools.ietf.org/rfc/rfc1122.txt.

20. Deering, S. a Hinden, R. RFC2460: IPv6 Specification. *IETF.* [Online] 12 1998. [Citace: 18. 06 2016.] https://www.ietf.org/rfc/rfc2460.txt.

21. *Ignoring the Great Firewall of China.* Clayton, Richard, Murdoch, Steven J. a Watson, Robert N. M. 3, 2007, A Journal of Law and Policy for the Information Society, Sv. 2, stránky 70-77.

22. Touch, Joe, Lear, Eliot a Mankin, Allison. Service Name and Transport Protocol Port Number Registry. *IANA.* [Online] 13. 06 2016. [Citace: 18. 06 2016.] https://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.txt.

23. Postel, J. RFC768: User Datagram Protocol. *IETF.* [Online] 28. 08 1980. [Citace: 18. 06 2016.] https://www.ietf.org/rfc/rfc768.txt.

24. Information Sciences Institute (University of Southern California). RFC793: Transmission Control Protocol. *IETF.* [Online] 09 1981. [Citace: 18. 06 2016.] https://www.ietf.org/rfc/rfc793.txt.

25. WhatIsMyIpAddress. Geolocation Providers. *WhatIsMyIpAddress.* [Online] [Citace: 2016. 3 6.] http://whatismyipaddress.com/geolocation-providers.

26. Bittau, Andrea, a další. The case for ubiquitous transport-level encryption. [Online] 2008. [Citace: 07. 03 2016.] https://www.usenix.org/legacy/events/sec10/tech/full_papers/Bittau.pdf.

27. Porter, Thomas. The Perils of Deep Packet Inspection. *Symantec.* [Online] Symantec, 10. 01 2005. [Citace: 01. 03 2016.] http://www.symantec.com/connect/articles/perils-deep-packet-inspection.

28. Nowakowski, Tomasz, a další. *Safety and Reliability: Methodology and Applications.* London : CRC Press, 2014. 9781138026810.

29. Wiley, Brandon. Blocking-Resistant Protocol Classification Using Bayesian Model Selection. [Online] 2011. http://blanu.net/BayesianClassification.pdf.

30. Winter, Philipp a Crandall, Jedidiah R. The Great Firewall of China: How it Blocks Tor and Why it is Hard to Pinpoint. *USENIX ;login:*. 2012, Sv. 37, 6, stránky 42-45.

31. Anderson, Collin. Dimming the Internet: Detecting Throttling as a Mechanism of Censorship in Iran. 2013.

32. Lowe, Graham, Winters, Patrick a Marcus, Michael L. The Great DNS Wall of China. [Online] 21. 12 2007. [Citace: 2016. 03 16.] http://cs.nyu.edu/~pcw216/work/nds/final.pdf.

33. RIPE NCC. YouTube Hijacking: A RIPE NCC RIS case study. *RIPE NCC.* [Online] 17. 03 2008. [Citace: 16. 03 2016.] https://www.ripe.net/publications/news/industry-developments/youtube-hijacking-a-ripe-ncc-ris-case-study.

34. Cowie, Jim. Egypt Leaves the Internet. *DynResearch.* [Online] 27. 01 2011. [Citace: 16. 03 2016.] http://research.dyn.com/2011/01/egypt-leaves-the-internet/.

35. —. The Battle for Tripoli's Internet. *DynReseach.* [Online] 21. 08 2011. [Citace: 16. 03 2016.] http://research.dyn.com/2011/08/the-battle-for-tripolis-intern/.

36. Dingledine, Roger. Tor and censorship: lessons learned [26C3]. *YouTube.* [Online] 27. 12 2009. [Citace: 24. 04 2016.] https://youtu.be/gQ5R4NzVlbE.

37. Jiroušek, Radim. *Principy digitální komunikace.* Voznice : Leda, 2006. ISBN 80-7335-084-X.

38. Dierks, T. a Rescorla, E. RFC5246: The Transport Layer Security (TLS) Protocol Version 1.2. *IETF.* [Online] 08 2008. [Citace: 20. 03 2016.] https://tools.ietf.org/html/rfc5246.

39. Ylonen, T. a Lonvick, E. RFC4253: The Secure Shell (SSH) Transport Layer Protocol. *IETF.* [Online] 01 2006. [Citace: 20. 03 2016.] https://tools.ietf.org/html/rfc4253.

