(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 111314301 A
(43)申请公布日 2020.06.19

(21)申请号 202010051822.8

(22)申请日 2020.01.17

(71)申请人 武汉思普崚技术有限公司
地址 430070 湖北省武汉市东湖开发区光谷大道308号光谷动力节能环保产业园一期11栋

(72)发明人 朱光原

(74)专利代理机构 武汉智嘉联合知识产权代理事务所(普通合伙) 42231
代理人 易贤卫

(51)Int.Cl.
H04L 29/06(2006.01)
H04L 29/12(2006.01)

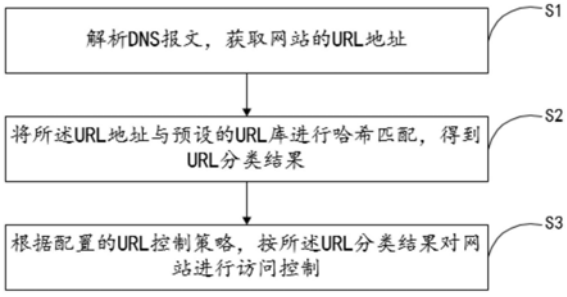
权利要求书1页 说明书4页 附图2页

(54)发明名称

一种基于DNS解析的网站访问控制方法及装置

(57)摘要

本发明涉及网站访问控制技术领域,公开了一种基于DNS解析的网站访问控制方法、装置以及计算机存储介质,其中,方法包括以下步骤:解析DNS报文,获取网站的URL地址;将所述URL地址与预设的URL库进行哈希匹配,得到URL分类结果;根据配置的URL控制策略,按所述URL分类结果对网站进行访问控制。本发明提供的基于DNS解析的网站访问控制方法、装置以及计算机存储介质,具有能够实现加密URL过滤,能够对各类协议下的网站访问进行控制的技术效果。



CN 111314301 A

1.一种基于DNS解析的网站访问控制方法,其特征在于,包括以下步骤:

解析DNS报文,获取网站的URL地址;

将所述URL地址与预设的URL库进行哈希匹配,得到URL分类结果;

根据配置的URL控制策略,按所述URL分类结果对网站进行访问控制。

2.根据权利要求1所述的基于DNS解析的网站访问控制方法,其特征在于,解析DNS报文,具体为:

判断URL控制是否打开,如果未打开则不进行网站访问控制,如果打开,则解析DNS报文。

3.根据权利要求1所述的基于DNS解析的网站访问控制方法,其特征在于,将所述URL地址与预设的URL库进行哈希匹配,得到URL分类结果之前,还包括:

将所述URL地址与URL白名单进行匹配,如果白名单匹配成功,则直接放通DNS报文,如果白名单匹配不成功,则将所述URL地址与URL黑名单进行匹配,如果黑名单匹配成功,阻断DNS报文,如果黑名单匹配不成功,则判断是否配置URL控制策略,如果未配置,则放通DNS报文,如果配置,则将所述URL地址与预设的URL库进行哈希匹配,得到URL分类结果。

4.根据权利要求3所述的基于DNS解析的网站访问控制方法,其特征在于,所述白名单匹配和黑名单匹配均采用哈希匹配。

5.根据权利要求1所述的基于DNS解析的网站访问控制方法,其特征在于,还包括,记录控制日志。

6.根据权利要求1所述的基于DNS解析的网站访问控制方法,其特征在于,所述URL库包括不同类型的URL;所述URL控制策略包括不同类型的URL及其对应的策略动作。

7.根据权利要求6所述的基于DNS解析的网站访问控制方法,其特征在于,所述策略动作包括放通、阻断以及日志警告级别。

8.一种基于DNS解析的网站访问控制装置,其特征在于,包括处理器以及存储器,所述存储器上存储有计算机程序,所述计算机程序被所述处理器执行时,实现如权利要求1-7任一所述的基于DNS解析的网站访问控制方法。

9.一种计算机存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时,实现如权利要求1-7任一所述的基于DNS解析的网站访问控制方法。

一种基于DNS解析的网站访问控制方法及装置

技术领域

[0001] 本发明涉及网站访问控制技术领域，具体涉及一种基于DNS解析的网站访问控制方法、装置以及计算机存储介质。

背景技术

[0002] 在目前常见的网络部署中，存在着企业对员工上网行为的控制管理的需要，通常采用URL过滤技术。例如：企业不允许研发员工在上班时间访问娱乐网站，在下班时间则允许；或者企业不允许市场人员访问研发内部网站等等。这些基于不同的用户组、不同的时间段可访问网页有区别的问题，可以采用URL过滤技术实现。

