Цензурирование в интернете: контраргументы технические

Продолжение статьи о цензуре в интернете.

Контраргументы

Существуют три основные группы аргументов, которые можно использовать для борьбы с идеями цензуры и наблюдения в Интернете:

- Технические и технологические обоснования.
- Экономика реализации цензуры.
- Философские и гуманитарные причины (в том числе базовые принципы, существующие во всех конституциях демократических государств: права человека, свобода слова, свобода вероисповедания и т. п.).

В конце этого раздела мы приведем некоторые полезные соображения и конкретные аргументы.

Хорошая новость в том, что заключения, которые рассматриваются во всех трёх областях, содержат очень сильные аргументы против цензуры. Но есть и плохая, плохая новость, которая заключается в том, что все три типа необходимо перевести или использовать в эмоциональных аргументах, чтобы они как-то могли повлиять на общественное мнение.

Опять же, как правило, ни общественность, ни политики и государственные служащие, которые продвигают интернет-цензурирование, не имеют достаточного и глубокого понимания затронутых вопросов. Постановка вопросов в легкодоступных и эмоционально нагруженных примерах или метафорах – очевидно, будет достаточно мощной тактикой.

Несколько контр-аргументов (например, "подвержение опасности развития цифровой экономики" или "вытеснение заблокированного контента в darknet", как описывается ниже) связаны с "Законом о непредвиденных последствиях", который проявляется в экономической, политической и социальной сферах деятельности людей. Концепция (закон) формулируется так: "совокупность свойств как результат действий, имеющих ненамеренные (неожидаемые) эффекты в дополнение к тем, которые были целью этих действий". Или "любые действия с достаточно сложными (в том числе – социальными и/или экономическими) системами, основанные на неполной по факту информации, могут привести к непредсказуемым результатам". Введение цензуры в Интернете – именно такой случай.

Кроме того, стоит иметь в виду, что противодействие введению цензуры нельзя отнести к какому-либо конкретному "политическому уклону". Цензуре могут противостоять как "левые", так и "либеральные" партии и движения. Цензура и свобода слова – это вопросы, представляющие интерес для людей с любой стороны политического спектра, что позволяет охватить социальные группы, которые не желают соглашаться с вами по другим вопросам. Возможно, это имеет важное значение.

Технические аргументы

Из-за исторически сложившихся архитектурных и технологических особенностей существования "феномена Интернета", есть несколько серьезных технических аргументов против интернет-цензуры. К основным категориям относятся:

- Введение цензурирования требует далеко идущих инфраструктурных и топологических изменений в сети;
- Тотальная фильтрация требует использования специального оборудования высокого класса и требований к производительности, которое, в любом случае не сможет эффективно обрабатывать весь интернет-трафик;
- Любые фильтры и блокировки не работают: они легко обходятся при необходимости, при этом никогда не будет блокироваться все, что нужно, но при этом будут блокироваться ресурсы, которые не должны быть этому подвержены.

Существует несколько технологических способов блокировки/фильтрации контента в Интернете. Кроме того, существует несколько уровней, на которых действует цензура. Каждый из способов фильтрации имеет свои сильные и слабые стороны, но ни один из них не может гарантировать стопроцентную эффективность.

При этом нужно учесть, что все способы блокировок имеют проблемы как с чрезмерной блокировкой, так и с недоблокировкой. Все системы фильтрации стоят денег, все они требуют обслуживания, энергоресурсов и общественного мониторинга корректной работы.

При этом следует учесть:

Эффективность интернет-цензуры никогда не бывает полной, поскольку существует множество способов обхода любой системы блокировки (в зависимости от предпринятых мер).

Пере-блокировки (Over-Blocking) возникают, когда под фильтр попадает законный контент, который не следует блокировать. Обычно переблокировки возникают случайно. Но возможно проявление и злонамеренных блокировок легального контента путём вмешательства в работу систем фильтрации. В зависимости от выбранной технологической схемы организации фильтров, переблокировки могут оказаться более или менее выраженной проблемой. Но риск переблокировок присутствует всегда и он неизбежен. Метод борьбы с переблокировками – так называемые, "белые списки". Но следует понимать, что "белые списки" это еще одна точка нагрузки на системы фильтрации, которые ведут к еще большему усложнению систем блокирования.