40. Rescorla, E. a Modadugu, N. RFC4347: Datagram Transport Layer Security. *IETF.* [Online] 04 2006. [Citace: 20. 03 2016.] https://tools.ietf.org/html/rfc4347.

41. Kent, S. a Seo, K. RFC4301: Security Architecture for the Internet Protocol. [Online] 12 2005. [Citace: 20. 03 2016.] https://tools.ietf.org/html/rfc4301.

42. OTR Development Team. Off-the-Record Messaging Protocol version 3. *CypherPunks.* [Online] OTR Development Team. [Citace: 20. 03 2016.] https://otr.cypherpunks.ca/Protocol-v3-4.1.1.html.

43. Callanan, Cormac, a další. *Leaping Over the Firewall: A Review of Censorship Circumvention Tools.* Washington D.C. : Freedom House, 2010. str. 76.

44. Cisco Systems, Inc. How Virtual Private Networks Work. *CISCO.* [Online] 13. 10 2008. [Citace: 14. 04 2016.] https://www.cisco.com/c/en/us/support/docs/security-vpn/ipsec-negotiation-ike-protocols/14106-how-vpn-works.pdf. 14106.

45. Callas, J., a další. RFC4880: OpenPGP Message Format. *IETF.* [Online] 11 2007. [Citace: 20. 03 2016.] https://tools.ietf.org/html/rfc4880.

46. Ramsdell, B. a Turner, S. RFC5751: Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.2. *IETF.* [Online] 01 2010. [Citace: 20. 03 2016.] https://tools.ietf.org/html/rfc5751.

47. Zmijewski, Earl. Turkish Internet Censorship Takes a New Turn. *DynResearch.* [Online] 30. 03 2014. [Citace: 21. 03 2016.] http://research.dyn.com/2014/03/turkish-internet-censorship/.

48. Arends, R., a další. RFC4033: DNS Security Introduction and Requirements. *IETF.* [Online] IETF, 03 2005. [Citace: 22. 03 2016.] https://tools.ietf.org/html/rfc4033.

49. IANA. Trust Anchors and Keys. *IANA.* [Online] [Citace: 22. 03 2016.] https://www.iana.org/dnssec/files.

50. Lamb, Rick. DNSSEC Deployment Report. [Online] 18. 03 2016. [Citace: 22. 03 2016.] http://rick.eng.br/dnssecstat/.

51. CZ.NIC. Statistiky - Domény podle DNSSEC. *CZ.NIC.* [Online] CZ.NIC, 21. 03 2016. [Citace: 22. 03 2016.] https://stats.nic.cz/stats/domains_by_dnssec/?rd=2016-03-21&dr=1y&tp=i-1d&ss=0&ds=normal&da=table.

52. Bernstein, J. Daniel. *DNSCurve.* [Online] [Citace: 02. 04 2016.] https://dnscurve.org/.

53. Wander, Matthäus. An overview of secure name resolution [29c3]. *Youtube.* [Online] 29. 12 2012. [Citace: 02. 04 2016.] 29th Chaos Communication Congress [29c3]. http://www.youtube.com/watch?v=eOGezLjlzFU&feature=youtu.be&t=44m55s.

54. Denis, Frank a Gacogne, Remi. DNSCrypt version 2 protocol specification. [Online] 29. 01 2016. [Citace: 01. 04 2016.] https://github.com/jedisct1/dnscrypt-proxy/blob/master/DNSCRYPT-V2-PROTOCOL.txt.

55. DNSCrypt. *DNSCrypt.* [Online] [Citace: 01. 04 2016.] https://dnscrypt.org/.

56. Dynamic Internet Technology. Dynaweb. *Dynamic Internet Technology.* [Online] [Citace: 12. 04 2016.] http://www.dit-inc.us/dynaweb.html.

57. Fielding, E. a Reschke, J. RFC 7231: Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content. *IETF.* [Online] 06 2014. [Citace: 05. 04 2016.] https://www.ietf.org/rfc/rfc7230.txt.

58. Leech, M., a další. RFC 1928: SOCKS Protocol Version 5. *IETF.* [Online] 03 1996. [Citace: 05. 04 2016.] https://www.ietf.org/rfc/rfc1928.txt.

59. Lee, Ying-Da. SOCKS 4A: A Simple Extension to SOCKS 4 Protocol. *OpenSSH.* [Online] [Citace: 11. 04 2016.] http://www.openssh.com/txt/socks4a.protocol.

