Internet Research Task Force (IRTF)
Request for Comments: 9505
Category: Informational
ISSN: 2070-1721

J. L. Hall
Internet Society
M. D. Aaron
CU Boulder
A. Andersdotter

B. Jones

N. Feamster
U Chicago
M. Knodel
Center for Democracy & Technology

November 2023

A Survey of Worldwide Censorship Techniques

Обзор методов цензуры в мире

Аннотация

В этом документе описаны технические механизмы сетевой цензуры, применяемые режимами по всему миру для блокировки или нарушения трафика Internet. Целью документа является ознакомление разработчиков, внедренцев и пользователей протоколов Internet со свойствами и механизмами, применяемые для цензурирования доступа конечных пользователей к информации. Документ не содержит предложений по рассмотрению отдельных протоколов и является лишь информационным, предназначенным служить справкой. Документ является результатом работы исследовательской группы IRTF по повышению и оценке приватности (Privacy Enhancement and Assessment или PEARG).

Статус документа

Документ не относится к категории Internet Standards Track и публикуется для информации.

Документ является результатом работы IRTF[1]. IRTF публикует результаты относящихся к Internet исследований и разработок. Эти результаты могут оказаться не пригодными для реализации. Данный RFC представляет согласованное мнение исследовательской группы Privacy Enhancements and Assessmentsв рамках IRTF. Документы, одобренные для публикации IRSG, не претендуют на статус Internet Standard (см. раздел 2 в RFC 7841).

Информацию о текущем статусе документа, ошибках и способах обратной связи можно найти по ссылке https://www.rfc-editor.org/info/rfc9505.

Оглавление

[1]Internet Research Task Force - комиссия по исследовательским задачам Internet.

1. Введение

Цензурой называют подавление сочтённых нежелательными, вредными, деликатными или неудобными [WP-Def-2020] субъектом, обладающий властью (например, правительством, организацией или частным лицом). Хотя цензоры, занимающиеся соответствующими делами, должны делать это с помощью правовых, военных или иных мер, этот документ сосредоточен на технических механизмах, применяемых для цензуры в сети.

Документ описывает технические механизмы, которые режимы с цензурой по всему миру применяют для блокировки и ограничения трафика Internet. В [RFC7754] обсуждается блокировка и фильтрация с точки зрения влияния на архитектуру Internet, а не на доступ конечных пользователей к содержимому и услугам. Расширяются и академические исследования обхода цензуры (см. обзорную статью [Tschantz-2016]), результаты которых приводятся здесь для информирования разработчиков протоколов и их реализаций.

Обход цензуры влияет на стоимость реализации цензурных мер и здесь рассматриваются такие расходы применительно к рассматриваемым техническим методам.

Документ прошёл широкое обсуждение и рецензирование в исследовательской группе PEARG и представляет согласованное мнение группы. Это не результат работы IETF и не стандарт.

2. Терминология

Здесь описывается три элемента цензуры в Internet - предписания (prescription), идентификация (identification) и вмешательство (interference). Документ содержит три основных раздела в соответствии с этими элементами. Предписания - это процесс, с помощью которого цензоры определяют, какие типы материалов следует подвергать цензуре (например, классификация порносайтов как нежелательных). Идентификация - это процесс, с помощью которого цензоры классифицируют конкретный трафик, или идентификаторы трафика для блокировки или ограничения (например, принимается решение о нежелательности web-страниц, содержащих sex в заголовке HTTP, или трафика от URL www.sex.example). Вмешательство - это процесс, с помощью которого цензоры вмешиваются во взаимодействие и предотвращают доступ к недозволенным материалам путём блокирования доступа или ухудшения соединения (например, за счёт реализации технического решения, способного идентифицировать заголовки HTTP или URL и обеспечивать для недозволенных материалов полную или частичную недоступность).

3. Технические предписания

Предписания - это процесс выяснения что цензоры хотели бы заблокировать [Glanville-2008]. Обычно цензоры собирают сведения «для блокирования» в списки блокировки (blocklist), базы хешей изображений [ekr-2021] или применяют эвристическую оценку содержимого в реальном масштабе времени [Ding-1999]. Некоторые национальные сети устроены для более естественного функционирования в качестве точек контроля [Leyba-2019]. Есть признаки применения сетевыми цензорами методов вероятностного машинного обучения [Tang-2016]. В некоторых юрисдикциях активными направлениями исследований в части выявления содержимого, представляющегося аморальным или коммерчески вредным для компаний или потребителей, являются сканирование web (crawling) и машинное обучение [SIDN-2020].

Обычно в списках блокировки имеется несколько элементов - ключевое слово, доменное имя, протокол или адрес IP. Блокирование по ключевым словам и именам доменов выполняется на прикладном уровне (например, HTTP), для блокирования по протоколам часто применяется глубокий просмотр пакетов (deep packet inspection или DPI) для определения запрещённых протоколов, блокирование по IP обычно выполняется по адресам IP в заголовках IPv4/IPv6. Некоторые цензоры используют присутствие определённых ключевых слов для подключения более жёстких списков блокировки [Rambert-2021] или более щадящего отношения к содержимому [Knockel-2021].

Механизмы создания списков блокировки могут быть разными. Цензоры могут приобретать у частных компаний программы контроля содержимого, позволяющие фильтровать трафик из широких категорий, которые хотелось бы заблокировать, таких как азартные игры или порнография [Knight-2005]. В таких случаях эти частные службы пытаются классифицировать каждый сомнительный (semi-questionable) web-сайт для обеспечения возможности блокировки по

метатегам. Аналогичным способом настраиваются эвристические системы для сопоставления оценок с категориями нежелательного или спорного содержимого.

В странах, заинтересованных в сохранении определённого политического контроля, обычно имеются министерства или организации, поддерживающие списки блокировок. Примерами являются Министерство промышленности и информационных технологий (Ministry of Industry and Information Technology) в Китае, Министерство культуры и мусульманского наставничества (Ministry of Culture and Islamic Guidance) в Иране, организации, занимающиеся вопросами авторских прав во Франции [HADOPI] и защитой прав потребителей в Евросоюзе (EU) [Reda-2017].

Фильтрация содержимого в изображениях и видео отребует от организаций хранить в базах данных для блокировки хэш-значения изображений или видео, с которыми может (с некоторой точностью) сопоставляться содержимое, которое передаётся, принимается или записывается с использованием централизованных приложений и служб для работы с содержимым [ekr-2021].

4. Техническая идентификация

4.1. Точки контроля

Цензура Internet происходит во всех частях топологии сети. Это может быть реализовано в самой сети 9например, на локальном или транзитном канале), на серверной стороне коммуникаций (например, web-хосты, облачные провайдеры, сети доставки содержимого), в экосистеме вспомогательных служб (например, DNS или удостоверяющий центр), на стороне клиента (например, в пользовательском устройстве, таком как смартфон, переносный или настольный компьютер, программы, выполняемые на устройствах). Важным аспектом повсеместного технического перехвата является необходимость полагаться на программы или оборудование для перехвата интересующего цензора содержимого. Имеются различные физические и логические точки контроля, где серверы могут применять механизмы перехвата. Некоторые из таких точек перечислены ниже.

Магистраль Internet

Если цензор контролирует элементы сетевой инфраструктуры Internet, такие как международные шлюзы в регионе или точки обмена трафиком (Internet Exchange Point или IXP), эти точки могут служит для фильтрации нежелательного трафика в регион или из него путём перехвата пакетов и отображения (mirroring) портов. Цензура на шлюзах наиболее эффективна для контроля потока информации между регионом и остальной частью Internet, не неэффективна для идентификации содержимого между пользователями внутри региона, который следует реализовать в точках обмена трафиком или иных точках агрегирования. Некоторые национальные сети эффективно реализуют точки «дросселирования» (choke) и иного контроля [Leyba-2019].

Internet-провайдеры (ISP)

ISP часто служат точками контроля. Их преимущество заключается в простоте учёта цензором, они часто оказываются в юрисдикции или оперативном управлении цензора бесспорным образом, а дополнительным свойством является возможность ISP идентифицировать локальный и международный трафик всех своих пользователей. Механизмы фильтрации могут быть установлены цензором у ISP на основе государственных указаний, владения, добровольно или принудительно.

Организации

Частные организации, такие как корпорации, школы, Internet-кафе, могут применять механизмы фильтрации. Эти механизмы иногда создаются по запросу государственного цензора, но могут внедряться и для достижения целей организации, например, формирования у школьников определённых моральных установок, независимо от целей общества или государства.

Сети распространения содержимого (CDN)

CDN стремятся сжать топологию сети для размещения содержимого ближе к пользователям сервиса. Это сокращает задержки на передачу содержимого и повышает QoS. Серверы содержимого службы CDN размещаются близко к пользователю в сетевом смысле и могут стать мощными точками контроля для цензоров, особенное если размещение репозиториев CDN позволяет легко вмешиваться в работу.

УЦ (CA) в инфраструктуре открытых ключей (PKI)

Органы, выпускающие криптографически защищённые ресурсы, могут быть значимыми точками контроля. CA, выдающий держателям доменов сертификаты для TLS/HTTPS (Web PKI) а также региональные или локальные регистраторы (Regional/ Local Internet Registry или RIR/LIR) выдающие полномочия на создание маршрутов (Route Origin Authorization или ROA) операторам BGP, могут быть вынуждены выпускать неконтролируемые сертификаты, что может приводить к компрометации (т. е. позволять программам цензоров осуществлять идентификацию и вмешательство так, где раньше это было невозможно). CA можно также вынудить к отзыву сертификатов. Это может приводить к недобросовестной маршрутизации трафика, возможности перехвата TLS или невозможности защищённого взаимодействия между легитимным отправителем и получателем потока трафика.

Службы

На поставщиков услуг приложений может оказываться давление, принуждение или правовое воздействие, вынуждающее их к цензуре определённого содержимого или потоков данных. Поставщики услуг, естественно, заинтересовано в максимально широкой базе потенциальных клиентов и возможное прекращение предоставления услуг или правовые последствия в связи с цензурой могут быть для них менее важны, нежели возможное исключение содержимого, пользователей или применения их услуг. Службы все чаще становятся центральной точкой дискуссий о цензуре, а также обсуждения моральных императивов применения инструментов цензуры.

Сайты содержимого

На серверной стороне коммуникаций находится множество платформ, публикующих созданное пользователями содержимое и требующих соблюдения условий обслуживания для всего содержимого и учётных записей пользователей, чтобы избавить владельцев хостинга от юридической ответственности. В совокупности эти правила, действия и средства защиты называют модерацией содержимого. Эта модерация осуществляется выше уровня служб или приложений, но механизмы настроены так, чтобы фильтровать, сортировать и блокировать содержимое и пользователей, что делает такие их подверженными цензуре через прямое давление на частную организацию.

Персональные устройства

Цензоры могут потребовать установки программ для цензуры на уровне устройств. С этим связано много недостатков в плане расширяемости, простоты обхода и требований к операционной системе (конечно если персональное устройство перед продажей обрабатывается программами для цензуры и его сложно

перенастроить, это может сыграть на руку тем, кто стремится контролировать информацию, например, для детей, студентов, клиентов или сотрудников). Появление мобильных устройств усугубило эти проблемы. Упомянутые программы могут быт сделаны обязательными для использования институциональными субъектами, действующими в соответствии с моральными императивами, не задаваемым государством.

На всех уровнях сетевой иерархии механизмы фильтрации, применяемые для цензурирования нежелательного трафика, по сути, одинаковы - цензор или напрямую указывает нежелательное содержимое с помощью описанных ниже идентификаторов и применяет блокировку или формовку (например, как описано ниже для предотвращения или затруднения доступа) или передаёт выполнение этих функций другому субъекту, не являющемуся цензором (например, частной компании). Идентификация нежелательного трафика может происходить на прикладном, транспорном или сетевом уровне стека IP. Цензоры часто сосредоточены на web-трафике, поэтому связанные с ним протоколы фильтруются предсказуемыми способами (параграфы 4.2.1 и 4.2.2). Например, подрывное изображение может пройти через фильтр ключевых слов, однако если потом это изображение будет сочтено нежелательным, цензор может внести в список блокировки IP-адрес сайта-провайдера.