Недоблокировки (Under-blocking) – случай, когда контент, который официально должен быть заблокирован, не блокируется. Имеется в виду – без использования средств обхода блокировок. Различные информационные ресурсы могут использовать средства по противодействию блокировкам, причём, в ряде случаев, цензурным ведомствам просто нечего противопоставить таким действиям, либо ущерб от блокировки конкретного ресурса может быть слишком большим.

Любые блокировки требуют усилий со стороны эксплуатирующей фильтры организации и **затрачиваемые ресурсы** (например, специальное оборудование, его вычислительная мощность, пропускная способность каналов связи) не всегда адекватны полученному эффекту.

Независимо от того, применяется ли метод "глубокой инспекции пакетов" (DPI) или используются более простые схемы блокировок – фильтрация является интрузивным (навязчивым) и ресурсоемким мероприятием.

Ниже приводится краткое описание возможных методов блокировок с разбором вышеприведенных факторов. Возможные методы обхода фильтрации будут приведены в конце этого раздела.

Блокировки по DNS:

- Вероятность переблокировок: высокая.
- Вероятность недоблокировок: средняя.
- Требуемые ресурсы: небольшие.
- Стоимость решения: средняя.
- Обход: очень легко.
- Использует DPI: НЕТ.

Блокировки "по DNS" базируются на том, что интернет-провайдеры, которые, как правило, используют собственные DNS-серверы по умолчанию для своих клиентов, исключают из разрешения (resolving) список определённых доменов. То есть, при обращении к заблокированному ресурсу провайдерский DNS-сервер просто выдаёт собственный IP-адрес ресурса, где может быть размещено, например, объявление "ресурс заблокирован". Это означает, что затраты провайдера на реализацию такого решения незначительны – нужно просто исправить (заменить или совсем удалить) соответствующую запись в локальной версии DNS-реестра.

Однако, пользователи могут легко использовать другие DNS-серверы, просто переконфигурировав сетевое подключение на, например, адрес публичного DNS-сервиса (8.8.8.8 или 8.8.4.4), что не является очень сложной технической задачей. И, таким образом, метод блокировок "по DNS" чрезвычайно легко обойти.

При этом, метод обладает огромным потенциалом для чрезмерной блокировки, поскольку из-за определённого содержимого на отдельных страницах, блокируются целые домены. Это означает, что в случае цензурного недовольства отдельными публикациями на каком-то домене будет заблокирован весь домен целиком. Возможно, из-за всего одной страницы или сообщения на форуме.

Кроме того, действительно криминальные сайты, целенаправленно публикующие незаконный контент, просто изменят домен или добавят еще одно доменное имя-синоним, и снова станут доступными. Причем, для целенаправленной борьбы с блокировками, такие сайты могут часто (иногда в течение нескольких часов) изменять доменные имена, и таким образом оставаться доступными для всех пользователей провайдера. Затраты на обработку списков блокировки и риск недоблокировки при этом являются средними.

Блокировка по IP-адресу:

- Вероятность переблокировок: высокая.
- Вероятность недоблокировок: средняя.

- Требуемые вычислительные ресурсы: небольшие.
- Стоимость решения: средняя.
- Обход: достаточно легко.
- Использует DPI: НЕТ.

Блокировки по IP-адресам основаны на требовании, чтобы интернет-провайдеры либо блокировали определенные IP-адреса внутри собственной сети, либо маршрутизировали все исходящие соединения через центральную, уполномоченную правительством, систему цензурирования. Данный способ фильтрации обходится несколько труднее, чем блокирование по DNS, но все равно достаточно просто.

Блокировка по IP-адресам, также, как и блокировка на базе DNS, не требуют внедрения систем DPI.

Веб-сайты, которые целенаправленно публикуют "запрещенный правительством контент", легко обходят блокировки по IP, изменив собственный IP-адрес в A-записях DNS. Что технически лишь немного сложнее, чем изменение доменного имени, но при этом пользователи, желающие получить доступ к заблокированным веб-сайтам, могут использовать множество методов обхода цензурных фильтров.