60. Dynamic Internet Technology. Freegate. *Dynamic Internet Technology.* [Online] [Citace: 12. 04 2016.] http://www.dit-inc.us/freegate.html.

61. Ultrareach Internet Corporation. *Ultrasurf.* [Online] [Citace: 12. 04 2016.] http://ultrasurf.us/.

62. AnchorFree, Inc. . *Hotspot Shield.* [Online] [Citace: 12. 04 2016.] https://www.hotspotshield.com/.

63. Reymond, Eric S. *The cathedral & the bazaar: musings on Linux and open source by an accidental revolutionary.* Cambridge : O'Reilly, 1999. str. 30. 1-56592-724-9.

64. The Tor Project. *Tor.* [Online] [Citace: 12. 04 2016.] https://www.torproject.org/.

65. Psiphon Inc. *Psiphon.* [Online] [Citace: 12. 04 2016.] https://psiphon.ca/.

66. Brave New Software Project, Inc. *Lantern.* [Online] [Citace: 23. 04 2016.] https://getlantern.org/.

67. The Tor Project, Inc. Tor: Overview. *Tor.* [Online] [Citace: 13. 04 2016.] https://www.torproject.org/about/overview.html.en.

68. —. Tor directory protocol, version 3. *Tor.* [Online] 13. 04 2016. [Citace: 13. 04 2016.] https://gitweb.torproject.org/torspec.git/tree/dir-spec.txt.

69. Psiphon Inc. Psiphon 3 Circumvention System README. [Online] 13. 04 2016. [Citace: 13. 04 2016.] https://bitbucket.org/psiphon/psiphon-circumvention-system.

70. —. Open Source. *Psiphon* . [Online] [Citace: 13. 04 2016.] https://www.psiphon3.com/en/open-source.html.

71. Team Lantern. Lantern: Questions and Answers. *GitHub.* [Online] 18. 09 2015. [Citace: 23. 04 2016.] https://github.com/getlantern/lantern/wiki/%5Bdevelopers%5D-Questions-and-Answers.

72. The Freenet Project Inc. *Freenet.* [Online] [Citace: 13. 04 2016.] https://freenetproject.org.

73. I2P Project. *The Invisible Internet Project.* [Online] [Citace: 13. 04 2016.] https://geti2p.net.

74. The Freenet Project Inc. Documentation. *Freenet.* [Online] [Citace: 13. 04 2016.] https://freenetproject.org/documentation.html.

75. Dictionary.com . dark web. *Dictionary.com.* [Online] Random House, Inc. [Citace: 14. 04 2016.] http://www.dictionary.com/browse/dark-web.

76. BrightPlanet. Clearing Up Confusion – Deep Web vs. Dark Web. *BrightPlanet.* [Online] 27. 03 2014. [Citace: 14. 04 2016.] http://www.brightplanet.com/2014/03/clearing-confusion-deep-web-vs-dark-web/.

77. Rouse, Margaret. darknet. *TechTarget.* [Online] 12 2011. [Citace: 14. 04 2016.] http://searchnetworking.techtarget.com/definition/darknet.

78. The Tor Project, Inc. Tor Rendezvous Specification. *Tor.* [Online] 06. 04 2016. [Citace: 14. 04 2016.] https://gitweb.torproject.org/torspec.git/tree/rend-spec.txt.

79. Fifield, David, a další. Evading Censorship with Browser-Based Proxies. *Privacy Enhancing Technologies.* Lecture Notes on Computer Science, Sv. 7384.

80. *Decoy Routing: Toward Unblockable Internet Communication.* Karlin, Josh, a další. San Francisco : autor neznámý, 2011. FOCI'11 - Proceedings of the USENIX Workshop on Free and Open Communications on the Internet.

81. *Telex: Anticensorship in the Network Infrastructure.* Wustrow, Eric, a další. San Francisco : autor neznámý, 2011. USENIX Security Symposium.

82. *Cirripede: Circumvention Infrastructure using Router Redirection with Plausible Deniability.* Houmansadr, Amir, a další. 2011. CCS '11 Proceedings of the 18th ACM conference on Computer and communications security. stránky 187-200. ISBN: 978-1-4503-0948-6 doi>10.1145/2046707.2046730.

83. *TapDance: End-to-Middle Anticensorship without Flow Blocking.* Wustrow, Eric, Swanson, Colleen M. a Halderman, J. Alex. 2014. 23rd USENIX Security Symposium.