4.2. Прикладной уровень

В последующих параграфах описываются свойства и компромиссы, связанные с цензурой на основе фильтрации по сведениям прикладного уровня. В каждый параграф включены Примеры из практики, описывающие распространённые варианты поведения.

4.2.1. Идентификация заголовков запросов HTTP

Заголовок HTTP содержит много полезных для идентификации трафика сведений. Хотя единственным обязательным полем заголовка запроса HTTP (начиная с HTTP/1.1) является host, поле метода HTTP требуется для выполнения каких-либо полезных действий. Поэтому для вездесущей цензуры чаще всего применяются поля method и host. Цензор может просматривать (sniff) трафик и идентифицировать конкретное доменное имя (host), а обычно и имя страницы (напрмиер, GET /page). Этот метод идентификации обычно применяется в паре с идентификацией транспортных заголовков (см. параграф 4.3.1) для повышения надёжности.

Компромиссы. Идентификация заголовков запросов HTTP технически проста и может быть легко реализована на уровне магистрали или ISP. Нужное для этого оборудование дёшево и легко доступно, что делает его привлекательным, когда важную роль играет бюджет и масштабы. Протокол защищённой доставки гипертекста HTTPS (Hypertext Transport Protocol Secure) шифрует соответствующие поля запросов и откликов, поэтому для фильтрации HTTPS нужна ещё идентификация транспорта (см. параграф 4.3.1). Однако некоторые контрмеры могут тривиально преодолеть простые формы идентификации заголовков запросов HTTP. Например, взаимодействующие конечные точки (web-сервер и клиент) могут шифровать или иным способом скрывать (obfuscate) поле host в запросе, что может помешать методам сопоставления по значению поля host в заголовке запроса.

Примеры из практики. Исследования механизмов цензуры выявили факты фильтрации по заголовкам HTTP и/или URL во многих странах, включая Бангладеш, Бахрейн, Китай, Индию, Иран, Малайзию, Пакистан, Россию, Саудовскую Аравию, Южную Корею, Тайланд, Турцию [Verkamp-2012] [Nabi-2013] [Aryan-2013]. Цензоры часто покупают коммерческие технологии [Dalek-2013], сочетающие идентификацию заголовков запросов HTTP и транспортных заголовков для фильтрации конкретных URL. Dalek с соавторами [Dalek-2013] и Jones с соавторами [Jones-2014] выявили использование таких решений на практике.

4.2.2. Идентификация заголовков откликов HTTP

Если идентификация заголовков запросов HTTP полагается на сведения из запроса клиента к серверу HTTP, то для идентификации заголовков в откликах HTTP, позволяющей выявить нежелательное содержимое, служат сведения, передаваемые сервером клиенту.

Компромиссы. Как и методы идентификации заголовков запросов HTTP, методы идентификации трафика HTTP общеизвестны, дешёвы и сравнительно просты в реализации. Однако при использовании HTTPS они бесполезны, поскольку HTTPS шифрует отклик и его заголовки.

Поля отклика менее полезны для идентификации содержимого по сравнению с полями запроса, поскольку значение поля Server легко определить при идентификации заголовков запроса HTTP, а поле Via редко бывает важным. Механизмы цензуры откликов HTTP обычно пропускают первые n пакетов, пока обрабатывается отображённый трафик. Это может вести к пропуску части содержимого, что позволяет пользователю заметить активное вмешательство цензора в нежелательное содержимое.

Примеры из практики. А 2009 г. Jong Park с соавторами в университете New Mexico показали, что китайский межсетевой экран Great Firewall of China (GFW) использует идентификацию заголовков в откликах [Crandall-2010]. Однако Jong Park и др. обнаружили, что GFW прекратил это в процессе изучения. Ввиду перекрытия фильтрации по откликам HTTP и ключевым словам (см. параграф 4.2.4), можно предположить с достаточным основанием, что большинство цензоров полагается на фильтрацию потоков TCP по ключевым словам, а не фильтрацию откликов HTTP.

4.2.3. Защита на транспортном уровне (TLS)

Как и для HTTP, цензоры применяют множество методов для цензурирования TLS (и HTTPS). Большинство из этих методов основано на поле индикации имени сервера (Server Name Indication или SNI), включая цензурирование SNI, зашифрованных SNI (Encrypted SNI или ESNI) и пропущенных SNI. Цензоры могут также отслеживать содержимое HTTPS по сертификатам серверов. Отметим, что TLS 1.3 выступает как защитный компонент в протоколе QUIC.

4.2.3.1. Индикация имени сервера (SNI)

В зашифрованных соединениях TLS могут участвовать серверы, на которых размещается множество виртуальных серверов с одним сетевым адресом и клиент должен указать в сообщении ClientHello доменное имя, с которым он хочет соединиться (чтобы сервер мог ответить подходящим сертификатом TLS), используя расширение TLS SNI [RFC6066]. В TLS на основе протокола TCP сообщения ClientHello не шифруются. При использовании протокола QUIC сообщение ClientHello шифруется, но это не обеспечивает его эффективной защиты, поскольку начальные ключи шифрования выводятся с использованием значения, доступного в линии. Поскольку SNI часто передаётся в открытом

виде (как и передаваемые в ответ поля сертификата), цензоры и программы фильтрации могут использовать его в целях блокировки, фильтрации или вмешательства путём сброса соединений с доменами, соответствующими запретному содержимому (например, bad.foo.example можно фильтровать, а good.foo.example - пропускать) [Shbair-2015]. В рабочей группе TLS предпринимаются усилия по стандартизации шифрования SNI [RFC8744] [TLS-ESNI] и недавние исследования дают многообещающие результаты при использовании ESNI в условиях фильтрации по SNI [Chai-2019] в некоторых странах.

Одним из популярных способов избежать идентификации цензорами является подстановка домена (domain fronting) [Fifield-2015]. Чтобы избежать идентификации приложения указывают в расширении SNI подставное имя домена, а настоящее указывают в заголовке host, защищённом HTTPS. Видимое значение SNI указывает разрешенный домен, а заблокированный скрывается в зашифрованном заголовке приложения. Некоторые службы шифрованных сообщений полагались на подстановку доменов в странах, использующих фильтрацию по SNI. Эти службы использовали для прикрытия домены, на уровне которых задавать блокирование нежелательно. Однако компании, владеющие большинством популярных доменов, перенастроили свои программы так, чтобы предотвратить такие злоупотребления. Возможно в будущем удастся достичь похожих результатов путём шифрования SNI.

Компромиссы. Некоторые клиенты не передают расширение SNI (например, клиенты, поддерживающие лишь SSL, но не TLS), что делает указанный метод неэффективным (см. параграф 4.2.3.3). Кроме того, для этого метода нужен глубокий просмотр пакетов (DPI), что может быть дорогостоящим в плане вычислительных ресурсов и инфраструктуры, особенно при использовании с протоколом QUIC, где для DPI требуется извлечение ключа и расшифровка ClientHello, чтобы прочесть SNI. Неаккуратная настройка блокировки по SNI может приводить к избыточной блокировке трафика, например, в результате случайной блокировки домена второго уровня вроде populardomain.example. В случае применения ESNI давление на цензуру может быть перенесено в другие точки вмешательства, такие как поставщики содержимого и приложений.

Примеры из практики. Имеется множество фирм, предлагающих средства фильтрации на основе SNI [Trustwave-2015] [Sophos-2023] [Shbair-2015]. Правительства Китая, Египта, Ирана, Катара, Южной Кореи, Турции, Туркменистана и Объединенных Арабских Эмиратов широко применяют фильтрацию и блокировку SNI [OONI-2018] [OONI-2019] [NA-SK-2019] [CitizenLab-2018] [Gatlan-2019] [Chai-2019] [Grover-2019] [Singh-2019]. Блокировка SNI для трафика QUIC впервые была отмечена в России в марте 2022 г. [Elmenhorst-2022].

4.2.3.2. Зашифрованные SNI (ESNI)

С учётом утечки данных SNI естественной реакцией является шифрование поля, что и было сделано в TLS 1.3 с помощью шифрованных приветствий клиента (Encrypted Client Hello или ECH). До ECH для предотвращения утечек, вызываемых SNI, применялось расширение ESNI, где шифровалось само поле SNI. К сожалению, цензоры могут настроиться на блокировку соединений, использующих ESNI. Это гарантированно вызовет избыточную блокировку, но может иметь смысл для цензоров, пока ESNI ещё не получили широкого распространения внутри страны. ECH является новым стандартом для защиты всего TLS ClientHello, но ещё не получил широкого распространения.

Компромиссы. Стоимость цензуры ESNI значительно выше по сравнению с SNI, поскольку цензор уже не может направить свои действия на конкретные домены и гарантированно вызовет избыточную блокировку. В таких случаях цензор использует избыточную блокировку для полного запрета использования ESNI.

Примеры из практики. В 2020 г. в Китае началось цензурирование всех применений ESNI [Bock-2020b], даже для безобидных соединений. Механизм цензуры ESNI в Китае отличается от механизма для соединений на основе SNI, что позволяет предположить развёртывание новых промежуточных устройств специально для соединений ESNI.

4.2.3.3. Отсутствие SNI

Исследователи заметили, что некоторые клиенты полностью опускают расширение SNI, что ограничивает доступные цензору сведения. Как и в случае с ESNI, цензоры могут полностью блокировать соединения без SNI, хотя это и вызовет избыточную блокировку.

Компромиссы. Блокирование всех соединений без поля SNI ведёт к избыточной блокировке, однако соединения без SNI на практике встречаются сравнительно редко.

Примеры из практики. В прошлом исследователи наблюдали в России блокировку цензорами соединений без поля SNI [Bock-2020b].

4.2.3.4. Сертификат отклика сервера

В процессе согласования TLS после TLS ClientHello сервер будет отвечать сертификатом TLS. Этот сертификат содержит домен, к которому пытается обратиться клиент, открывая дополнительную возможность для цензуры. Этот метод не будет работать с TLS 1.3, где сертификат шифруется.

Компромиссы. Цензура на основе сертификатов сервера требует методов DPI, что может быть более ресурсоёмким по сравнению с другими методами. Кроме того, сертификат в согласовании TLS передаётся позже по сравнению с полем SNI, что требует от цензора отслеживать соединение дольше.

Примеры из практики. Исследователи наблюдали, как Reliance Jio ISP в Индии использует поля отклика с сертификатом для цензурирования соединений [Satija-2021].

4.2.4. Инструментальное воздействие на распространителей содержимого

Правительства многих стран вынуждают поставщиков содержимого вводить самоцензуру или создают правовую базу, в рамках которой распространители содержимого (content distributor) стимулируются следовать предпочтениям по ограничению содержимого от внешних агентов [Boyle-1997]. По причине обширности сферы применения такой цензуры распространителями содержимого считаются любые службы, предоставляющие услуги пользователю, от web-сайтов до хранилищ локально установленных программ.

Общепринятым методом инструментального воздействия на распространителей содержимого является идентификация ключевых слов для обнаружения запретных слов на платформах распространителей. Правительства могут предоставлять списки таких слов, а распространители могут создавать свои списки.

Все чаще применяется метод контроля за распространением содержимого на основе сопоставления хэш-значений для обнаружения и выполнения действий применительно к изображениям и видео, для которых заданы ограничения правительством, учреждениями, организациями или самими распространителями содержимого.

Другим методом влияния на распространителей содержимого является требование не вступать во взаимодействие с некоторыми категориями пользователей (см. параграф 6.4).

Компромиссы. Предоставляя распространителям содержимого средства идентификации ограниченного содержимого или его поставщиков, цензоры могут получать новые сведения за счёт политического капитала компаний, которые они заставляют или поощряют участвовать в цензуре. Например, цензор может получить представление о содержимом шифрованного трафика, вынудив web-сайты идентифицировать содержимое, на которое наложены ограничения. Принуждение распространителей содержимого влиять на пользователей, категории пользователей, содержимое и его поставщиков в части применения самоцензуры является дополнительным средством для цензоров (см. параграф 6.2). Компромисс при инструментальном воздействии на распространение содержимого сильно зависит от поставщика содержимого и требуемой поддержки. Типичная проблема заключается в том, что целевые наборы ключевых слов или категорий слишком широки и создают риск избыточной блокировки или не проходят достаточно строгой правовой проверки перед обязательным применением (см. стр. 8 в [EC-2012]).