Существует возможно повысить эффективность блокировок по IP-адресу, и тем самым сформировать систему фильтрации более трудную для обхода – создать фильтры, блокирующие целые диапазоны (блоки) IP, однако это резко повышает вероятность чрезмерных блокировок.

URL-фильтрация:

- Вероятность переблокировок: низкая.
- Вероятность недоблокировок: высокая.
- Требуемые вычислительные ресурсы: средние.
- Стоимость решения: высокая.
- Обход: средней сложности.
- Использует DPI: ДА.

Метод фильтрации URL (Uniform Resource Locator – единый указатель ресурса) использует DPI (Deep Packet Inspection) – технологию накопления статистических данных, проверки и фильтрации сетевых пакетов по их содержимому.

Поскольку этот метод фильтрации блокирует только определенный контент с URL-адресами, а не целые веб-сайты или серверы, как и методы фильтрации по DNS и IP, то метод имеет гораздо более низкую вероятность случайного блокирования. Но это же влечет, что у метода имеется более высокий потенциал для недоблокировок, поскольку контент может быть доступен на одном и том же сервере под разными URL-адресами. Изменение лишь небольшой части имени ссылки, делает фильтр бесполезным.

И нужно отметить, что интернет-пользователи, которые хотят получить доступ к заблокированному контенту – имеют множество методов обхода, включая прокси, VPN, TOR, Darknet’ы – все это мы обсудим ниже.

Динамические блокировки:

- Вероятность переблокировок: высокая;
- Вероятность недоблокировок: высокая;
- Требуемые вычислительные ресурсы: очень много;
- Стоимость решения: высокая;
- Обход: средней сложности;
- Использует DPI: ДА.

Этот метод использует систему DPI для чтения содержимого передаваемых данных и сравнивает его со списком ключевых слов или с образцами изображений или видео (в зависимости от типа содержимого).

Динамические блокировки имеют очень большую вероятность чрезмерной блокировки. В качестве примера: вы не получите доступа к, например, сайтам про "Формулу-1", только потому, что самый популярный чемпион гонок немец по фамилии Шумахер. Или множество сайтов, включая Википедию, где имеется медицинская информация. Но при этом, это совершенно не избавит вас от контента, написанного эзоповым языком или специфичными арго – недоблокировки тут будут очень значительными, особенно, если авторы "нежелательного контента" знают о попытках их

блокировать.

При этом, нужно понимать, что борьба с недоблокировками путём расширения списков ключевых слов только усугубляет проблему с переблокировками. Чем больше вы добавляете новых слов в стоп-листы, тем больше вероятность того, что под "запрет" попадет вполне легальный, а иногда даже полезный контент.

И хотя прямые затраты на обработку (создание и редакцию) списков сравнительно низки, но для технической реализации этого метода требуются огромные ресурсы по вычислительным мощностям и пропускной способности каналов. Это связано с тем, что каждый поток данных в сети необходимо проверить и сравнивать с ключевыми словами и шаблонами. И тем более такая система будет очень требовательной для цензуры изображений, видео и других нетекстовых сообщений. Даже без учета шифрования.

И пользователи по-прежнему смогут обойти фильтр множеством способов.

Блокировки на хэш-функциях:

- Вероятность переблокировок: низкая;
- Вероятность недоблокировок: высокая;
- Требуемые вычислительные ресурсы: много;
- Стоимость решения: высокая;
- Обход: средней сложности;
- Использует DPI: ДА.

Суть метода блокировок, основанных на хэшах – используемая DPI-система осуществляет проверку содержимого потоков данных, хэширует их криптографическими хэш-функциями и сравнивает с известной базой данных хэшей, которые необходимо заблокировать. Данный метод имеет низкую вероятность для чрезмерной блокировки (в зависимости от качества используемых хэш-функций), но очень высокий потенциал для недоблокировок, поскольку одно небольшое изменение содержания передаваемой информации, влечет за собой изменение хэша и, следовательно, контент не будет заблокирован.

При этом, метод имеет очень высокие потребности в вычислительных ресурсах – все потоки данных необходимо кодировать, а полученные хэши сверять с базой заблокированного в режиме реального времени. И нужно понимать, что издержки вычислительных ресурсов на обработку хэш-данных также вполне весомы.