84. Akamai Technologies. *Akamai.* [Online] [Citace: 18. 04 2016.] https://www.akamai.com/.

85. Amazon Web Services, Inc. Amazon CloudFront – Content Delivery Network (CDN). *Amazon Web Services.* [Online] [Citace: 18. 04 2016.] https://aws.amazon.com/cloudfront/.

86. CloudFlare, Inc. *CloudFlare.* [Online] [Citace: 18. 04 2016.] https://www.cloudflare.com/.

87. Google Inc. Google App Engine. *Google Cloud Platform.* [Online] [Citace: 18. 04 2016.] https://cloud.google.com/appengine/.

88. Eastlake, D. RFC6066: Transport Layer Security (TLS) Extensions: Extension Definitions. *IETF.* [Online] 01 2011. [Citace: 18. 04 2016.] https://tools.ietf.org/rfc/rfc6066.txt. ISSN: 2070-1721.

89. Fielding, R. a Reschke, J. RFC7230: Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing. *IETF.* [Online] 06 2014. [Citace: 18. 04 2016.] https://tools.ietf.org/rfc/rfc7230.txt. ISSN: 2070-1721.

90. Fifield, David, a další. Blocking-resistant communication through domain fronting. *Proceedings on Privacy Enhancing Technologies.* 06 2015, Sv. 2015, 2, stránky 46-64.

91. The Tor Project, Inc. Meek. *Tor.* [Online] 07. 04 2016. [Citace: 18. 04 2016.] https://trac.torproject.org/projects/tor/wiki/doc/meek.

92. The Tor Project. GoAgent. *Tor.* [Online] 21. 09 2015. [Citace: 23. 04 2016.] https://trac.torproject.org/projects/tor/wiki/doc/GoAgent.

93. Fifield, David, Nakibly, Gabi a Boneh, Dan. OSS: Using Online Scanning Services for Censorship Circumvention. *Privacy Enhancing Technologies.* 2013, Sv. 7981, stránky 185-204.

94. Brubaker, Chad, Houmansadr, Amir a Shmatikov, Vitaly. CloudTransport: Using Cloud Storage for Censorship-Resistant Networking. *Privacy Enhancing Technologies.* 2014, Sv. 8555, stránky 1-20.

95. Burnett, Sam, Feamster, Nick a Vempala, Santosh. Chipping Away at Censorship Firewalls with User-Generated Content. *USENIX Security Symposium.* 2010.

96. Invernizzi, Luca, Kruegel, Christopher a Vigna, Giovanni. Message In A Bottle: Sailing Past Censorship. *Annual Computer Security Applications Conference.* 2013.

97. *Seeing through Network-Protocol Obfuscation.* Wang, Liang, a další. Denver : ACM, 2015. CCS '15 Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. stránky 57-69. 978-1-4503-3832-5.

98. Appelbaum, Jacob a Dingledine, Roger. How governments have tried to block Tor [28C3]. *YouTube.* [Online] 28. 12 2011. [Citace: 20. 04 2016.] https://youtu.be/GwMr8Xl7JMQ.

99. Wilde, Tim. Knock Knock Knockin' on Bridges' Doors. *Tor.* [Online] 07. 01 2012. [Citace: 20. 04 2016.] https://blog.torproject.org/blog/knock-knock-knockin-bridges-doors.

100. Wiley, Brandon. *Dust: A Blocking-Resistant Internet Transport Protocol.* School of Information, University of Texas. Austin : autor neznámý, 2011.

101. Leidl, Bruce. obfuscated-openssh/PROTOCOL. *GitHub.* [Online] 07. 04 2010. [Citace: 21. 04 2016.] https://github.com/brl/obfuscated-openssh.

102. The Tor Project. obfs2 (The Twobfuscator). *Tor.* [Online] 23. 01 2015. [Citace: 21. 04 2016.] https://gitweb.torproject.org/pluggable-transports/obfsproxy.git/tree/doc/obfs2/obfs2-protocol-spec.txt.

103. —. obfs3. *Tor.* [Online] 23. 01 2015. [Citace: 21. 04 2016.] https://gitweb.torproject.org/pluggable-transports/obfsproxy.git/tree/doc/obfs3/obfs3-protocol-spec.txt.

104. —. obfs4. *Tor.* [Online] 13. 04 2016. [Citace: 21. 04 2016.] https://gitweb.torproject.org/pluggable-transports/obfs4.git/tree/doc/obfs4-spec.txt.

