Рекомендации по ограничению доступа к информации, распространяемой посредством информационно-телекоммуникационной сети «Интернет», в порядке, установленном Федеральным законом от 27 июля 2006 года № 149-ФЗ «Об информации, информационных технологиях и о защите информации»

1. В настоящих Рекомендациях по ограничению доступа операторами связи, оказывающими услуги по предоставлению доступа к сети «Интернет», к информации, распространяемой посредством информационно-телекоммуникационной сети «Интернет» (далее – сеть «Интернет»), в порядке, установленном Федеральным законом от 27 июля 2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации» (далее - Рекомендации), описываются мероприятия по ограничению доступа к сайтам в сети «Интернет», осуществляемые в рамках взаимодействия Федеральной службы по надзору в сфере связи, информационных технологий и массовых коммуникаций с операторами связи, оказывающими услуги по предоставлению доступа к сети «Интернет».

2. Основные понятия, используемые в Рекомендациях:
   Все ссылки на RFC взяты как можно более свежие.

   оператор связи – юридическое лицо или индивидуальный предприниматель, оказывающие услуги связи на основании соответствующей лицензии;

   выгрузка – информация из единого реестра доменных имен, указателей страниц сайтов в сети «Интернет» и сетевых адресов, позволяющих идентифицировать сайты в сети «Интернет», содержащие информацию, распространение которой в Российской Федерации запрещено, предоставляемая Роскомнадзором операторам связи для принятия мер по ограничению доступа к информации;
   Изначальное определение — ад и немцы. Я косноязычный, но это лучше.

   доменное имя - символьное имя из множества иерархии доменных имен в сети «Интернет» (в строгом соответствии с текущей поддерживаемой иерархией доменных имен https://www.iana.org/domains), служащее для идентификации ресурсов в сети «Интернет», представленное в формате ACE в соответствии с документами IETF RFC1035 и IETF RFC5890;
   Определение из закона никуда не годится. Это не противоречит, но зато применимо в нормативе.

   указатель страницы сайта в сети «Интернет» - строка, состоящая из протокола (HTTP или HTTPS), доменного имени, порта и символов, определенных владельцем сайта в сети «Интернет», позволяющая идентифицировать сайт в сети «Интернет», сформированная в соответствии с разделом 2.7 документа IETF RFC 7230;
   Дополненное уточнением формата определение из закона.

   IETF - Рабочая группа по инженерным проблемам сети Интернет (Internet Engineering Task Force)

   RFC -обозначение документа IETF (Request For Comments)

строка HTTP запроса — набор заголовков протокола HTTP, позволяющий идентифицировать запрашиваемый ресурс, как это описано в разделе 3.1.1 и в секции 5.3 документа IETF RFC7230;
Важный термин. Там есть такая штука, как Request в разделе 3.1.1., который потом в секции 5.3 подробно разбирается как может выглядеть. Это про полный URL в GET и про заголовок Host.

TLS — криптографические протоколы, обеспечивающие защищенную передачу в сети «Интернет»;

SNI — расширение протокола TLS, описанное в документе IETF RFC 6066, которое позволяет клиентам сообщать доменное имя, с которым он желает соединиться во время процесса «рукопожатия»;

сетевой адрес — уникальный сетевой адрес из множества адресов сети «Интернет» (в строгом соответствии с текущим распределенным адресным пространством https://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xhtml и https://www.iana.org/assignments/ipv6-address-space/ipv6-address-space.xhtml), представленный в виде четырех групп десятичных чисел от 0 до 255, разделенных точками для протокола IPv4, и в виде восьми групп четырехзначных шестнадцатеричных чисел, разделенных двоеточием для протокола IPv6 в соответствии с документом IETF RFC5952;
Определение искореняет всякие 127.0.0.1 и 192.168.0.0 и прочие спекуляции. Это конечно иногда смешно, но крайне разрушительно.

3. Ограничение доступа к информации, распространяемой посредством сети «Интернет», сайтам в сети «Интернет» и информационным ресурсам (далее – ограничение доступа) операторами связи осуществляется на абонентских сетях доступа. Ограничение доступа операторами связи не осуществляется на транспортах сетях, а также на межоператорских стыках, кроме случаев, когда такое ограничение осуществляется в рамках гражданских договоров.
Про транзит

4. Ограничение доступа к информации, распространяемой посредством сети «Интернет», сайтам в сети «Интернет» и информационным ресурсам (далее – ограничение доступа) операторами связи рекомендуется осуществлять в строгом соответствии с выгрузкой.

5. Экземпляр выгрузки содержит следующие реквизиты:
   - номер версии формата, в котором сформирована выгрузка;
   - момент времени, когда был сформирован данный экземпляр выгрузки;
   - момент времени, когда в выгрузку последний раз были внесены изменения, требующие незамедлительного реагирования.

6. Каждый экземпляр выгрузки подписывается усовершенствованной электронной подписью Федеральной службы по надзору в сфере связи, информационных технологий и массовых коммуникаций с меткой времени.

7. Формирование нового экземпляра выгрузки производится ежечасно. При внесении в выгрузку записей с указанием (маркером) на высокую срочность реагирования (раздел 2 и Приложение 1 Памятки операторам связи о получении доступа к выгрузке, размещенной на официальном сайте http://vigruzki.rkn.gov.ru) формируется внеочередной экземпляр выгрузки.