Гибридные решения (например на основе IP и hash):

- Вероятность переблокировок: низкая;
- Вероятность недоблокировок: высокая;
- Требуемые вычислительные ресурсы: средние;
- Стоимость решения: высокая;
- Обход: средней сложности;
- Использует DPI: ДА.

Разумеется, может быть найден некий компромисс между высокоуровневыми и относительно затратными блокировками на основе хэшей и низкоуровневыми и недорогими решениями по фильтрации контента на базе IP или DNS. Обычно, эти решения подразумевают, что существует некий "черный список" IP-адресов или доменных имен, для которых включена блокировка на основе хеша, поэтому система работает только для небольшой части передаваемого трафика. При этом, метод всё равно должен использовать "глубокую проверку пакетов".

И в данном случае необходимые ресурсы и затраты на обработку списков весьма значительны, вероятность переблокировок слишком высока, а обход таких фильтров на уровне пользователей не потребует глубоких знаний сетевых технологий.

Способы обхода блокировок

Все вышеперечисленные способы фильтрации контента и интернет-цензуры бессильны, если пользователи решат эти "преграды" обойти. Существует огромное количество методов обхода блокировок, использование которых зависит только от желания и квалификации пользователя. Причём замечено, что чем больше применяется блокировок, тем более изощрёнными становятся методы их обхода. И тем больше людей начинает ими интересоваться и использовать на практике.

Например, для обхода **фильтров по DNS** применяются пользовательские настройки DNS-сервера.

Метод не требует никакого технического мастерства или глубоких знаний – достаточно "прописать в настройках" компьютера и/или домашнего маршрутизатора собственный, а не провайдерский, сервер DNS, которых в настоящее время существует огромное количество. Самые известные из которых – так называемые, "публичные DNS-серверы" от крупнейших интернет-компаний (например, Яндекс.DNS, Google или Cloud Flare). И для того, чтоб в действительности использовать метод блокировок по DNS, необходимо полноценное развертывание методов цензуры, включая карательные методы.

Прокси-серверы, особенно анонимные (или анонимизирующие), расположенные за пределами страны, где установлены системы фильтрации, очень просты в использовании для обхода любых цензурирующих систем. Для того, чтобы начать пользоваться прокси-сервером достаточно немного изменить настройки в своей операционной системе или скорректировать настройки браузера. Не менее тривиально можно установить специальный плагин (дополнение) для браузера и использовать его при необходимости. При этом, можно заблокировать сами прокси-серверы (посредством блокировки, например, соответствующих IP-адресов), однако их количество по всему миру столь велико, что физически невозможно не только заблокировать их, но и просто выявить.

Виртуальные частные сети / VPN (в том числе, VPN "для бедных", например, SSH-туннели) – для настройки этой системы обхода блокировок требуется несколько больше технических знаний и мастерства, чем есть у среднестатистического пользователя. Однако, существует целый сектор услуг коммерческих VPN-сервисов, в том числе и для целей обхода блокировок. И следует оговориться, что VPN-сервисы совсем не обязательно созданы именно для антицензурных целей. Чаще всего, их создают корпорации и компании для обеспечения безопасного доступа к внутренним ресурсам и/или обеспечения информационной безопасности.

Блокировка всего трафика VPN/SSH потенциально возможна, но требует установки мощнейших DPI-систем. Кроме того, и VPN, и тем более SSH, повсеместно используется в чисто деловых целях огромным количеством людей и, что особенно важно, различными устройствами, которым необходим удалённый доступ к различным корпоративными ресурсам. Например, это могут быть банкоматы, терминалы оплаты в магазинах или любые другие промышленные устройства. И даже транспортные средства. Заблокировав весь VPN-трафик, цензура рискует оставить страну вообще без экономики.