[0003] 常见的URL过滤功能可以归纳为3点：

[0004] 黑/白名单功能：如把钓鱼网站、黄色网站等列入黑名单，可以保护公司内网的安全；而把一些畅通无阻的网页加入白名单，就不需要进行分类查询，提高了访问速度。

[0005] 分类访问功能：对于黑/白名单无法匹配的网站，采取分类查询的功能。网站分类可以用户自己配置，也可以向第三方的分类查询服务器进行查询，如surfcontrol的分类服务器。总而言之，查询到分类结果后，可以与本地的用户组和时间段关联，判断该网站http请求是否应该放行。

[0006] 页面推送：若是被阻断的页面，需要对用户进行通知，可以采用页面推送的方式。此时，需要对发起http请求的用户推送一个页面，通知客户，访问被阻断，并且断开http请求。

[0007] 如何从网站中提取URL呢?我们知道http get请求的格式中包括host，我们把host字段提取出来，与配置的控制策略进行对比匹配，决定是放行还是过滤。

[0008] 目前的这种URL过滤方法，主要存在以下问题：

[0009] 1.当前越来越多的网站采用加密URL(https)进行访问，常见的获取host的方法因为加密的原因无法再获取到host字段，导致过滤失效。如果要进行解密和重组行为，则会导致系统资源的大量使用，影响实际效率。

[0010] 2.域名部分是通过HTTP头的Host字段来获取的，其他字段均不能保证能正确获取域名，而这个字段有的服务器并不检查，可以随便填个别的域名，服务器也可以正确返回。例如在HTTP/1.0中，这个字段更加不是必须的，因此根本不能保证获取正确的域名。

[0011] 3.只能限制http和https，其余协议的访问无法控制。

发明内容

[0012] 本发明的目的在于克服上述技术不足，提供一种基于DNS解析的网站访问控制方法、装置以及计算机存储介质，解决现有技术中无法对加密URL进行过滤，只能对特定协议网站访问进行控制的技术问题。

[0013] 为达到上述技术目的，本发明的技术方案提供一种基于DNS解析的网站访问控制方法，包括以下步骤：

[0014] 解析DNS报文,获取网站的URL地址;

[0015] 将所述URL地址与预设的URL库进行哈希匹配,得到URL分类结果;

[0016] 根据配置的URL控制策略,按所述URL分类结果对网站进行访问控制。

[0017] 本发明还提供一种基于DNS解析的网站访问控制装置,包括处理器以及存储器,所述存储器上存储有计算机程序,所述计算机程序被所述处理器执行时,实现所述基于DNS解析的网站访问控制方法。

[0018] 本发明还提供一种计算机存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时,实现所述基于DNS解析的网站访问控制方法。

[0019] 与现有技术相比,本发明的有益效果包括:本发明在DNS解析时,即进行URL过滤,从DNS报文中解析出URL,从而实现域名提取判断,由于DNS报文一般是UDP包,不需要进行解密,从而可以实现加密URL的过滤,同时,UDP包不需要报文重组,从而降低了过滤过程对各种资源的消耗;再次,DNS报文的解析并不局限于http和https,因此基于DNS报文解析进行URL过滤适用于不同的协议;最后,哈希匹配的速度较快,通过哈希匹配的方式实现URL分类,可以提高控制效率。

附图说明

[0020] 图1是本发明提供的基于DNS解析的网站访问控制方法一实施方式的流程图;

[0021] 图2是本发明中DNS报文一实施方式的格式示意图;

[0022] 图3是本发明中DNS报文解析得到的资源记录一实施方式的格式示意图。

具体实施方式

[0023] 为了使本发明的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本发明进行进一步详细说明。应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0024] 实施例1

[0025] 如图1所示,本发明的实施例1提供了基于DNS解析的网站访问控制方法,以下简称本方法,包括以下步骤:

[0026] S1、解析DNS报文,获取网站的URL地址;

[0027] S2、将所述URL地址与预设的URL库进行哈希匹配,得到URL分类结果;