8. В случае недоступности выгрузки оператор связи обязан уведомить об этом ситуационный центр Роскомнадзора по своему региону в течение часа с момента обнаружения недоступности.
   Тут бы конечно чего как, но это пусть пока они теперь думают

9. Получение и обработка выгрузки производится операторами связи не позднее, чем через 25 часов после формирования выгрузки. В случае формирования внеочередного экземпляра выгрузки, в котором присутствуют записи выгрузки с указанием (маркером) на высокую срочность реагирования, получение и обработка выгрузки операторами связи в соответствии с пунктом 3 статьи 15 Федерального закона от 27 июля 2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации» производится не позднее, чем через два часа после формирования выгрузки.
   Раз в час оператор проверяет, нет ли чего срочного по «момент времени, когда в выгрузку последний раз были внесены изменения, требующие незамедлительного реагирования». Т.е. если он проверил, а через минуту выгрузка, то пятьдесят девять минут он её получит. И час на обработку и распространение.
   25 часов сделано из этого же принципа. Хотя, положа руку на сердце, вот без вот этих «срочно-срочно» можно просто раз в час и час на обработку — равно два часа и всё.

10. Запись выгрузки может содержать:
    - основание ограничения доступа;
    - момент времени, с которого возникает необходимость ограничения доступа;
    - тип срочности реагирования; реквизиты решения о необходимости ограничения доступа;
    - значение хэш-кода, позволяющего проверить внесение изменений в запись выгрузки;
    - одно или несколько доменных имен;
    - один или несколько указателей страниц сайтов в сети «Интернет»;
    - один или несколько сетевых адресов.

11. При любом способе ограничения доступа к сайту, операторам связи рекомендуется производить анализ и фильтрацию с целью ограничения доступа к сайту только трафика к сетевым адресам, указанным в записи выгрузки для данного сайта.
    Это об «ресолвит РКН»

12. При наличии в записи выгрузки информации о доменном имени с указанием маски (вида *.domain.com), относящейся к доменным именам нижестоящего уровня, в тех случаях, когда фильтрация с использованием доменного имени применима, рекомендуется ограничиваться доступ к основному доменному имени, а также ко всем доменным именам, подпадающим под маску.

В (13) и (14) везде проставлено, что сетевой адрес должен быть из той же строки, что и всё остальное, и никак иначе.

13. При наличии в записи выгрузки информации об указателях страниц сайта в сети «Интернет», доменном имени и сетевом адресе оператору связи рекомендуется ограничить доступ к указателям страниц сайта в сети «Интернет» следующим образом:
А вот тут домены используются только как справочная информация. Они вообще бессмысленны. Я не смог подогнать.

    13.1. Для указателей страниц сайта в сети «Интернет», содержащих сетевой протокол, не поддерживающий шифрование (HTTP), производится фильтрация трафика к сетевым адресам из данной записи выгрузки с целью ограничения доступа только к этим указателям страниц сайта путем анализа и сопоставления строки HTTP запроса с указателями страниц сайта из выгрузки.

    13.2. Для указателей страниц сайта в сети «Интернет», содержащих сетевой протокол, поддерживающий шифрование (HTTPS), производится фильтрация трафика к сетевым адресам из данной записи выгрузки с целью ограничения доступа к сайту путем анализа заголовка SNI и сопоставлением его с доменным именем вычисленном из указателя страниц сайта из записи выгрузки. При этом доступ ограничивается ко всему сайту.

14. При наличии в записи выгрузки информации о доменном имени и сетевом адресе при отсутствии информации об указателе страницы сайта в сети «Интернет», операторам связи рекомендуется ограничивать доступ к сайту следующим образом:
А тут кстати пришлось изобретать извращение. По-хорошему, этот пункт выглядит как «задайте нормально URL». И тогда бы только предудущим и следующим всё обошлось. Кстати. Вот это тот момент, когда «а как же если HTTP на другом порту?». Правильно — просто нормально задаём список URL по предыдущему пункту.

    14.1. Трафик к сетевым адресам из данной записи выгрузки по протоколу не поддерживающему шифрование (HTTP) на стандартном порту 80 фильтруется с целью ограничения доступа к сайту путем анализа и сопоставления строки HTTP запроса и доменного имени из записи выгрузки. При этом доступ ограничивается ко всему сайту.

    14.2. Трафик к сетевым адресам из данной записи выгрузки по протоколу поддерживающему шифрование (HTTPS) на стандартном порту 443 фильтруется с целью ограничения доступа к сайту путем анализа заголовка SNI и сопоставлением его с доменным именем из записи выгрузки. При этом доступ ограничивается ко всему сайту.

15. При наличии в записи выгрузки информации о сетевом адресе и отсутствии информации о доменном имени и указателе страницы сайта в сети «Интернет», операторам связи рекомендуется ограничить доступ к указанному сетевому адресу, включая все сетевые порты.

Выкинута вся методика. Нечего путать это в рекомендации. Вместе с методикой выкинута неявная возможность DPI фильтровать весь трафик и вылавливать из него HTTP запросы на любой порт.