Примеры из практики. Исследователи обнаружили идентификацию ключевых слов поставщиками содержимого от платформ обмена мгновенными сообщениями [Senft-2013] до поисковых систем [Rushe-2014] [Cheng-2010] [Whittaker-2013] [BBC-2013] [Condliffe-2013]. Для демонстрации распространённости этого типа идентификации ключевых слов следует рассмотреть цензуру в поисковых системах.

Цензура поисковых систем демонстрирует идентификацию ключевых слов поставщиками содержимого и может носить как региональный, так и планетарный характер. Иногда реализация бывает добровольной, но обычно она основана на законах и правилах страны, где работает поисковая машина. Списки блокировки по ключевым словам, скорей всего, поддерживаются провайдерами поисковых систем. Известно, что в Китае от провайдеров поисковых машин требуют «добровольно» поддерживать списки блокировки поисковых слов для получения и сохранения лицензии поставщика содержимого (Internet Content Provider или ICP) [Cheng-2010]. Очевидно, что такие списки поддерживает каждый провайдер поисковой машины, судя по незначительным вариациям перехватываемого поиска [Zhu-2011] [Whittaker-2013]. В Великобритании подталкивают поисковые системы к самоцензуре, угрожая судебным разбирательством в случае отказа - Google и Microsoft согласились блокировать более 100 000 запросов в Великобритании, чтобы помочь в борьбе со злоупотреблениями [BBC-2013] [Condliffe-2013]. Законодательство Европейского союза и США требует изменять результаты поиска в ответ на претензии в части авторских прав, торговых знаков, защиты данных и дискредитации [EC-2012].

Сложность обнаружения идентификации ключевых слов зависит от поисковой выдачи. В некоторых случаях специализированный или пустой результат позволяет легко обнаружить блокировку, но более тонкую цензуру заметить сложнее. В феврале 2015 г. поисковую машину Bing компании Microsoft обвинили в цензуре китайского содержимого за пределами Китая [Rushe-2014], поскольку выдача Bing различалась для запросов на китайском и английском языке. Однако не исключено, что цензура самой большой базы китайских пользователей поиска - Китая - исказила результаты так, что более популярные в Китае результаты (без цензуры) оказались более популярны у говорящих на китайском языке и за пределами Китая.

Дистанцирование дистрибьюторов содержимого от некоторых категорий пользователей произошло, например, в Испании в результате конфликта движения за независимость Каталонии с испанской правовой презумпцией унитарного государства [Lomas-2019]. Организаторы соревнования по киберспорту (E-sport) отмежевались от ведущих игроков, выразивших свои политические взгляды в связи с протестами 2019 г. в Гонконге [Victor-2019] (см. параграф 5.3.1).

4.2.5. DPI-идентификация

К DPI технически относится любой анализ пакетов глубже адресов IP и номеров портов и в последние годы это стало реализуемо вычислительными средствами в качестве механизма цензуры [Wagner-2009]. В отличие от других методов, DPI восстанавливает (reassemble) сетевые потоки для проверки разделов данных приложения, а не только заголовков, поэтому часто служит для идентификации ключевых слов. DPI отличается от других методов идентификации использованием дополнительных методов идентификации содержимого, например, размера и временных характеристик пакетов. Для предотвращения значительного влияния на QoS в DPI обычно анализируются копии данных, а исходные пакеты маршрутизируются обычным путем. Трафик обычно расщепляется отображающим коммутатором или оптическим разветвителем (splitter) и анализируется кластером машин, на которых работают системы обнаружения вторжений (Intrusion Detection System или IDS), настроенные для цензуры.

Компромиссы. DPI является одним из наиболее дорогих маханизмов идентификации и сильно влияет на QoS [Porter-2005]. При использовании для фильтрации ключевых слов в потока TCP системы DPI могут приводить к избыточной блокировке. Как и другие мтоды, DPI мало подходит для шифрованных данных, хотя для идентификации трафика DPI может использовать для идентификации трафика нешифруемые элементы потока шифрованных данных (например, SNI в TLS) или метаданные шифрованного потока (например, размеры пакетов, которые различаются для текстовых и видео-потоков). Дополнительные сведения о фильтрации по SNI приведены в параграфе 4.2.3.1.

Дополнительные сведения могут быть получены при сравнении некоторых нешифруемых элементов согласования TLS с аналогичными элементами данных из известных источников. Такая практика, называемая «отпечатками TLS» (fingerprinting), позволяет вероятностно идентифицировать операционные системы сторон, браузеры и приложения на основе конкретных комбинаций версии TLS, шифров, опций сжатия и т. п., передаваемых в сообщениях ClientHello, путём сравнения с похожими сигнатурами нешифрованного трафика [Husak-2016].

Несмотря на отмеченные сложности, DPI является мощным методом идентификации и широко применяется на практике. Великий китайский брандмауэр GFW - крупнейшая система цензуры в мире - использует DPI для обнаружения запретного содержимого в HTTP и DNS, а также внедрения в соединения пакетов TCP RST и негативных откликов DNS [Crandall-2010] [Clayton-2006] [Anonymous-2014].

Примеры из практики. В ряде исследований были обнаружены свидетельства применения DPI для цензуры содержимого и воздействия на него. Clayton с соавторами, Crandal и др., Anonymous и Khattak с соавторами исследовали GFW [Crandall-2010] [Clayton-2006] [Anonymous-2014]. Khattak с соавтоорами даже исследовали

брандмауэр для обнаружения деталей реализации, таких как число хранимых состояний [Khattak-2013]. Проект Tor заявляет, что Китай, Иран, Эфиопия и другие страны наверняка применяли DPI для блокировки протокола obfs2 [Wilde-2012]. Малайзию обвиняют в использовании целевой инспекции DPI в сочетании с DDoS для идентификации и последующей атаки на материалы оппозиции [Wagstaff-2013]. Представляется вероятным, что организации, не блокирующие содержимое в реальном масштабе времени, могут применять DPI для сортировки и классификации собранного трафика с использованием высокоскоростной обработки пакетов [Hepting-2011].

4.3. Транспортный уровень

4.3.1. Неглубокая проверка пакетов и идентификация заголовков транспорта

Среди методов поверхностной проверки пакетов идентификация транспортных заголовков является наиболее распространенной, надежной и предсказуемой. Транспортные заголовки содержат несколько бесценных элементов информации, которые наиболее очевидны для успешной маршрутизации трафика - адреса IP и номера портов отправителя и получателя. Поля IP отправителя и получателя ценны вдвойне, поскольку они не только позволяют цензору блокировать нежелательное содержимое по IP, но и дают цензору возможность определить адрес пользователя, передающего запрос, и адрес вызываемой службы, по которому в большинстве случаев можно определить посещенный домен [Patil-2019]. Номер порта полезен для списков разрешённых приложений.

Сочетание адресов IP, портов и сведений о протоколе в транспортном заголовке позволяет цензору использовать поверхностную инспекцию пакетов для идентификации конкретных конечных точек TCP или UDP. Блокировка конечных точек UDP наблюдалась в контексте блокирования QUIC [Elmenhorst-2021].

Компромиссы. Идентификация заголовков популярна из-за её простоты, доступности и надёжности. Идентификация заголовков тривиально реализуется в некоторых маршрутизаторах, но сложна для реализации в маршрутизаторах магистралей и ISP, поэтому там она обычно реализуется с помощью DPI. Включение IP в список блокировки эквивалентно установке конкретного маршрута в маршрутизаторе (например, маршрута /32 для IPv4 или /128 для IPv6). Однако из-за ограниченного размера таблицы потоков этот в список блокировки невозможно включить более нескольких тысяч IP. Кроме того, такая блокировка является достаточно грубой и часто избыточна для некоторых служб вроде сетей CDN, где содержимое связано с сотнями или тысячами адресов IP. Несмотря на эти недостатки, блокировка по IP очень эффективна, поскольку для её обхода пользователю приходится передавать трафик через прокси. Кроме того, блокировка по IP работает против любых протоколов, работающих на основе IP, например, TCP или QUIC.

Блокровка портов обычно бесполезна, поскольку один порт может использоваться для разных типов содержимого, а приложения могут менять номер порта. Например, большая часть трафика HTTP идёт через порт 80, поэтому цензор не может различить нежелательное и разрешённое содержимое web лишь по номеру порта. HTTPS работает через порт 443, что влечёт для цензова аналогичные последствия, которому, кроме того, доступна лишь часть метаданных. Иногда применяются списки портов, с помощью которых цензор ограничивает взаимодействие (например, разрешается лишь порт 80 для трафика HTTP), и эти списки наиболее эффективны в сочетании с другими методами идентификации. Например, цензор может блокировать принятый по умолчанию порт 443 для HTTPS, вынуждая многих пользователей вернуться к HTTP. В качестве контрпримера можно привести блокировку порта 25 (SMTP), давно применяемую в сетях домашних провайдеров для снижения объёма спама, однако это не позволяет клиентам таких ISP запускать свои почтовые серверы.

4.3.2. Идентификация протокола

Иногда цензоры блокируют некоторые протоколы целиком, используя различные характеристики трафика. Например, Иран снижает производительность трафика протокола HTTPS, препятствующего дальнейшему анализу, чтобы вынудить пользователей перейти на протокол HTTP, который можно анализировать [Aryan-2013]. Простая идентификация протокола будет распознавать трафик TCP через порт 443 как HTTPS, но изощренный анализ статистических свойств данных содержимого и поведения потоков будет более эффективным, даже если порт 443 не используется [Hjelmvik-2010] [Sandvine-2015].

Если цензоры обнаружат средства обхода, они могут блокировать их. Поэтому цензоры, такие как Китай, очень заинтересованы в идентификации протоколов для инструментов обхода цензуры. В последние годы это переросло в соперничество между цензорами и разработчиками средств обхода. Это соперничество привело к разработке в Китае чрезвычайно эффективной методики идентификации протоколов, называемой исследователями активным зондированием (active probing) или активным сканированием (active scanning). При активном зондировании цензоры определяют, используется ли на хосте протокол обхода, пытаясь инициировать взаимодействие с использованием этого протокола. Если хост и цензор успешно согласуют соединение, цензор узнает, что на хосте применяется средство обхода. В Китае используется активное сканирование для эффективного блокирования Tor [Winter-2012].

Компромиссы. Идентификация протокола даёт лишь представление о способе передачи информации, но не о самой информации. Идентификация протокола полезна для обнаружения и блокировки средств обхода (таких как Tor) или трафика, который трудно проанализировать (например, VoIP или SSL), поскольку цензор может предположить, что этот трафик следует блокировать. Однако блоикировка может оказаться избыточной при использовании для популярных протоколов. Методы дороги в вычислительном и финансовом смысле из-за применения статистическго анализа, а также могут быть малоэффективны по причине низкой точности.

В прошлом цензоры использовали идентификацию протоколов для фильтрации по разрешительному списку (allowlist), например, разрешая использовать лишь конкретные, предварительно проверенные протоколы и блокируя любые нераспознанные протоколы [Bock-2020]. Такой подход к фильтрации протоколов может приводить к чрезмерной блокировке, если списки разрешённых протоколов слишком малы, поскольку многие стандартные «разрешенные» протоколы легко идентифицируются (например, HTTP).

Примеры из практики. Иденификацию протокола можно легко обнаружить, если она происходит в реальном масштабе времени и блокируется лишь определённый протокол. Однако некоторые типы идентификации протокола, такие как активное сканирование, могут быть более сложны для обнаружения. Идентификация протоколов использовалась Ираном для обнаружения и подавления трафика протокола SSH (Secure Shell), чтобы осложнить его использование [Van-der-Sar-2007], а также Китаем для идентификации и блокировки ретрансляторов Tor [Winter-2012]. Идентификация

протоколов применялась также для управления трафиком, например, в 2007 г., когда компания Comcast из США использовала внедрение пакетов TCP RST в потоки для прерывания трафика BitTorrent [Winter-2012]. В 2020 г. Иран развернул фильтр, который разрешал использовать лишь три протокола (DNS, TLS, HTTP) на конкретных портах, подвергая цензуре любые соединения, которые не идентифицированы [Bock-2020]. В 2022 г. Россия, возможно, использовала идентификацию протоколов для блокировки большинства соединений HTTP/3 [Elmenhorst-2022].