[0028] S3、根据配置的URL控制策略,按所述URL分类结果对网站进行访问控制。

[0029] 本发明实施例首先通过解析DNS报文获取URL。DNS即域名系统(Domain Name System),是互联网的一项服务。它作为将域名和IP地址相互映射的一个分布式数据库,能够使人更方便地访问互联网。DNS允许用户终端设备将给定的人类可读URL转换为网络可以理解的机器可用IP地址(也可以将IP地址转换为相应的URL地址),这就是域名解析。

[0030] DNS报文格式如图2所示,其中,查询问题区域(Queries)中的资源记录区域(包括回答区域,授权区域和附加区域),对其进行解析后,得到如图3所示的资源记录,图3中域名(name)即URL地址。图2、图3中0、15、16、31表示字节标尺。

[0031] 解析得到URL地址后,将其与用户配置的URL库进行哈希匹配,主要是将DNS报文中获取到的URL地址的hash值,与用户配置的URL库进行匹配,获取匹配结果。Hash,一般翻译

做“散列”,也有直接音译为“哈希”的,就是把任意长度的输入(又叫做预映射,pre-image),通过散列算法,变换成固定长度的输出,该输出就是散列值。这种转换是一种压缩映射,也就是,散列值的空间通常远小于输入的空间,不同的输入可能会散列成相同的输出,而不可能从散列值来唯一的确定输入值。简单的说就是一种将任意长度的消息压缩到某一固定长度的消息摘要的函数。哈希表是根据设定的哈希函数H(key)将一组关键字映射到一个有限的地址区间上,并以关键字在地址区间中的象作为记录在表中的存储位置,这种表称为哈希表或散列,所得存储位置称为哈希地址或散列地址。哈希表作为一种线性数据结构,与表格结构和队列结构等相比,哈希表无疑是查找速度比较快的一种。

[0032] 将URL地址与URL库进行hash比对,得到URL地址的分类结果,例如是金融类、教育类等等,然后基于URL控制策略,依据分类结果判断进行阻断还是放通。

[0033] 本发明在域名解析时就进行限制,在DNS报文请求包中就把域名提取出来进行判断。由于DNS报文一般是UDP包,一个包中就包含了所有信息,不需要重组;其次,对于UDP包,各种资源的消耗都很少,客户端发UDP包的资源消耗远小于TCP,服务器根本就没消耗,防火墙跟踪UDP也比跟踪TCP要简单得多;再次,限制得可以更加全面,通常限制URL只限制了HTTP,而限制DNS则把该域名对应的所有服务都可以限制住;最后,DNS限制就没有IP访问的漏洞,因为本来就得不到IP。因此,基于DNS协议的URL过滤,在域名解析阶段进行控制,对http和https的url均有效,不需要解密以及报文重组,所以对性能影响小,同时对该域名对应的所有服务都可以控制。

[0034] 优选的,解析DNS报文,具体为:

[0035] 判断URL控制是否打开,如果未打开则不进行网站访问控制,如果打开,则解析DNS报文。

[0036] 在进行URL过滤之前,首先检查URL控制是否打开,如未打开,说明不需要进行URL过滤控制,此时则可直接进行报文放通,不会进行DNS解析过滤步骤;只有在URL控制打开,需要进行URL过滤控制时,才会执行后续DNS解析以及控制步骤。

[0037] 优选的,将所述URL地址与预设的URL库进行哈希匹配,得到URL分类结果之前,还包括:

[0038] 将所述URL地址与URL白名单进行匹配,如果白名单匹配成功,则直接放通DNS报文,如果白名单匹配不成功,则将所述URL地址与URL黑名单进行匹配,如果黑名单匹配成功,阻断DNS报文,如果黑名单匹配不成功,则判断是否配置URL控制策略,如果未配置,则放通DNS报文,如果配置,则将所述URL地址与预设的URL库进行哈希匹配,得到URL分类结果。

[0039] 本优选实施例中,增加了URL白名单/黑名单配置功能,其优先级高于URL控制策略,因此先进行白名单以及黑名单的匹配,再实施URL控制策略。

[0040] 具体的,首先按照URL白名单进行匹配,如果白名单匹配成功,说明该报文可放通,因此直接放通DNS报文。如白名单未匹配成功,则继续查看是否匹配黑名单,如果黑名单匹配,则说明配置了恶意URL,直接进行阻断,同时记录日志。如黑名单匹配不成功,即未匹配恶意URL,则判断是否配置了URL控制策略,如未配置,说明没有可执行的URL控制策略,则放通DNS报文,如配置有URL控制策略,则进行DNS报文解析获取URL,获取URL地址,进而与URL库进行哈希匹配,得到URL分类结果,最后根据URL控制策略,按URL分类结果进行访问控制。