TOR, или Onion Router или "луковичная маршрутизация". Очень эффективный способ обхода вообще любых цензурных систем. Но основной его недостаток – относительно медленная работа Сети. TOR очень просто настроить – пользователи могут просто загрузить пакет TOR Browser Bundle и получить доступ в Интернет. Поскольку это программное обеспечение специально разрабатывалось для обхода блокировок и для анонимизации пользователей, его почти невозможно заблокировать. Трафик TOR для любых систем "глубокой инспекции пакетов" выглядит как обычнейший (ванильный) трафик HTTPS, которым пользователи обмениваются между собой. Кроме того, используются всевозможные методы маскировки трафика и "запутывание маршрутизации". В итоге, TOR прекрасно работает даже в странах с агрессивной политикой интернет-цензуры, таких как Китай, Иран или даже КНДР. Этот инструмент пока является самым надежным способом обхода любых технических блокировок.

И ни одно из решений по интернет-фильтрации не может блокировать контент в так называемых Darknets – виртуальных частных сетях, соединения в которых устанавливаются только между доверенными пирами, иногда именующимися как "друзья". Причем, для организации таких сетей используются весьма нестандартные и даже нетривиальные протоколы организации связи. Для создания таких сетей используют весь доступный инструментарий анонимизации, шифрования и специализированного программного обеспечения (например, TOR, I2P, FreeNet). При этом гарантируется высокая устойчивость к цензуре посредством технической организации самих сетей.

Единственным недостатком использования "даркнетов" является их низкая пропускная способность, происходящая из архитектурно-технических особенностей организации таких сетей.

На самом деле, развёртывание систем интернет-фильтрации подталкивает потенциально подцензурный контент уходу в TOR и "даркнеты". На такой контент устанавливают атрибуты приватности в социальных сетях или используют максимально доступные настройки шифрования в мессенджерах. В итоге, все это только усложняет сбор доказательств правоохранительными органами и скрывает противоправные действия в интернете. Этот аспект будет раскрыт далее в разделе "философских аргументов".

Особо нужно отметить все более широкое использование **протоколов шифрования TLS/SSL**, которые совсем не обязательно являются инструментом обхода блокировок. Но фильтрация контента, передаваемого в зашифрованном виде, невозможна априори. Любые системы DPI могут только определить "направление трафика" – IP-адрес или вычленить домен назначения через механизм Server Name Indication (SNI). То есть, любая из существующих на сегодня система блокировок может только определить "куда" (к какому ресурсу) обращается пользователь, что делает невозможным частичную блокировку ресурса – только целиком.

Протоколы TLS/SSL обеспечивают сквозную шифрацию всего потока данных между "источником" и "клиентом" и первоначально такой способ связи задумывался для организации безопасных соединений с банковскими системами и сайтами электронной коммерции. Однако, в настоящее время количество интернет-ресурсов, использующих SSL-шифрование, превысило число "обычных интернет-ресурсов", включая, например, сервисы электронной почты или социальные сети. Отличить сайт, использующий сертификат шифрования от "обычного сайта" можно по зеленому значку "защищено" в строке браузера или по использованию символов https:// вместо http:// в адресной строке.

В общем объёме трафика, протокол https (или 443 порт в TCP/IP соединениях) сегодня занимает более половины и доля шифрованного трафика постоянно растет.

После развёртывания цензурного решения на основе DPI, заблокированные пользователи и/или сервисы будут постепенно и естественно стремиться к простому, но очень эффективному решению. Это означает, что любая DPI-система блокировок должна обрабатывать TLS/SSL соединения. И это возможно сделать только двумя способами:

- полностью заблокировать использование SSL;
- выполнить атаку "человек-в-середине" (или MiTM) на зашифрованные потоки данных.

При этом, нужно учесть, что блокировка ресурсов, использующих SSL-шифрование, не может стать жесткой – пусть бы даже и TLS/SSL-соединения технически довольно легко отфильтровываются. Причина невозможности заблокировать шифрование проста: использование TLS/SSL в принципе является допустимым, законным и часто используемым способом обеспечения информационной безопасности пользователей, особенно в легальном корпоративном секторе (и особенно в банковских и финансовых структурах). Полная блокировка любых алгоритмов шифрования не может быть жизнеспособным решением – это вызовет возмущение пользователей, компаний по безопасности и особенно финансовый сектор.