4.4. Остаточная цензура

Ещё одной особенностью некоторых современных систем цензуры является остаточная цензура - карательная форма цензуры, при которой после прерывания цензором запретного соединения сохраняется отслеживание последующих соединений, даже если они безвредны [Bock-2021]. Остаточная цензура может принимать разные формы и часто основана на методах технического вмешательства, описанных в следующем разделе. Важным аспектом остаточной цензуры является состав того, что цензор продолжает блокировать после её включения. Имеется три основных варианта - адреса IP у сервера и клиента (2-tuple), адреса сервера и клиента плюс порт сервера (3-tuple), адреса и номера портов у клиента и сервера (4-tuple). Будущие соединения, соответствующие отслеживаемому кортежу будут прерываться цензором [Bock-2021].

Остаточную цензуру порой сложно выявить и оценка её действенности может быть затруднена.

Компромиссы. Влияние остаточной цензуры заключается в том, что она ещё больше отбивает у пользователей желание обращаться к запретному содержимому, хотя неясно, насколько такая цензура успешна.

Примеры из практики. В Китае многие годы применялась остаточная цензура 3-tuple в сочетании с цензурой HTTP и исследователи сообщали об аналогичной остаточной цензуре для HTTPS. По всей видимости В Китае применяется сочетание остаточной цензуры 3-tuple и 4-tuple для цензуры HTTPS с использованием ESNI. Некоторые цензоры, (включая Казахстан и Иран), выполняющие цензуру через отбрасывание пакетов, зачастую случайно применяют остаточную цензуру 4-tuple [Bock-2021].

5. Техническое вмешательство

5.1. Прикладной уровень

5.1.1. Вмешательство в DNS

Имеется ряд механизмов, которые цензоры могут применять для блокировки или фильтрации доступа к содержимому путём изменения откликов DNS [AFNIC-2013] [ICANN-SSAC-2012], включая блокировку откликов, в также возврат сообщений об ошибках или некорректных адресов. Отметим, что в настоящее время существует защищённая доставка запросов DNS через HTTPS [RFC8484] и TLS [RFC7858], которая может смягчить помехи для запросов DNS между «заглушкой» (stub) и распознавателем.

Отклики на запросы DNS с возвратом некорректного адреса можно реализовать путём перехвата на пути, отравления кэшей вне пути или выдачи ложных откликов на стороне сервера.

DNS mangling (мскажение) - это метод перехвата на сетевом уровне, где возвращается некорректный адрес IP в отклике на запрос DNS для запретного адресата. Так поступают, например, в некоторых китайских сетях (авторам неизвестно о других столь же масштабных примерах), где каждый передаваемый запрос DNS проверяется (предположительно с помощью таких технологий, как DPI) и при соответствии списку цензуры на него возвращается ложный отклик. Конечные пользователи могут увидеть действие этого метода, просто передав запрос DNS для любого адреса, не используемого в Китае (см. пример ниже). Если доменное имя не подвергается цензуре, отклика просто не будет, в ином случае возвращается обманный отклик. Ниже приведён пример запросов с помощью утилиты dig у сервера с неиспользуемым в Китае адресом IP 192.0.2.2[1] для имен www.uncensored.example (нет отклика) и www.censored.example (подвергался цензуре на момент написания, получен обманный адрес IP 198.51.100.0).

```
% dig +short +nodnssec @192.0.2.2 A www.uncensored.example
;; connection timed out; no servers could be reached

% dig +short +nodnssec @192.0.2.2 A www.censored.example
198.51.100.0
```

Отравление кэшей DNS происходит вне пути и является механизмом вмешательства цензора в отклики полномочных серверов имён DNS рекурсивным распознавателям за счёт более быстрой отправки откликов (с ложными адресами IP) [Halley-2008]. Отравление кэша происходит после того, как серверы имён запрашиваемого сайта распознают запрос и пытаются передать запрашивающему устройству отклик с корректным адресом IP. На пути возврата отклика распознанный адрес IP рекурсивно кэшируется каждым сервером DNS, который ранее пересылал запрос. В процессе кэширования при распознавании нежелательного ключевого слова распознанный адрес IP «отравляется» и возвращается другой адрес IP (или ошибка NXDOMAIN) быстрее, чем мог ответить исходный распознаватель, что ведёт к кэшированию ложено адреса IP (возможно, рекурсивно). Ложные адреса IP обычно направляют в несуществующий домен или на страницу с предупреждением. Как вариант, иранская цензура препятствует взаимодействию, не позволяя передавать отклик [Aryan-2013].

Известны также случаи DNS-обмана (DNS lying), когда цензор требует предоставлять (оператором рекурсивного распознавателя или ISP) отклики DNS, отличающиеся от тех, которые даёт полномочный сервер [Bortzmeyer-2015].

Компромиссы. Эти формы вмешательства в DNS требуют от цензора вынуждать пользователей проходить через контролируемую иерархию DNS (или промежуточную сеть, где цензор является активным всемогущим «атакующим» [RFC7624] для переписывания откликов DNS), чтобы механизм мог работать. Вмешательство в DNS можно обойти, используя другие распознаватели DNS (например, общедоступные DNS), которые могут находиться вне сферы контроля цензора, или технологию виртуальных частных сетей (Virtual Private Network или VPN). Искажение DNS и отравление кэша предполагают возврат некорректных адресов IP при попытках распознавания доменных имён, но в некоторых случаях адресат может быть технически доступен. Например, для протокола HTTP у пользователя может быть другой способ получения IP-адреса нужного сайта и пользователь сможет получить доступ к нему, если сайт

[1] Этот адрес из блока, зарезервированного для документации, см. RFC 6890. *Прим. перев.*

настроен на использование по умолчанию для данного IP. Проблема возникает и при блокировке по целям, поскольку иногда пользователи, находящиеся за пределами региона цензуры, будут направляться через серверы DNS или переписывающее DNS оборудование, контролируемые цензором, что вызовет отказы при выполнении запросов. Простота обхода в сочетании с большим риском блокировки содержимого и блокировки по целям делает вмешательство в DNS неполным, сложным и не самым идеальным механизмом цензуры.

Кроме того, описанные выше механизмы предполагают отсутствие DNSSEC или отключенную проверку DNSSEC на стороне клиента или рекурсивного распознавателя (это вполне возможно с учётом ограниченного внедрения DNSSEC и проверки DNSSEC на стороне клиентов). Отметим, что при попытке блокировать распознавание могут предоставляться записи DNSSEC, которые не пройдут проверку, если клиент или рекурсивный распознаватель выполняет ее.

Ранее для цензуры применялись методы, основанные на передаче запросов DNS в открытом виде (cleartext) через порт 53 [SSAC-109-2020]. С внедрение ширования в DNS (например, DNS через HTTPS [RFC8484]) запросы все чаще передаются через порт 443 вместе с другим трафиком HTTPS или в зашифрованном виде при работе DNS через TLS [RFC7858] (см. параграф 4.3.1).

Примеры из практики. Вмешательство в DNS при подобающей его реализации легко обнаружить по отмеченным выше недостаткам. В марте 2014 г. Турция в течение почти недели полагалась на вмешательство в DNS для блокировки web-сайтов, включая Twitter и YouTube. Простота обхода блокировки привела к росту популярности Twitter, пока турецкие ISP не ввели блокировку по IP для выполнения требований правительства [Zmijewski-2014]. В итоге турецкие ISP стали перехватывать все запросы к международным распознавателям DNS компаний Google и Level 3 [Zmijewski-2014]. Некорректная реализация вмешательства в DNS привела к крупным бедствиям для цензуры. В январе 2014 г. Китай начал направлять все запросы, проходящие через GFW, в домен dongtaiwang.com из-за неподобающей настройки отравления DNS. Этот инцидент считается крупнейшим в истории отказом служб Internet [AFP-2014] [Anon-SIGCOMM12]. В таких странах, как Китай, Турция и США обсуждался вопрос о бликировке целиком доменов верхнего уровня (Top-Level Domain или TLD) [Albert-2011]. Блокировка DNS часто применяется в европейских странах для оборбы с нежелательным содержимым.

- Насилие (abuse) над детьми (Норвегия, Великобритания, Бельгия, Дания, Финляндия, Франция, Германия, Ирландия, Италия, Мальта, Нидерланды, Польша, Испания, Швеция [Wright-2013] [Eneman-2010]).
- Азартные игры (Бельгия, Болгария, Чехия, Кипр, Дания, Эстония, Франция, Греция, Венгрия, Италия, Латвия, Литва, Польша, Португалия, Румыния, Словакия, Словения, Испания; параграф 6.3.2 в [EC-gambling-2012], [EC-gambling-2019]).
- Нарушаюения авторских прав (Все страны Европейской экономической зоны).
- Разжигание ненависти и экстремизм (Франция [Hertel-2015]).
- Терроризм (France [Hertel-2015]).

5.2. Транспортный уровень

5.2.1. Снижение производительности

Хотя некоторые из описанных в этом разделе методов направлены в основном на блоткрование или предотвращение доступа к содержимому, в некоторых случаях эеевективной стратегией цензуры может быть не полное блокирование доступа к данному адресату или службе, а снижение производительности соответствующих сетевых соединений. В результате восприятие сайта пользователем становится настолько негативным, что тот предпочтет обратиться к другому сайту, службе или способу связи, а при отсутствии их - совсем отказаться от работы с таким сайтом. Одним из способов снижения производительности является формовка (shaping) трафика, ограничивающая пропускную способность для отдельных типов трафика.

Компромиссы. Хотя снижение производительности не всегда устранает возможность доступа к соответствующему ресурсу, это может вынудить пользователей применять иные средства коммуникаций, где цензуру (или слежку) организовать проще.

Примеры из практики. Известно, что Иран сокращает пропускную способность для трафика HTTPS, чтобы шире применялся нешифрованный трафик HTTP [Aryan-2013].

5.2.2. Отбрасывание пакетов

Отбрасывание пакетов - это простой механизм предотвращения нежелательного трафика. Цензов идентифицирует нежелательный трафик и решает не пересылать должным образом никакие пакеты, связанные с доставкой такого трафика, как этого требует обычная маршрутизация. Это может сочетаться с любым из описанных выше механизмов, если цензору известно, что пользователь должен направить трафик через контролируемый маршрутизатор.

Компромиссы. Отбрасывание пакетов наиболее успешно в случаях, когда каждый пакет имеет доступные сведения, связанные с нежелательным содержимым (например, IP-адрес получателя). Одним из недостатков метода является необходимость блокирования всего содержимого, включая разрешённый трафик с тем же адресом IP (например. для для всех блогов или репозиториев GitHub на одном сервере). Известно, что Китай в течение 3 дней отбрасывал все пакеты GitHub из-за одного репозитория с нежелательным содержимым [Anonymous-2013]. Необходимость проверки каждого пакета в близком к реальному масштабе времени делает отбрасывание пакетов проблематичным решением с точки зрения QoS.

Примеры из практики. Отбрасывание пакетов является очень распространённой формой технического вмешательства и поддаётся точному обнаружению из-за оставляемых уникальных следов в тайм-аутах запросов. Замечено, что Great Firewall в Китае использует отбрасывание пакетов в качестве одного из основных технических механизмов цензуры [Ensafi-2013]. Иран применяет отбрасывание пакетов в качестве механизма поддавления SSH [Aryan-2013]. Это лишь два примера из практики повсеместной цензуры. Примечательно, что отбрасывание пакетов в процессе согласования или при работе соединений - это единственный метод технического вмешательства, отмеченный к настоящему времени для протокола QUIC (например, в Индии, Иране, России и Уганде [Elmenhorst-2021] [Elmenhorst-2022]).

5.2.3. Внедрение пакетов RST

Внедрением пакетов обычно называют метод нарушения работы сети с участием машины на пути (machine-in-the-middle или MITM), которая подменят пакеты в существующем потоке трафика. Пакеты RST обычно служат для информирования одной из сторон соединения TCP о прекращении передачи другой стороной и рекомендации закрыть соединение. Внедрение пакетов RST - это особый тип атак с внедрением, используемый для прерывания существующего потока путём отправки пакетов RST обеим сторонам соединения TCP. Каждый получатель считает, что соединение прервано другой стоной и сессия прерывается.