105. *ScrambleSuit: A Polymorphic Network Protocol to Circumvent Censorship.* Winter, Philipp, Pulls, Tobias a Fuss, Juergen. New York : autor neznámý, 2013. WPES '13 Proceedings of the 12th ACM workshop on Workshop on privacy in the electronic society. stránky 213-224. ISBN: 978-1-4503-2485-4 doi>10.1145/2517840.2517856.

106. *Protocol Misidentification Made Easy with Format-Transforming Encryption.* Dyer, Kevin P., a další. Berlin : autor neznámý, 2013. CCS’13.

107. *StegoTorus: A Camouflage Proxy for the Tor Anonymity System.* Weinberg, Zachary, a další. Raleigh : autor neznámý, 2012. CCS’12.

108. Mohajeri, Hooman a Li, Baiyu. Code Talker Tunnel. *GitHub.* [Online] 20. 08 2014. [Citace: 21. 04 2016.] https://github.com/blizzardplus/Code-Talker-Tunnel.

109. *SkypeMorph: Protocol Obfuscation for Tor Bridges.* Moghaddam, Hooman Mohajeri, a další. Raleigh : autor neznámý, 2012. CCS’12.

110. *I want my voice to be heard: IP over Voice-over-IP for unobservable censorship circumvention.* Houmansadr, Amir, a další. 2013. Network and Distributed System Security Symposium - NDSS’13.

111. *Towards Illuminating a Censorship Monitor’s Model to Facilitate Evasion.* Khattak, Sheharbano, a další. Washington, D.C. : USENIX, 2013. 3rd USENIX Workshop on Free and Open Communications on the Internet.

112. *BridgeSPA: Improving Tor Bridges with Single Packet Authorization.* Smits, Rob, a další. Chicago : autor neznámý, 2011. WPES’11.

113. West Chamber: Season 3. *GitHub.* [Online] 06. 11 2015. [Citace: 21. 04 2016.] https://github.com/liruqi/west-chamber-season-3/.

114. *CensorSpoofer: Asymmetric Communication using IP Spoofing for Censorship-Resistant Web Browsing.* Wang, Qiyan, a další. Raleigh : autor neznámý, 2012. CCS’12.

115. The Tor Project. Tor Browser Bundle. *Tor.* [Online] 23. 04 2016. [Citace: 23. 04 2016.] https://gitweb.torproject.org/builders/tor-browser-bundle.git/.

116. —. A Childs Garden of Pluggable Transports. *Tor.* [Online] 06. 09 2015. [Citace: 23. 04 2016.] https://trac.torproject.org/projects/tor/wiki/doc/AChildsGardenOfPluggableTransports.

117. —. GetTor. *Tor.* [Online] 2016. [Citace: 23. 04 2016.] https://gettor.torproject.org/.

118. Psiphon Inc. Download Psiphon. *Psiphon.* [Online] [Citace: 24. 04 2016.] https://psiphon.ca/en/download.html.

119. The Tor Project. BridgeDB. *Tor.* [Online] 2016. [Citace: 24. 04 2016.] https://bridges.torproject.org/.

120. Paar, Christof a Pelzl, Jan. *Understanding cryptography : a textbook for students and practitioners.* Berlin : Springer, c2010. ISBN 978-3-642-04100-6.

121. The Tor Project. How to verify signatures for packages. *Tor.* [Online] [Citace: 24. 04 2016.] https://www.torproject.org/docs/verifying-signatures.html.en.

122. Making and verifying signatures. *The GNU Privacy Handbook.* [Online] [Citace: 24. 04 2016.] https://www.gnupg.org/gph/en/manual/x135.html.

123. Ministerstvo zahraničních věcí ČR. státy světa - informace na cesty. *Ministerstvo zahraničních věcí ČR.* [Online] [Citace: 12. 05 2016.] http://www.mzv.cz/jnp/cz/encyklopedie_statu.

124. Central Intelligence Agency. The World Factbook 2016-17. *Central Intelligence Agency.* [Online] 2016. [Citace: 22. 06 2016.] https://www.cia.gov/library/publications/resources/the-world-factbook/rankorder/2153rank.html.

125. Ministerstvo zahraničních věcí ČR. Čína: Specifika, bezpečnostní situace, doporučení turistům. *Ministerstvo zahraničních věcí ČR.* [Online] [Citace: 03. 05 2016.] http://www.mzv.cz/jnp/cz/encyklopedie_statu/asie/cina/cestovani/other.html.