[0041] 优选的,所述白名单匹配和黑名单匹配均采用哈希匹配。

[0042] 由于哈希匹配具有速度快的优点，因此本优选实施例中，白名单匹配以及黑名单匹配与URL库匹配一样，也采用哈希匹配的方式实现。

[0043] 优选的，本方法还包括，记录控制日志。

[0044] 按照URL控制策略执行相应策略动作时，对策略动作进行日志记录，便于后续对URL过滤过程进行追溯、分析等。

[0045] 优选的，所述URL库包括不同类型的URL；所述URL控制策略包括不同类型的URL及其对应的策略动作。

[0046] URL控制策略规定了哪一类URL对应哪一种动作，因此，通过DNS报文解析以及哈希匹配得到相应的URL的类别后，即按照URL控制策略为该类别URL配置的策略动作，对报文进行阻断或放通，同时记录日志，实现URL的过滤预警，实现访问控制。

[0047] 优选的，所述策略动作包括放通、阻断以及日志警告级别。

[0048] 本优选实施例中，URL控制策略包括需要控制的具体URL或者URL分类，及其对应的策略动作是放通还是阻断，及其对应的日志告警级别。

[0049] 实施例2

[0050] 本发明的实施例2提供了基于DNS解析的网站访问控制装置，包括处理器以及存储器，所述存储器上存储有计算机程序，所述计算机程序被所述处理器执行时，实现实施例1提供的基于DNS解析的网站访问控制方法。

[0051] 本发明实施例提供的基于DNS解析的网站访问控制装置，用于实现基于DNS解析的网站访问控制方法，因此，基于DNS解析的网站访问控制方法所具备的技术效果，基于DNS解析的网站访问控制装置同样具备，在此不再赘述。

[0052] 实施例3

[0053] 本发明的实施例3提供了计算机存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时，实现实施例1提供的基于DNS解析的网站访问控制方法。

[0054] 本发明实施例提供的计算机存储介质，用于实现基于DNS解析的网站访问控制方法，因此，基于DNS解析的网站访问控制方法所具备的技术效果，计算机存储介质同样具备，在此不再赘述。

[0055] 以上所述本发明的具体实施方式，并不构成对本发明保护范围的限定。任何根据本发明的技术构思所做出的各种其他相应的改变与变形，均应包含在本发明权利要求的保护范围内。

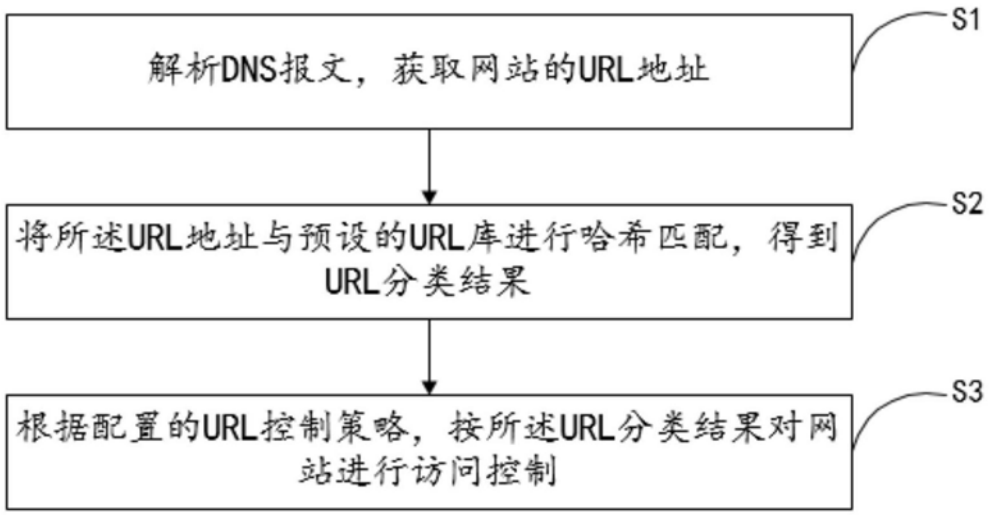

解析DNS报文，获取网站的URL地址
S1
将所述URL地址与预设的URL库进行哈希匹配，得到URL分类结果
S2
根据配置的URL控制策略，按所述URL分类结果对网站进行访问控制
S3