Протокол QUIC после организации соединения не чувствителен к таким атакам с внедрением пакетов. Хотя в QUIC реализован механизм сброса без учёта состояния, каждая из партнёров воспринимает такой сброс лишь в том случае, когда пакет завершается выданным ранее маркером (stateless reset), который трудно угадать. В процессе согласования QUIC обеспечивает эффективную защиту лишь от злоумышленников вне пути, но уязвим к атакам с внедрением пакетов со стороны злоумышленников, проанализировавших предшествующие пакеты (см. подробности в [RFC9000]).

Компромиссы. Несмотря на неэффективность для протоколов, отличных от TCP (QUIC, IPsec), внедрение пакетов RST имеет ряд преимуществ, которые делают этот метод чрезвычайно популярным для цензуры. Внедрение пакетов RST является вмешательством, позволяющим избежать узких мест QoS, с которыми можно столкнуться при использовании таких методов, как отбрасывание пакетов. Это свойство «внеполосности» (out-of-band) позволяет цензору просматривать копию данных, обычно отображаемую оптическим ответвителем, что делает метод идеальным в сочетании с DPI и идентификацией протоколов [Weaver-2009]. Такой асинхронный вариант часто называют «машиной в стороне» (machine-on-the-side или MOTS). Преимуществом внедрения пакетов RST является также то, что для прерывания соединения достаточно восприятия внедренного пакета лишь одной из двух конечных точек.

Сложность внедрения пакетов RST заключается в подделке «достаточного» объёма данных, чтобы конечная точка сочла пакет RST легитимным, - обычно это включает корректный адрес IP, порт и порядковый номер TCP. Порядковый номер подделать сложней всего, поскольку [RFC9293] указывает, что для восприятия пакета RST ему следует быть упорядоченным, хотя в том же RFC рекомендуется воспринимать пакеты из окна (in-window). Эта рекомендация важна и её выполнение позволяет организовать атаку с внедрением RST вслепую (Blind RST Injection) [Netsec-2011]. Когда номера из окна разрешены, организация вставки RST вслепую становится тривиальной задачей. Хотя термин «внедрение вслепую» предполагает, что у цензора нет никаких конфиденциальных (sensitive) сведений о порядковых номерах потока TCP, в который он внедряется, можно просто перебрать все ~70000 возможных окон. Это особенно полезно для прерывания шифрованных и запутанных (obfuscated) протоколов, таких как SSH и Tor [Gilad]. Некоторые системы обхода цензуры пытаются запутать цензора, заставляя того отслеживать некорректную информацию, что делает внедрение пакетов RST бесполезным [Khattak-2013] [Wang-2017] [Li-2017] [Bock-2019] [Wang-2020].

Внедрение пакетов RST предназначено для сетей с поддержкой состояния, что делает этот метод бесполезным для соединения UDP. Внедрение пакетов RST является одним из наиболее популярных методов цензуры, применяемых в настоящее время, благодаря его универсальности и эффективности для всех типов трафика TCP. Недавние исследования показали, что атаки с внедрением пакетов TCP RST возможны даже при размещении атакующего вне пути [Cao-2016].

Примеры из практики. Как отмечено выше, внедрение пакето RST чаще всего применяется в комбинации с методами, требующими расщепления (отвода) трафика, такими как DPI и идентификация протоколов. В 2007 г. компанию Comcast обвинили в применении внедрения пакетов RST для прерывания трафика, идентифицированного как BitTorrent [Schoen-2007], что привело к постановлению Федеральной комиссии США по связи (US Federal Communications Commission) против Comcast [VonLohmann-2008]. Известно также, что Китай часто применяет внедрение пакетов RST для цензуры. Особенно наглядно это проявляется в прерывании работы шифрованных и запутанных (obfuscated) протоколов, вроде применяемых в Tor [Winter-2012].

5.3. Уровень маршрутизации

5.3.1. Изоляция сетей

Хотя это, пожалуй, самый грубый метод цензуры, нет более эффективного способа предотвратить распространения нежелательных сведений в web, чем отключение сети. Сеть в том или ином регионе можно логически изолировать путём отзыва цензором всех префиксов протокола граничного шлюза (Border Gateway Protocol или BGP), проходящих через страну цензора.

Компромиссы. Последствия изоляции сети в регионе огромны и несомненны, цензор расплачивается за абсолютный контроль над цифровой информацией утратой преимуществ доступа в глобальную сеть Internet. Изоляция сетей дорого обходится и в политическом смысле, поскольку граждане, привыкшие к использованию платформ и услуг Internet воспринимают такое отключение как потерю своей гражданской свободы. Изоляция сети редко бывает длительной и обычно применяется лишь как крайняя мера в период серьёзных гражданских волнений в стране.

Примеры из практики. Отключение сетей, как правило, происходит лишь во время серьёзных беспорядков, что обусловлено огромными социальными, политическими и экономическими последствиями такого шага. Одним из первых, получивших широкое освещение, случаев стало отключение сети хунтой Мьянмы при подавлении восстания в 2007 г. [Dobie-2007]. Китай отключал сеть в районе Синьцзян во время беспорядков 2009 г., чтобы предотвратить распространение протестов в другие регионы [Heacock-2009]. Наиболее часто отключение сетей применялось во время Арабской весны в Египте и Ливии в 2011 г. [Cowie-2011], в Сирии в 2012 г. [Thomson-2012]. Россия заявила о попытке отключения своих сетей от Internet в апреле 2019 г. в рамках проверки сетевой независимости страны. Сообщалось, что в рамках тестового отключения российские телекоммуникационные компании должны маршрутизировать весь трафик на государственные точки мониторинга [Cimpanu-2019]. Наибольшее число отключений сетей наблюдалось в Индии в 2016 и 2017 гг. [Dada-2017].

5.3.2. Анонсирование обманных маршрутов

Более тонко настраиваемая и широкомасштабная цензура может быть реализована путём захвата BGP (hijacking), где префиксы IP преднамеренно маршрутизируется некорректно в пределах региона и вне его с помощью BGP. Это ограничивает и эффективно цензурирует корректно известные местоположения информации, поступающей в юрисдикцию или из неё, а также предотвращает людям, находящимся вне юрисдикции, просматривать содержимое,

созданное в этой юрисдикции, по мере распространения искажённых маршрутов. Первое может быть достигнуто с помощью анонсирования через BGP некорректных маршрутов, которые не предназначены для выхода за пределы юрисдикции, а второе - путём преднамеренного внедрения обманных маршрутов, распространяемых в Internet.

Компромиссы. Глобальное распространение ложного маршрута к web-сайту может перегрузить ISP, если сайт был популярным. Это не является постоянным решением, поскольку некорректные маршруты BGP с глобальным распространением могут быть исправлены, но маршруты с локальным влиянием (внутри юрисдикции) могут быть исправлены ISP/IXP лишь для локальных пользователей.

Примеры из практики. В 2008 г. компания Pakistan Telecom по требованию правительства Пакистана ввела цензуру для YouTube путём изменения маршрутов BGP для этого сайта. Новые маршруты анонсировались восходящим ISP и не только. В результате вся сеть Internet стала направлять маршруты для YouTube на Pakistan Telecom и это продолжалось в течение многих часов. В 2018 г. почти все службы Google и клиенты Google Cloud (например, Spotify), не могли работать более часа после того, как компания Google потеряла контроль над несколькими миллионами своих адресов IP. Эти префиксы IP некорректно маршрутизировались в China Telecom (государственный китайский ISP) [Google-2018], аналогично захвату BGP для правительственных и военных web-сайтов США компанией China Telecom в 2010 г. ISP в России (2022) и Мьянме (2021) неоднократно пытались захватить один префикс Twitter [Siddiqui-2022].

5.4. Воздействие на других уровнях

5.4.1. Распределенный отказ в обслуживании (DDoS)

Распределенные атаки для отказа в обслуживании (Distributed Denial of Service или DDoS) являются основным механизмом, применяемым «хактивистами» и злонамеренными хакерами. Цензоры в прошлом также применяли DDoS по разным причинам. Существует большое разнообразие атак DDoS [Wikip-DoS]. Однако на верхнем уровне возникает два варианта - «наводнение» (flood) пакетами ведёт к непригодности службы для использования, а атаки с авариями (crash) нацелены на отказ службы, чтобы ресурсы можно было выделить в другом месте без «освобождения» службы.

Компромиссы. DDoS является привлекательным механизмом в случаях, когда цензор хочет на ограниченное время полностью (а не только локально) прекратить доступ к нежеланному содеримому. Временный характер является, пожалуй, единственным уникальным преимуществом DDoS как метода цензуры. Ресурсы, требуемые для организации успешной DDoS-атаки против крупной цели, требуют больших вычислительных затрат, для чего обычно нужно владеть или арендовать распределенную вредоносную платформу, такую как сеть ботов (botnet), и оценка требуемых ресурсов неточна. DDoS является очень грубым методом цензуры и, судя по всему, применяется в основном как быстрый и легкодоступный механизм блокирования нежелательного содержимого в течение ограниченного срока.

Примеры из практики. В 2012 г. британская организация по разведке сигналов, Штаб правительственной связи (Government Communications Headquarters или GCHQ) использовала DDoS для временной остановки чатов IRC (Internet Relay Chat), посещаемых членами группы Anonymous, с использованием метода Syn Flood DDoS. Этот метод использует согласование TCP для перегрузки сервера-жертвы таким числом запросов, что легитимный трафик сильно замедляется или прекращается [NBC-2014] [CERT-2000]. Сайты-диссиденты часто становятся жертвами DDoS-атак во время политически важных событий, например, DDoS в Бирме [Villeneuve-2011]. Правящие партии в России [Kravtsova-2012], Зимбабве [Orion-2013] и Малайзии [Muncaster-2013] обвинялись в применении DDoS для предотвращения поддержки и доступа к сайтам оппозиции во время выборов. В 2015 г. в Китае была организована DDoS-атака с использованием системы MITM (названной «Великая пушка» - Great Cannon), размещённой на Great Firewall и позволяющей внедрять код JavaScript при посещении китайской поисковой системы, что вело к передаче пользовательскими агентами трафика DDoS на различные сайты [Marczak-2015].

5.4.2. Глубокая цензура

Зачастую цензоры применяют сразу несколько методов, организуя глубокую цензуру (censorship in depth). Такая цензура может иметь разные формы - некоторые цензоры блокируют одно и то же содержимое несколькими методами (например, блокирование домена в DNS, блокирование по IP и HTTP), другие организуют распараллеливание для повышения надёжности (например, применение нескольких разных систем цензуры для блокировки одного и того же домена), третьи могут применять дополнительные системы для снижения возможностей обхода (например, полная блокировка нежелательных протоколов, вынуждающая пользователей переходить на другие протоколы).

Компромиссы. Глубокая цензура может быть привлекательной для цензоров, поскольку она даёт дополнительные гарантии и при обходе одного метода может сработать другой. Основным недостатком такого подхода является стоимость внедрения, поскольку требуется реализация нескольких систем цензуры, работающих в тандеме.

Примеры из практики. Глубокая цензура сегодня применяется многими крупными государствами. Исследователи отмечают, что Китай развернул крупную систему глубокой цензуры, часто подвергая цензуре по нескольким протоколам один и тот же ресурс [Chai-2019] [Bock-2020b] или реализуя несколько систем цензуры для одного содержимого и протокола [Bock-2021b]. В Иране внедрён дополняющий фильтр протоколов для ограничения использования протоколов на некоторых портах, вынуждающий пользователей применять протоколы, которые система цензуры может фильтровать [Bock-2020].

6. Иные виды вмешательства

6.1. Ручная фильтрация

Иногда ручная настройка является простейшим способом блокировки содержимого. Ручная фильтрация отличается от обычной тактики создания списков блокировки тем, что она не обязательно направлена на конкретный IP или DNS, а удаляет или помечает содержимое. С учётом неточности автоматической фильтрации ручная сортировка содержимого и маркировка нежелательных web-сайтов, блогов и других сред может быть эффективным средством как сама по себе, так и в сочетании с автоматическими методами обнаружения, за которыми следуют действия, требующие подтверждения вручную. Такая фильтрация может выполняться в магистрали или у ISP. Хорошим примером является китайская армия мониторов [BBC-2013b], но чаще ручная фильтрация происходит на институциональном уровне. Для ICP, таких как Google и Weibo, требуется получение лицензии на работу в Китае. Одним из предварительных условий для получения такой лицензии является согласие подписать «добровольное обязательство», известное как Public

Pledge on Self-discipline for the Chinese Internet Industry. Неспособность «энергично поддерживать» заявленные возможности может приводить к ответственности ICP за недопустимое (offending) содержимое со стороны правительства Китая [BBC-2013b].