126. Xu, Xueyang , Mao, Z. Morley a Halderman, J. Alex. Internet Censorship in China: Where Does the Filtering Occur? *Proceedings of the 12th International Conference on Passive and Active Measurement.* 2011, stránky 133-142.

127. Fifield, David. [tor-talk] meek-google suspended for terms of service violations. *The Tor Project.* [Online] 02. 06 2016. [Citace: 07. 06 2016.] https://lists.torproject.org/pipermail/tor-talk/2016-June/041057.html.

128. Internet Info, s.r.o. Slovníček pojmů: blacklist, whitelist, greylist. *Root.cz.* [Online] [Citace: 21. 06 2016.] https://www.root.cz/slovnicek/blacklist-whitelist-greylist/.

129. Walton, Andy. What Are Domain Name System (DNS) Resolvers and How Do They Work? *Chron.* [Online] [Citace: 23. 06 2016.] http://smallbusiness.chron.com/domain-name-system-dns-resolvers-work-76639.html.

130. Bernat, Michal. *Fenomén počítačového pirátství.* Právnická fakulta, Masarykova univerzita v Brně. Brno : autor neznámý, 2008. Diplomová práce.

131. Kaspersky Lab. Získejte více informací o malwaru a způsobech ochrany proti němu. *Kaspersky.* [Online] [Citace: 21. 06 2016.] http://www.kaspersky.com/cz/internet-security-center/internet-safety/what-is-malware-and-how-to-protect-against-it.

132. Spurgeon, E. Charles a Zimmerman, Joann. *Ethernet switches.* Beijing : Oreilly & Associates Inc, 2013. ISBN 9781449367305.

133. Šafránek, David. Firewall - paketový filtr. *MUNI.* [Online] 12. 12 2000. [Citace: 21. 06 2016.] https://www.fi.muni.cz/~xsafran1/linux/netfilter.html.

134. Lammle, Todd. *CCNA routing and switching study guide.* Indianapolis : Wiley & Sons, 2013. ISBN 9781118749708.

135. Slížek, David. Jak si ministerstvo financí představuje blokování webů? *Lupa.cz.* [Online] 12. 04 2016. [Citace: 18. 06 2016.] http://www.lupa.cz/clanky/jak-si-ministerstvo-financi-predstavuje-blokovani-webu/.

136. Beal, Vangie. domain name resolvers. *Webopedia.* [Online] [Citace: 23. 06 2016.] http://www.webopedia.com/TERM/D/domain_name_resolvers.html.

Terminologický slovník

Termín	Význam
blacklist	Seznam objektů (IP adres, klíčových slov atd.) na něž jsou uvaleny restrikce (např. zákaz přístupu či využití služby). (128)
DNS resolver	Server, který přijímá DNS dotazy od klientských stanic, následně provádí překlad *doménových jmen* dotazováním jiných DNS serverů nebo z lokální cache a vrací odpověď klientovi. (129)
filtrace komunikace	Aplikování metod manipulace ze strany cenzora (viz 2.4) na komunikaci. V textu (kromě kapitoly 2.4) volně zaměnitelné s termínem *blokace komunikace.* (vlastní definice autora)
Internetové pirátství	Veškerá protiprávní činnost, která je realizována pomocí výpočetní techniky na Internetu a jejím důsledkem je porušování práv k duševnímu vlastnictví. (130)
klíčové slovo	Ve smyslu cenzury označuje textový řetězec, který bývá využit k identifikaci nežádoucího obsahu. Většinou se jedná o slova cenzurou zakázaná nebo reprezentující nežádoucí obsah. (vlastní definice autora)
malware	*Malware* označuje jakýkoli počítačový program, jehož cílem je provádět nežádoucí činnosti či poškozovat oprávněného uživatele počítače. (131)
nežádoucí komunikace	Komunikace, která je v rozporu s pravidly pro povolenou komunikaci nastavené cenzorem. V práci volně zaměnitelná s termínem *závadná komunikace.* (vlastní definice autora)