图1

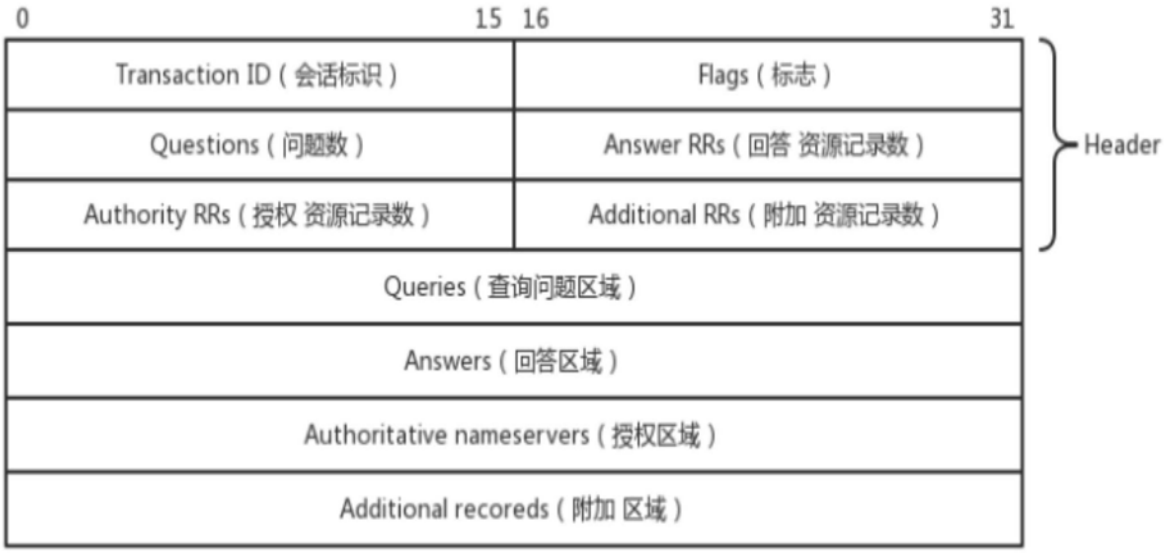

0
15
16
31
Transaction ID (会话标识)
Flags (标志)
Questions (问题数)
Answer RRs (回答 资源记录数)
Authority RRs (授权 资源记录数)
Additional RRs (附加 资源记录数)
Header
Queries (查询问题区域)
Answers (回答区域)
Authoritative nameservers (授权区域)
Additional recoreds (附加 区域)

图2

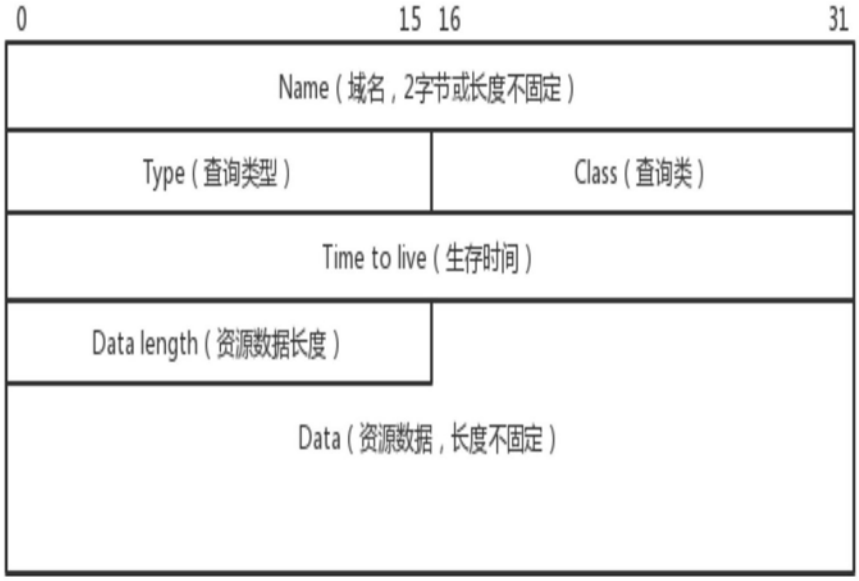

0
15 16
31
Name (域名 , 2字节或长度不固定)
Type (查询类型)
Class (查询类)
Time to live (生存时间)
Data length (资源数据长度)
Data (资源数据 , 长度不固定)

图3