6.2. Самоцензура

Самоцензуру трудно документировать, поскольку она проявляется, прежде всего, в отсутствии нежелательного содержимого. Средства, способствующие самоцензуре, могут потребовать от предполагаемого автора поверить, что выступление повышает для него риск неблагоприятных последствий (технический мониторинг, требования к идентификации и т. п.). Методы навязывания самоцензуры «Репортеров без границ» (Reporters Without Borders) приводятся в ежегодных отчётах World Press Freedom Index [RWB-2020].

6.3. Изъятие сервера

Как уже упоминалось в [Murdoch-2008], серверы должны иметь физическое местоположение где-то в мире. Если нежелательное содержимое размещено в стране, где действует цензура, серверы могут быть изъяты физически или от хостинг-провайдера могут потребовать запрета доступа (если сервер виртуальный и размещён в облачной инфраструктуре, где может не быть фиксированного местоположения.

6.4. Уведомления и удаление содержимого

Во многих странах имеются механизмы, позволяющие частному лицу или иному поставщику содержимого направить держателю хостинга юридический запрос на удаление содержимого. Примеры включают системы, развёрнутые такими компаниями, как Google, для соблюдения «права быть забытым» (Right to be Forgotten) в Европейском союзе [Google-RTBF], правила ответственности для провайдеров электронных платформ [EC-2012], ориентированные на авторские права уведомления и удаление содержимого, предусмотренные разделом 512 Закона США об авторских правах в цифровую эпоху (United States Digital Millennium Copyright Act или DMCA) [DMLP-512].

6.5. Изъятие доменного имени

Доменные имена каталогизируются на серверах имён, управляемых юридическими лицами (регистраторами). Эти реестры можно вынудить передать управление доменом от зарегистрировавшего домен лица кому-либо иному с помощью правовой процедуры, основанной на частных договорах или законе. Изъятие доменного имени все чаще применяется государственными органами и частными организациями для борьбы с распространением нежелательного содержимого [ICANN-2012] [EFF-2017].

7. Продолжение работы

Помимо создания обстоятельного ресурса, описывающего методы цензуры, этот документ определяет важные направления будущей работы.

В целом видимые затраты на реализацию методов цензуры указывают необходимость более точной классификации режимов цензуры по мере их становления и развития, а также более точных описаний методов обхода цензуры. Под зрелостью цензора понимается техническая зрелость, требуемая для применения конкретных методов. В будущем можно будет классифицировать методы цензуры по интенсивности работы цензоров, включая требуемую инфраструктуру, для успешной цензуры содержимого, пользователей или услуг.

В части обхода цензуры рост числа протоколов, применяющих шифрование, является эффективной мерой противодействия некоторым формам цензуры, описанным здесь, однако подробное исследование обхода и шифрования оставлено для другого документа. Кроме того, сообщество разработчиков средств обхода цензуры развивает направление по исследованию подключаемого (pluggable) транспорта, в рамках которого собираются, документируются и совершенствуются методы запутывания (obfuscating) трафика в пути, чтобы сделать его для цензуры неотличимым от других видов трафика [Tor-2019]. Для этих методов будет полезна работа сообщества разработчиков стандартов Internet.

Эмпирические примеры показывают, что методы цензуры могут быстро развиваться, а опыт говорит, что этот документ может быть актуален лишь в течение короткого времени. В будущем документ может быть дополнен новыми методами, описанными с использованием сравнимой методологии.

8. Взаимодействие с IANA

Этот документ не требует действий IANA.

9. Вопросы безопасности

Этот документ является обзором имеющейся литературы по методам цензуры в сети, поэтому он не касается каких-либо вопросов безопасности, помимо тех, которые уже рассмотрены в упомянутых здесь работах.

10. Литература

[AFNIC-2013] AFNIC, "Report of the AFNIC Scientific Council: Consequences of DNS-based Internet filtering", January 2013, <http://www.afnic.fr/medias/documents/conseilscientifique/SC-consequences-of-DNS-based-Internet-filtering.pdf>.

[AFP-2014] AFP, "China Has Massive Internet Breakdown Reportedly Caused By Their Own Censoring Tools", January 2014, <http://www.businessinsider.com/chinas-internet-breakdown-reportedly-caused-by-censoring-tools-2014-1>.

[Albert-2011] Albert, K., "DNS Tampering and the new ICANN gTLD Rules", June 2011, <https://opennet.net/blog/2011/06/dns-tampering-and-new-icann-gtld-rules>.

[Anon-SIGCOMM12] Anonymous, "The Collateral Damage of Internet Censorship by DNS Injection", July 2012, <http://www.sigcomm.org/sites/default/files/ccr/papers/2012/July/2317307-2317311.pdf>.

[Anonymous-2013] Anonymous, "GitHub blocked in China - how it happened, how to get around it, and where it will take us", January 2013, <https://en.greatfire.org/blog/2013/jan/github-blocked-china-how-it-happened-how-get-around-it-and-where-it-will-take-us>.

[Anonymous-2014] Anonymous, "Towards a Comprehensive Picture of the Great Firewall's DNS Censorship", August 2014, <https://www.usenix.org/system/files/conference/foci14/foci14-anonymous.pdf>.

[Aryan-2013] Aryan, S., Aryan, H., and J. A. Halderman, "Internet Censorship in Iran: A First Look", 2012, <https://jhalderm.com/pub/papers/iran-foci13.pdf>.

[BBC-2013] BBC News, "Google and Microsoft agree steps to block abuse images", November 2013, <http://www.bbc.com/news/uk-24980765>.

[BBC-2013b] BBC, "China employs two million microblog monitors state media say", 2013, <https://www.bbc.com/news/world-asia-china-24396957>.

[Bock-2019] Bock, K., Hughey, G., Qiang, X., and D. Levin, "Geneva: Evolving Censorship Evasion Strategies", DOI 10.1145/3319535.3363189, November 2019, <https://geneva.cs.umd.edu/papers/geneva_ccs19.pdf>.

[Bock-2020] Bock, K., Fax, Y., Reese, K., Singh, J., and D. Levin, "Detecting and Evading Censorship-in-Depth: A Case Study of Iran's Protocol Filter", January 2020, <https://geneva.cs.umd.edu/papers/evading-censorship-in-depth.pdf>.

[Bock-2020b] Bock, K., iyouport, Anonymous, Merino, L-H., Fifield, D., Houmansadr, A., and D. Levin, "Exposing and Circumventing China's Censorship of ESNI", August 2020, <https://geneva.cs.umd.edu/posts/china-censors-esni/esni/>.

[Bock-2021] Bock, K., Bharadwaj, P., Singh, J., and D. Levin, "Your Censor is My Censor: Weaponizing Censorship Infrastructure for Availability Attacks", DOI 10.1109/SPW53761.2021.00059, May 2021, <https://geneva.cs.umd.edu/papers/woot21-weaponizing-availability.pdf>.

[Bock-2021b] Bock, K., Naval, G., Reese, K., and D. Levin, "Even Censors Have a Backup: Examining China's Double HTTPS Censorship Middleboxes", FOCI '21: Proceedings of the ACM SIGCOMM 2021 Workshop on Free and Open Communications on the Internet, Pages 1-7, DOI 10.1145/3473604.3474559, August 2021, <https://geneva.cs.umd.edu/papers/foci21.pdf>.

[Bortzmeyer-2015] Bortzmeyer, S., "DNS Censorship (DNS Lies) As Seen By RIPE Atlas", December 2015, <https://labs.ripe.net/Members/stephane_bortzmeyer/dns-censorship-dns-lies-seen-by-atlas-probes>.

[Boyle-1997] Boyle, J., "Foucault in Cyberspace: Surveillance, Sovereignty, and Hardwired Censors", 66 University of Cincinnati Law Review 177-205, 1997, <https://scholarship.law.duke.edu/faculty_scholarship/619/>.

[Cao-2016] Cao, Y., Qian, Z., Wang, Z., Dao, T., Krishnamurthy, S., and L. Marvel, "Off-Path TCP Exploits: Global Rate Limit Considered Dangerous", August 2016, <https://www.usenix.org/system/files/conference/usenixsecurity16/sec16_paper_cao.pdf>.

[CERT-2000] CERT, "CERT Advisory CA-1996-21 TCP SYN Flooding and IP Spoofing Attacks", 2000, <https://vuls.cert.org/confluence/display/historical/CERT+Advisory+CA-1996-21+TCP+SYN+Flooding+and+IP+Spoofing+Attacks>.

[Chai-2019] Chai, Z., Ghafari, A., and A. Houmansadr, "On the Importance of Encrypted-SNI (ESNI) to Censorship Circumvention", 2019, <https://www.usenix.org/system/files/foci19-paper_chai_update.pdf>.

[Cheng-2010] Cheng, J., "Google stops Hong Kong auto-redirect as China plays hardball", June 2010, <http://arstechnica.com/tech-policy/2010/06/google-tweaks-china-to-hong-kong-redirect-same-results/>.

[Cimpanu-2019] Cimpanu, C., "Russia to disconnect from the internet as part of a planned test", February 2019, <https://www.zdnet.com/article/russia-to-disconnect-from-the-internet-as-part-of-a-planned-test/>.

[CitizenLab-2018] Marczak, B., Dalek, J., McKune, S., Senft, A., Scott-Railton, J., and R. Deibert, "Bad Traffic: Sandvine's PacketLogic Devices Used to Deploy Government Spyware in Turkey and Redirect Egyptian Users to Affiliate Ads?", March 2018, <https://citizenlab.ca/2018/03/bad-traffic-sandvines-packetlogic-devices-deploy-government-spyware-turkey-syria/>.

[Clayton-2006] Clayton, R., Murdoch, S.J., and R.N.M. Watson, "Ignoring the Great Firewall of China", Lecture Notes in Computer Science, Volume 4258, DOI 10.1007/11957454_2, 2006, <https://link.springer.com/chapter/10.1007/11957454_2>.

[Condliffe-2013] Condliffe, J., "Google Announces Massive New Restrictions on Child Abuse Search Terms", November 2013, <http://gizmodo.com/google-announces-massive-new-restrictions-on-child-abus-1466539163>.

[Cowie-2011] Cowie, J., "Egypt Leaves The Internet", NANOG 51, February 2011, <https://archive.nanog.org/meetings/nanog51/presentations/Tuesday/LT-Cowie-Egypt%20Leaves%20The%20Internet.pdf>.

[Crandall-2010] Park, J.C. and J. Crandall, "Empirical Study of a National-Scale Distributed Intrusion Detection System: Backbone-Level Filtering of HTML Responses in China", June 2010, <http://www.cs.unm.edu/~crandall/icdcs2010.pdf>.

[Dada-2017] Dada, T. and P. Micek, "Launching STOP: the #KeepItOn internet shutdown tracker", September 2017, <https://www.accessnow.org/keepiton-shutdown-tracker/>.

[Dalek-2013] Dalek, J., Haselton, B., Noman, H., Senft, A., Crete-Nishihata, M., Gill, P., and R. J. Deibert, "A Method for Identifying and Confirming the Use of URL Filtering Products for Censorship", IMC '13: Proceedings of

the 2013 conference on Internet measurement conference, Pages 23-30, DOI 10.1145/2504730.2504763, October 2013, <http://conferences.sigcomm.org/imc/2013/papers/imc112s-dalekA.pdf>.

[Ding-1999] Ding, C., Chi, C. H., Deng, J., and C. L. Dong, "Centralized Content-Based Web Filtering and Blocking: How Far Can It Go?", IEEE SMC'99 Conference Proceedings, DOI 10.1109/ICSMC.1999.825218, October 1999, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.132.3302&rep=rep1&type=pdf>.

[DMLP-512] Digital Media Law Project, "Protecting Yourself Against Copyright Claims Based on User Content", May 2012, <https://www.dmlp.org/legal-guide/protecting-yourself-against-copyright-claims-based-user-content>.

[Dobie-2007] Dobie, M., "Junta tightens media screw", BBC News, September 2007, <http://news.bbc.co.uk/2/hi/asia-pacific/7016238.stm>.