paket	V sekcích popisující síťovou komunikaci (viz 2.2) je chápán ve významu datové jednotky vyměňované na síťové vrstvě (132). Ve zbytku práce označuje jakýkoliv datový blok přenášený sítí. (18)
paketový filtr	Zařízení nebo proces pracující na síťové vrstvě, které zkoumá hlavičky procházejících paketů a na základě zdrojových a cílových IP adres vyhodnocuje, zdali je pošle dál. (133)
přímo připojená síť	*Přímo připojená síť* (angl. directly-connected network) je logické vyčlenění té části sítě, která je dosažitelná na vrstvě síťového rozhraní, čili pro komunikaci mezi jednotlivými uzly není potřeba směrování. (vlastní definice autora)
signatura protokolu	*Signatura protokolu* je skupina dat, která charakterizuje určitý protokol. Může se jednat o souhrn několika identifikátorů jako je velikost paketů, čísla portů, specifická část kódu v hlavičce a další. (vlastní definice autora)
směrovač	Zařízení nebo proces založené na protokolech síťové vrstvy využívané k propojení sítí. (132)
směrování	*Směrování* (angl. routing) je proces přesunu paketů mezi sítěmi za využití směrovačů. (134)
transportní protokol	Protokol sloužící k přenosu dat způsobem, který přináší do komunikace přidanou hodnotu, ale zároveň nijak neovlivňuje přenášený obsah. Ve smyslu cenzury slouží k ovlivnění komunikace tak, aby jí nebyla zachycena. (vlastní definice autora)

Seznam obrázků

Obr. 1.1: Mapa cenzury Internetu ve světě ... 5
Obr. 2.1: Zapouzdření aplikačních dat na vrstvách TCP/IP ... 14
Obr. 2.2: Formát Ethernetového rámce ... 14
Obr. 2.3: Hlavička protokolu IPv6 ... 15
Obr. 2.4: Hlavička protokolu UDP ... 16
Obr. 2.5: Hlavička protokolu TCP ... 17
Obr. 2.6: HTTP GET žádost o stránku www.vse.cz ... 17
Obr. 4.1: Testované metody obcházení cenzury ... 48

Seznam tabulek

Tab. 4.1: Dostupnost webových stránek ... 53
Tab. 4.2: Dostupnost portů cílového serveru ... 54
Tab. 4.3: Výsledky měření za pomocí lokální DNS databáze ... 55
Tab. 4.4: Výsledky úspěšnosti DNS metod v ČLR ... 56
Tab. 4.5: Výsledky úspěšnosti proxy serverů ... 57
Tab. 4.6: Výsledky úspěšnosti transportních protokolů ... 58

Příloha A

Tab. A1: Země dle hodnocení FOTN 2015 – výběr kategorie Not Free ... 75

Příloha B

Tab. B1: Testované webové stránky ... 76
Tab. B2: Specializované programy pro obcházení cenzury ... 77
Tab. B3: Specifikace VPN a protokolů ... 77
Tab. B4: Testované transportní protokoly ... 78
Tab. B5: Nástroje pro měření ... 78

Příloha C

Tab. C1: Výsledky DNS dotazů přes výchozí DNS server (DHCP) ... 79
Tab. C2: Výsledky DNS dotazů přes otevřené servery (OpenDNS) ... 80
Tab. C3: Výsledky DNS dotazů přes DNS Crypt ... 81

Příloha A: Výběr testované země

Tab. A1: Země dle hodnocení FOTN 2015 – výběr kategorie Not Free

Země	FOTN 2015 skóre (0=nejsvobodnější, 100=nejméně svobodná)
China	88
Syria	87
Iran	87
Ethiopia	82
Cuba	81
Uzbekistan	78
Vietnam	76
Saudi Arabia	73
Bahrain	72
Pakistan	69
United Arab Emirates	68
Sudan	65
The Gambia	65
Belarus	64
Thailand	63
Myanmar	63
Russia	62
Kazakhstan	61
Egypt	61

Zdroj: Freedom House (4)

Příloha B: Metodika praktického test obcházení cenzury

Tab. B1: Testované webové stránky

HTTP	typ	Herdict ID
https://www.iplocation.net/find-ip-address	Zjištění veřejné IP	39434
https://www.dropbox.com	Cloudové úložiště	9178
https://twitter.com/LordSnow	Micro-blogovací sociální síť	2633
http://cn.nytimes.com/world/20160429/c29syria/	Zpravodajský web (čínština)	19501
http://www.nytimes.com/2016/05/05/technology/personaltech/the-smartphone-way-to-inner-calm.html	Zpravodajský web (angličtina)	2480
https://search.yahoo.com	Vyhledávací server	41715