Что касается использования (фактически, хакерской практики) атаки "человек посередине" (MiTM), то это означает способ получения шифрованного потока данных, его дешифровку и проверку содержимого. Затем следует повторное шифрование и отправка данных по назначению. И предпочтительно сделать это так, чтобы ни клиент, ни сервер не обнаружил вторжения. Потенциально, при правильно созданных и подписанных сертификатах, такое решение может быть жизнеспособным, если вся цензурная система, включая оборудование, будет иметь специальный цифровой сертификат, допускающий подобные практики.

Что не является однозначно безопасным.

Отмечены случаи, когда легальные MiTM-сертификаты утекали из скомпрометированных центров сертификации (или корневых центров сертификации) и использовались репрессивными режимами для атак MITM на TLS/SSL. Здесь следует привести пример голландской корпорации DigiNotar, которая попыталась провернуть подобное в интересах Иранского правительства (подробнее на английском – история 2011 года).

Известны случаи использования MiTM-сертификатов в некоторых цензурных системах – случай с подделкой сертификатов Google от того же DigNotar довольно долго муссировался в прессе. Но нужно отметить, что подобные случаи очень плохо отражаются на бизнесе производителей и операторов подобных систем. DigNotar был обанкрочен, партнеры такого удостоверяющего центра понесли убытки (иногда несовместимые с дальнейшим ведением деятельности).

Другая сторона вопроса: реализация MiTM-атаки на TLS/SSL-соединения – весьма ресурсоемкая операция. Представьте, что вам нужно не только отфильтровать контент, но и по каждому соединению совершить, как минимум две операции дешифрования-шифрования, потом умножьте количество операций на количество SSL-потока. И с учетом, что один пользователь одновременно может устанавливать десятки таких сессий, у вас получатся требования несовместимые с нормальной деятельностью интернет-провайдера. Оборудования по фильтрации будет сильно больше, чем оборудования для "просто передачи данных". Это довольно дорого.

Однако существует другой, гораздо более весомый аргумент. Исполнение атаки "человек посередине" еще более навязчиво и еще больше нарушает базовые права человека на тайну связи, чем использование DPI. Это сознательный акт нарушения тайны связи и по сути – взлом абонентских данных. Любые передаваемые по интернет сведения становятся незащищенными и небезопасными по определению. Причем, существует не так уж и много столь враждебных цифровому праву законодательных инициатив, которые правительство может навязать своим гражданам. MiTM-атака, пожалуй, самая враждебная инициатива.

Применение MiTM для всех подключений в сети – резко снижает уровень доверия к такому типу коммуникаций в целом. Что приводит к тому, что граждане не доверяют банковским, финансовым и

коммерческим сайтам и вообще ко всем сайтам, использующим TLS/SSL, что влечёт к тому, что подрывается вообще доверие к "цифровой экономике". Ну, потому что если государство, которое не является самым инновационным участником интернет-экономики, может проводить подобную политику, то уж более продвинутые технологические компании тем более могут делать что угодно с вашими данными. И никто не поручится, что данные не попадут в руки откровенных мошенников.

Интернет, при массовом внедрении MiTM и DPI, перестаёт быть безопасной доверенной средой. Это аксиоматическое утверждение.

И в заключение раздела – MiTM-атака технологически не может продолжаться достаточно долго и будет рано или поздно обнаружена. При наличии множества субъектов коммуникаций и корневых сертифицирующих центров, рано или поздно подмена сертификатов будет обнаружена и брешь в безопасности будет опубликована. Разумеется, что поддельный сертификат будет отозван владельцем (особенно, если речь идет о международном обмене данными) и MiTM-атака перестанет работать. Причём, такой отзыв будет иметь побочный эффект в виде цепной реакции отключения всех сайтов и web-систем от всех подключений к скомпрометированным удостоверяющим центрам. Производители программного обеспечения, которые дорожат своей репутацией, также отключат поддержку сертификатов такого "удостоверяющего центра". Пользователи же увидят, что "данный сертификат недействителен" и сообщение об ошибке. Именно так все и произошло со скандалом с DigNotar.

Никто не хочет иметь дел с мошенниками. Пусть бы даже и правительственными.

Продолжение следует. Поговорим об экономических и эмоциональных контраргументах