[EC-2012] European Commission, "Summary of the results of the Public Consultation on the future of electronic commerce in the Internal Market and the implementation of the Directive on electronic commerce (2000/31/EC)", January 2012, <https://ec.europa.eu/information_society/newsroom/image/document/2017-4/consultation_summary_report_en_2010_42070.pdf>.

[EC-gambling-2012] European Commission, "Online gambling in the Internal Market Accompanying the document Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions Towards a comprehensive framework for online gambling", 2012, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52012SC0345>.

[EC-gambling-2019] European Commission, "Evaluation of regulatory tools for enforcing online gambling rules and channelling demand towards controlled offers", January 2019, <https://ec.europa.eu/growth/content/evaluation-regulatory-tools-enforcing-online-gambling-rules-and-channelling-demand-towards-1_en>.

[EFF-2017] Malcom, J., Rossi, G., and M. Stoltz, "Which Internet registries offer the best protection for domain owners?", Electronic Frontier Foundation, July 2017, <https://www.eff.org/files/2017/08/02/domain_registry_whitepaper.pdf>.

[ekr-2021] Rescorla, E., "Overview of Apple's Client-side CSAM Scanning", August 2021, <https://educatedguesswork.org/posts/apple-csam-intro/>.

[Elmenhorst-2021] Elmenhorst, K., Schuetz, B., Aschenbruck, N., and S. Basso, "Web Censorship Measurements of HTTP/3 over QUIC", IMC '21: Proceedings of the 21st ACM Internet Measurement Conference, Pages 276-282, DOI 10.1145/3487552.3487836, November 2021, <https://dl.acm.org/doi/pdf/10.1145/3487552.3487836>.

[Elmenhorst-2022] Elmenhorst, K., "A Quick Look at QUIC Censorship", April 2022, <https://www.opentech.fund/news/a-quick-look-at-quic/>.

[Eneman-2010] Eneman, M., "Internet service provider (ISP) filtering of child-abusive material: A critical reflection of its effectiveness", DOI 10.1080/13552601003760014, June 2010, <https://www.tandfonline.com/doi/abs/10.1080/13552601003760014>.

[Ensafi-2013] Ensafi, R., Knockel, J., Alexander, G., and J.R. Crandall, "Detecting Intentional Packet Drops on the Internet via TCP/IP Side Channels: Extended Version", DOI 10.48550/arXiv.1312.5739, December 2013, <http://arxiv.org/pdf/1312.5739v1.pdf>.

[Fifield-2015] Fifield, D., Lan, C., Hynes, R., Wegmann, P., and V. Paxson, "Blocking-resistant communication through domain fronting", DOI 10.1515/popets-2015-0009, May 2015, <https://petsymposium.org/2015/papers/03_Fifield.pdf>.

[Gatlan-2019] Gatlan, S., "South Korea is Censoring the Internet by Snooping on SNI Traffic", February 2019, <https://www.bleepingcomputer.com/news/security/south-korea-is-censoring-the-internet-by-snooping-on-sni-traffic/>.

[Gilad] Gilad, Y. and A. Herzberg, "Off-Path TCP Injection Attacks", ACM Transactions on Information and System Security, Volume 16, Issue 4, Article No.: 13, pp. 1-32, DOI 10.1145/2597173, April 2014, <https://doi.org/10.1145/2597173>.

[Glanville-2008] Glanville, J., "The big business of net censorship", The Guardian, November 2008, <http://www.theguardian.com/commentisfree/2008/nov/17/censorship-internet>.

[Google-2018] "Google Cloud Networking Incident #18018", November 2018, <https://status.cloud.google.com/incident/cloud-networking/18018>.

[Google-RTBF] Google, Inc., "Search removal request under data protection law in Europe", 2015, <https://support.google.com/legal/contact/lr_eudpa?product=websearch>.

[Grover-2019] Grover, G., Singh, K., and E. Hickok, Ed., "Reliance Jio is using SNI inspection to block websites", November 2019, <https://cis-india.org/internet-governance/blog/reliance-jio-is-using-sni-inspection-to-block-websites>.

[HADOPI] Hadopi, "Hadopi | Haute Autorité pour la diffusion des oeuvres et la protection des droits sur internet", <https://www.hadopi.fr/>.

[Halley-2008] Halley, B., "How DNS cache poisoning works", October 2008, <https://www.networkworld.com/article/2277316/tech-primers/tech-primers-how-dns-cache-poisoning-works.html>.

[Heacock-2009] Heacock, R., "China shuts down Internet in Xinjiang region after riots", OpenNet Initiative, July 2009, <https://opennet.net/blog/2009/07/china-shuts-down-internet-xinjiang-region-after-riots>.

[Hepting-2011] Wikipedia, "Hepting v. AT&T", September 2023, <https://en.wikipedia.org/wiki/Hepting_v._AT%26T&oldid=1175143505>.

[Hertel-2015] Hertel, O., "Comment les autorités peuvent bloquer un site Internet" [How authorities can block a website], March 2015, <https://www.sciencesetavenir.fr/high-tech/comment-les-autorites-peuvent-bloquer-un-site-internet_35828>.

[Hjelmvik-2010] Hjelmvik, E. and W. John, "Breaking and Improving Protocol Obfuscation", Technical Report No. 2010-05, ISSN 1652-926X, July 2010, <https://www.iis.se/docs/hjelmvik_breaking.pdf>.

[Husak-2016] Husák, M., Čermák, M., Jirsík, T., and P. Čeleda, "HTTPS traffic analysis and client identification using passive SSL/TLS fingerprinting", DOI 10.1186/s13635-016-0030-7, February 2016, <https://link.springer.com/article/10.1186/s13635-016-0030-7>.

[ICANN-2012] ICANN Security and Stability Advisory Committee, "Guidance for Preparing Domain Name Orders, Seizures & Takedowns", January 2012, <https://www.icann.org/en/system/files/files/guidance-domain-seizures-07mar12-en.pdf>.

[ICANN-SSAC-2012] ICANN Security and Stability Advisory Committee (SSAC), "SAC 056: SSAC Advisory on Impacts of Content Blocking via the Domain Name System", October 2012, <https://www.icann.org/en/system/files/files/sac-056-en.pdf>.

[Jones-2014] Jones, B., Lee, T-W., Feamster, N., and P. Gill, "Automated Detection and Fingerprinting of Censorship Block Pages", IMC '14: Proceedings of the 2014 Conference on Internet Measurement Conference, Pages 299-304, DOI 10.1145/2663716.2663722, November 2014, <http://conferences2.sigcomm.org/imc/2014/papers/p299.pdf>.

[Khattak-2013] Khattak, S., Javed, M., Anderson, P.D., and V. Paxson, "Towards Illuminating a Censorship Monitor's Model to Facilitate Evasion", August 2013, <http://0b4af6cdc2f0c5998459-c0245c5c937c5dedcca3f1764ecc9b2f.r43.cf2.rackcdn.com/12389-foci13-khattak.pdf>.

[Knight-2005] Knight, W., "Iranian net censorship powered by US technology", June 2005, <https://www.newscientist.com/article/dn7589-iranian-net-censorship-powered-by-us-technology/>.

[Knockel-2021] Knockel, J. and L. Ruan, "Measuring QQMail's automated email censorship in China", FOCI '21: Proceedings of the ACM SIGCOMM 2021 Workshop on Free and Open Communications on the Internet, Pages 8-15, DOI 10.1145/3473604.3474560, April 2021, <https://dl.acm.org/doi/10.1145/3473604.3474560>.

[Kravtsova-2012] Kravtsova, Y., "Cyberattacks Disrupt Opposition's Election", The Moscow Times, October 2012, <http://www.themoscowtimes.com/news/article/cyberattacks-disrupt-oppositions-election/470119.html>.

[Leyba-2019] Leyba, K., Edwards, B., Freeman, C., Crandall, J., and S. Forrest, "Borders and gateways: measuring and analyzing national as chokepoints", COMPASS '19: Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies, pages 184-194, DOI 10.1145/3314344.3332502, July 2019, <https://doi.org/10.1145/3314344.3332502>.

[Li-2017] Li, F., Razaghpanah, A., Molavi Kakhki, A., Akhavan Niaki, A., Choffnes, D., Gill, P., and A. Mislove, "lib•erate, (n): a library for exposing (traffic-classification) rules and avoiding them efficiently", DOI 10.1145/3131365.3131376, November 2017, <https://david.choffnes.com/pubs/liberate-imc17.pdf>.

[Lomas-2019] Lomas, N., "Github removes Tsunami Democràtic's APK after a takedown order from Spain", October 2019, <https://techcrunch.com/2019/10/30/github-removes-tsunami-democratics-apk-after-a-takedown-order-from-spain/>.

[Marczak-2015] Marczak, B., Weaver, N., Dalek, J., Ensafi, R., Fifield, D., McKune, S., Rey, A., Scott-Railton, J., Deibert, R., and V. Paxson, "An Analysis of China's "Great Cannon"", August 2015, <https://www.usenix.org/system/files/conference/foci15/foci15-paper-marczak.pdf>.

[Muncaster-2013] Muncaster, P., "Malaysian election sparks web blocking/DDoS claims", The Register, May 2013, <http://www.theregister.co.uk/2013/05/09/malaysia_fraud_elections_ddos_web_blocking/>.

[Murdoch-2008] Murdoch, S. J. and R. Anderson, "Tools and Technology of Internet Filtering" in "Access Denied: The Practice and Policy of Global Internet Filtering", DOI 10.7551/mitpress/7617.003.0006, 2008, <https://doi.org/10.7551/mitpress/7617.003.0006>.

[NA-SK-2019] Morgus, R., Sherman, J., and S. Nam, "Analysis: South Korea's New Tool for Filtering Illegal Internet Content", March 2019, <https://www.newamerica.org/cybersecurity-initiative/c2b/c2b-log/analysis-south-koreas-sni-monitoring/>.

[Nabi-2013] Nabi, Z., "The Anatomy of Web Censorship in Pakistan", August 2013, <http://0b4af6cdc2f0c5998459-c0245c5c937c5dedcca3f1764ecc9b2f.r43.cf2.rackcdn.com/12387-foci13-nabi.pdf>.

[NBC-2014] NBC News, "Exclusive: Snowden Docs Show UK Spies Attacked Anonymous, Hackers", February 2014, <http://www.nbcnews.com/feature/edward-snowden-interview/exclusive-snowden-docs-show-uk-spies-attacked-anonymous-hackers-n21361>.

[Netsec-2011] n3t2.3c, "TCP-RST Injection", October 2011, <https://nets.ec/TCP-RST_Injection>.

[OONI-2018] Evdokimov, L., "Iran Protests: DPI blocking of Instagram (Part 2)", February 2018, <https://ooni.org/post/2018-iran-protests-pt2/>.

[OONI-2019] Singh, S., Filastò, A., and M. Xynou, "China is now blocking all language editions of Wikipedia", May 2019, <https://ooni.org/post/2019-china-wikipedia-blocking/>.

[Orion-2013] Orion, E., "Zimbabwe election hit by hacking and DDoS attacks", Wayback Machine archive, August 2013, <https://web.archive.org/web/20130825010947/http://www.theinquirer.net/inquirer/news/2287433/zimbabwe-election-hit-by-hacking-and-ddos-attacks>.

[Patil-2019] Patil, S. and N. Borisov, "What can you learn from an IP?", Proceedings of the Applied Networking Research Workshop, Pages 45-51, DOI 10.1145/3340301.3341133, July 2019, <https://irtf.org/anrw/2019/anrw2019-final44-acmpaginated.pdf>.

[Porter-2005] Porter, T., "The Perils of Deep Packet Inspection", 2010, <http://www.symantec.com/connect/articles/perils-deep-packet-inspection>.

[Rambert-2021] Rampert, R., Weinberg, Z., Barradas, D., and N. Christin, "Chinese Wall or Swiss Cheese? Keyword filtering in the Great Firewall of China", DOI 10.1145/3442381.3450076, April 2021, <https://www.andrew.cmu.edu/user/nicolasc/publications/Rambert-WWW21.pdf>.

[Reda-2017] Reda, F., "New EU law prescribes website blocking in the name of "consumer protection"", November 2017, <https://felixreda.eu/2017/11/eu-website-blocking/>.