Zdroj: autor, herdict.org

Tab. B2: Specializované programy pro obcházení cenzury

Název	Verze	Platforma
Freegate	7.58 Professional Edition	Windows
Hotspot Shield	5.3.2	Windows
Lantern	2.1.2	Windows
Psiphon 3	115	Windows
Tor Browser	5.5.5	Linux
UltraSurf	15.04	Windows
Your Freedom	20160419-01	Linux

Zdroj: autor

Tab. B3: Specifikace VPN a protokolů

Typ	Verze	Platforma
OpenVPN	2.3.10	Linux
OpenSSL	1.0.2g-fips	Linux

Zdroj: autor

Tab. B4: Testované transportní protokoly

Název
FTE
MEEK – AMAZON
MEEK - AZURE
MEEK – GOOGLE
OBFS3
OBFS4
SCRAMBLESUIT

Zdroj: autor

Tab. B5: Nástroje pro měření

Název	Verze	Platforma
DiG	9.10.3	Linux
Internet Explorer	11.306.10586.0	Windows
Nmap	7.01	Linux
ProxyChains	3.1	Linux
Tcpdump	4.7.4	Linux
Traceroute	2.0.21	Linux
Wget	1.17.1	Linux
Wireshark	2.0.3	Windows

Zdroj: autor

Příloha C: Výsledky praktického testu obcházení cenzury

Tab. C1: Výsledky DNS dotazů přes výchozí DNS server (DHCP)

	Hong Kong		Chang-čou		Si-an		Peking		Šanghaj	
cn.nytimes.com	54.239.174.79	✓	108.160.170.46	✖	199.16.158.76	✖	199.16.158.82	✖	108.160.166.10	✖
dropbox.com	108.160.172.206	✓	37.61.54.158	✖	37.61.54.158	✖	205.251.193.59	✖	37.61.54.158	✖
nytimes.com	170.149.159.130	✓	31.13.79.194	✖	199.101.134.172	✖	199.16.158.82	✖	199.16.158.77	✖
twitter.com	199.59.150.7	✓	37.61.54.158	✖	37.61.54.158	✖	208.78.70.34	✖	37.61.54.158	✖

Zdroj: autor (Příloha D:/dns/*/dig-DEFAULT; Prague/DEFAULT/)*

Tab. C2: Výsledky DNS dotazů přes otevřené servery (OpenDNS)

	Hong Kong		Chang-čou		Si-an		Peking		Šanghaj	
cn.nytimes.com	54.239.174.80	✓	31.13.68.36	✖	74.117.178.12	✖	69.171.235.80	✖	108.160.166.10	✖
dropbox.com	108.160.172.238	✓	37.61.54.158	✖	37.61.54.158	✖	37.61.54.158	✖	159.106.121.75	✖
nytimes.com	170.149.159.130	✓	31.13.74.32	✖	74.117.178.125	✖	199.16.158.180	✖	199.16.158.77	✖
twitter.com	199.59.148.82	✓	37.61.54.158	✖	93.46.8.89	✖	37.61.54.158	✖	37.61.54.158	✖

Zdroj: autor (Příloha D:/tcpdump/opendns.pcap; Prague/opendns/)*

Tab. C3: Výsledky DNS dotazů přes DNS Crypt

	Hong Kong		Chang-čou		Si-an		Peking		Šanghaj	
cn.nytimes.com	54.239.174.79	✓	54.230.191.26	✓	*nezměřeno*		54.230.254.191	✓	52.85.148.90	✓
dropbox.com	108.160.172.206	✓	108.160.172.238	✓	*nezměřeno*		108.160.172.238	✓	108.160.172.238	✓
nytimes.com	170.149.159.130	✓	170.149.159.130	✓	*nezměřeno*		170.149.159.130	✓	170.149.159.130	✓
twitter.com	199.59.150.39	✓	199.59.148.82	✓	*nezměřeno*		199.59.148.10	✓	199.59.150.7	✓

Pozn.: V lokalitě Si-an nebyl proveden test pomocí DNSCrypt, proto není možné prezentovat výsledky.
Zdroj: autor (Příloha D:/dns/*/dig-dnscrypt; Prague/dnscrypt/)*

Příloha D: Soubory s výsledky praktického testu

Součástí práce je DVD se soubory, které vznikly jako výstup praktického testu a slouží jako zdroj pro výsledky měření kapitoly 4.3. Členění složek odpovídá městům, ve kterých bylo měření provedeno, a následné podsložky reprezentují metody či nástroje, pomocí kterých byl získán výstup.