[RFC6066] Eastlake 3rd, D., "Transport Layer Security (TLS) Extensions: Extension Definitions", RFC 6066, DOI 10.17487/RFC6066, January 2011, <https://www.rfc-editor.org/info/rfc6066>.

[RFC7624] Barnes, R., Schneier, B., Jennings, C., Hardie, T., Trammell, B., Huitema, C., and D. Borkmann, "Confidentiality in the Face of Pervasive Surveillance: A Threat Model and Problem Statement", RFC 7624, DOI 10.17487/RFC7624, August 2015, <https://www.rfc-editor.org/info/rfc7624>.

[RFC7754] Barnes, R., Cooper, A., Kolkman, O., Thaler, D., and E. Nordmark, "Technical Considerations for Internet Service Blocking and Filtering", RFC 7754, DOI 10.17487/RFC7754, March 2016, <https://www.rfc-editor.org/info/rfc7754>.

[RFC7858] Hu, Z., Zhu, L., Heidemann, J., Mankin, A., Wessels, D., and P. Hoffman, "Specification for DNS over Transport Layer Security (TLS)", RFC 7858, DOI 10.17487/RFC7858, May 2016, <https://www.rfc-editor.org/info/rfc7858>.

[RFC8484] Hoffman, P. and P. McManus, "DNS Queries over HTTPS (DoH)", RFC 8484, DOI 10.17487/RFC8484, October 2018, <https://www.rfc-editor.org/info/rfc8484>.

[RFC8744] Huitema, C., "Issues and Requirements for Server Name Identification (SNI) Encryption in TLS", RFC 8744, DOI 10.17487/RFC8744, July 2020, <https://www.rfc-editor.org/info/rfc8744>.

[RFC9000] Iyengar, J., Ed. and M. Thomson, Ed., "QUIC: A UDP-Based Multiplexed and Secure Transport", RFC 9000, DOI 10.17487/RFC9000, May 2021, <https://www.rfc-editor.org/info/rfc9000>.

[RFC9293] Eddy, W., Ed., "Transmission Control Protocol (TCP)", STD 7, RFC 9293, DOI 10.17487/RFC9293, August 2022, <https://www.rfc-editor.org/info/rfc9293>.

[Rushe-2014] Rushe, D., "Bing censoring Chinese language search results for users in the US", The Guardian, February 2014, <http://www.theguardian.com/technology/2014/feb/11/bing-censors-chinese-language-search-results>.

[RWB-2020] Reporters Without Borders (RSF), "2020 World Press Freedom Index: 'Entering a decisive decade for journalism, exacerbated by coronavirus'", April 2020, <https://rsf.org/en/2020-world-press-freedom-index-entering-decisive-decade-journalism-exacerbated-coronavirus>.

[Sandvine-2015] Sandvine, "Internet Traffic Classification: A Sandvine Technology Showcase", 2015, <https://www.researchgate.net/profile/Nirmala-Svsg/post/Anybody-working-on-Internet-traffic-classification/attachment/59d63a5779197b807799782d/AS%3A405810988503040%401473764287142/download/traffic-classification-identifying-and-measuring-internet-traffic.pdf>.

[Satija-2021] Satija, S. and R. Chatterjee, "BlindTLS: Circumventing TLS-based HTTPS censorship", FOCI '21: Proceedings of the ACM SIGCOMM 2021 Workshop on Free and Open Communications on the Internet, Pages 43-49, DOI 10.1145/3473604.3474564, August 2021, <https://sambhav.info/files/blindtls-foci21.pdf>.

[Schoen-2007] Schoen, S., "EFF tests agree with AP: Comcast is forging packets to interfere with user traffic", October 2007, <https://www.eff.org/deeplinks/2007/10/eff-tests-agree-ap-comcast-forging-packets-to-interfere>.

[Senft-2013] Crete-Nishihata, M., Dalek, J., Hardy, S., Hilts, A., Kleemola, K., Ng, J., Poetranto, I., Senft, A., Sinpeng, A., Sonne, B., and G. Wiseman, "Asia Chats: Analyzing Information Controls and Privacy in Asian Messaging Applications", November 2013, <https://citizenlab.org/2013/11/asia-chats-analyzing-information-controls-privacy-asian-messaging-applications/>.

[Shbair-2015] Shbair, W. M., Cholez, T., Goichot, A., and I. Chrisment, "Efficiently Bypassing SNI-based HTTPS Filtering", May 2015, <https://hal.inria.fr/hal-01202712/document>.

[Siddiqui-2022] Siddiqui, A., "Lesson Learned: Twitter Shored Up Its Routing Security", March 2022, <https://www.manrs.org/2022/03/lesson-learned-twitter-shored-up-its-routing-security/>.

[SIDN-2020] Moura, G., "Detecting and Taking Down Fraudulent Webshops at the .nl ccTLD", February 2020, <https://labs.ripe.net/Members/giovane_moura/detecting-and-taking-down-fraudulent-webshops-at-a-cctld>.

[Singh-2019] Singh, K., Grover, G., and V. Bansal, "How India Censors the Web", DOI 10.48550/arXiv.1912.08590, December 2019, <https://arxiv.org/abs/1912.08590>.

[Sophos-2023] Sophos, "Sophos Firewall: Web filtering basics", 2023, <https://support.sophos.com/support/s/article/KB-000036518?language=en_US>.

[SSAC-109-2020] ICANN Security and Stability Advisory Committee (SSAC), "SAC109: The Implications of DNS over HTTPS and DNS over TLS", March 2020, <https://www.icann.org/en/system/files/files/sac-109-en.pdf>.

[Tang-2016] Tang, C., "In-depth analysis of the Great Firewall of China", December 2016, <https://www.cs.tufts.edu/comp/116/archive/fall2016/ctang.pdf>.

[Thomson-2012] Thomson, I., "Syria cuts off internet and mobile communication", The Register, November 2012, <http://www.theregister.co.uk/2012/11/29/syria_internet_blackout/>.

[TLS-ESNI] Rescorla, E., Oku, K., Sullivan, N., and C. A. Wood, "TLS Encrypted Client Hello", Work in Progress, Internet-Draft, draft-ietf-tls-esni-17, 9 October 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-tls-esni-17>.

[Tor-2019] Tor, "Tor: Pluggable Transports", 2019, <https://2019.www.torproject.org/docs/pluggable-transports.html.en>.

[Trustwave-2015] Trustwave, "Filter : SNI extension feature and HTTPS blocking", 2015, <https://www3.trustwave.com/software/8e6/hlp/r3000/files/1system_filter.html>.

[Tschantz-2016] Tschantz, M., Afroz, S., Anonymous, and V. Paxson, "SoK: Towards Grounding Censorship Circumvention in Empiricism", DOI 10.1109/SP.2016.59, May 2016, <https://oaklandsok.github.io/papers/tschantz2016.pdf>.

[Van-der-Sar-2007] Van der Sar, E., "How To Bypass Comcast's BitTorrent Throttling", October 2012, <https://torrentfreak.com/how-to-bypass-comcast-bittorrent-throttling-071021>.

[Verkamp-2012] Verkamp, J. P. and M. Gupta, "Inferring Mechanics of Web Censorship Around the World", August 2012, <https://www.usenix.org/system/files/conference/foci12/foci12-final1.pdf>.

[Victor-2019] Victor, D., "Blizzard Sets Off Backlash for Penalizing Hearthstone Gamer in Hong Kong", The New York Times, October 2019, <https://www.nytimes.com/2019/10/09/world/asia/blizzard-hearthstone-hong-kong.html>.

[Villeneuve-2011] Villeneuve, N. and M. Crete-Nishihata, "Open Access: Chapter 8, Control and Resistance, Attacks on Burmese Opposition Media", January 2011, <http://access.opennet.net/wp-content/uploads/2011/12/accesscontested-chapter-08.pdf>.

[VonLohmann-2008] VonLohmann, F., "FCC Rules Against Comcast for BitTorrent Blocking", August 2008, <https://www.eff.org/deeplinks/2008/08/fcc-rules-against-comcast-bit-torrent-blocking>.

[Wagner-2009] Wagner, B., "Deep Packet Inspection and Internet Censorship: International Convergence on an 'Integrated Technology of Control'", Global Voices Advocacy, 2009, <http://advocacy.globalvoicesonline.org/wp-content/uploads/2009/06/deeppacketinspectionandinternet-censorship2.pdf>.

[Wagstaff-2013] Wagstaff, J., "In Malaysia, online election battles take a nasty turn", NBC News, May 2013, <https://www.nbcnews.com/tech/tech-news/malaysia-online-election-battles-take-nasty-turn-flna6c9783842>.

[Wang-2017] Wang, Z., Cao, Y., Qian, Z., Song, C., and S.V. Krishnamurthy, "Your State is Not Mine: A Closer Look at Evading Stateful Internet Censorship", DOI 10.1145/3131365.3131374, November 2017, <https://www.cs.ucr.edu/~zhiyunq/pub/imc17_censorship_tcp.pdf>.

[Wang-2020] Wang, Z., Zhu, S., Cao, Y., Qian, Z., Song, C., Krishnamurthy, S.V., Chan, K.S., and T.D. Braun, "SYMTCP: Eluding Stateful Deep Packet Inspection with Automated Discrepancy Discovery", DOI 10.14722/ndss.2020.24083, February 2020, <https://www.cs.ucr.edu/~zhiyunq/pub/ndss20_symtcp.pdf>.

[Weaver-2009] Weaver, N., Sommer, R., and V. Paxson, "Detecting Forged TCP Reset Packets", September 2009, <http://www.icir.org/vern/papers/reset-injection.ndss09.pdf>.

[Whittaker-2013] Whittaker, Z., "1,168 keywords Skype uses to censor, monitor its Chinese users", March 2013, <http://www.zdnet.com/1168-keywords-skype-uses-to-censor-monitor-its-chinese-users-7000012328/>.

[Wikip-DoS] Wikipedia, "Denial-of-service attack", March 2016, <https://en.wikipedia.org/w/index.php?title=Denial-of-service_attack&oldid=710558258>.

[Wilde-2012] Wilde, T., "Knock Knock Knockin' on Bridges Doors", The Tor Project, July 2012, <https://blog.torproject.org/blog/knock-knock-knockin-bridges-doors>.

[Winter-2012] Winter, P. and S. Lindskog, "How China Is Blocking Tor", April 2012, <http://arxiv.org/pdf/1204.0447v1.pdf>.

[WP-Def-2020] Wikipedia, "Censorship", March 2020, <https://en.wikipedia.org/w/index.php?title=Censorship&oldid=943938595>.

[Wright-2013] Wright, J. and Y. Breindl, "Internet filtering trends in liberal democracies: French and German regulatory debates", DOI 10.14763/2013.2.122, April 2013, <https://policyreview.info/articles/analysis/internet-filtering-trends-liberal-democracies-french-and-german-regulatory-debates>.

[Zhu-2011] Zhu, T., Bronk, C., and D.S. Wallach, "An Analysis of Chinese Search Engine Filtering", DOI 10.48550/arXiv.1107.3794, July 2011, <http://arxiv.org/ftp/arxiv/papers/1107/1107.3794.pdf>.

[Zmijewski-2014] Zmijewski, E., "Turkish Internet Censorship Takes a New Turn", Wayback Machine archive, March 2014, <http://web.archive.org/web/20200726222723/ https://blogs.oracle.com/internetintelligence/turkish-internet-censorship-takes-a-new-turn>.

Благодарности

В работе над документом принимали участие David Belson, Stéphane Bortzmeyer, Vinicius Fortuna, Gurshabad Grover, Andrew McConachie, Martin Nilsson, Michael Richardson, Patrick Vacek, Chris Wood.

Соавтор документа Hall работал над документом до прихода в Internet Society и вовлеченность в эту организацию указана лишь для идентификации.

Адреса авторов

Joseph Lorenzo Hall
Internet Society
Email: hall@isoc.org

Michael D. Aaron
CU Boulder
Email: michael.drew.aaron@gmail.com

Amelia Andersdotter
Email: amelia.ietf@andersdotter.cc

Ben Jones
Email: ben.jones.irtf@gmail.com

Nick Feamster
U Chicago
Email: feamster@uchicago.edu

Mallory Knodel
Center for Democracy & Technology
Email: mknodel@cdt.org

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru