



(12)发明专利

(10)授权公告号 CN 104657432 B

(45)授权公告日 2018.01.23

(21)申请号 201510043556.3

(22)申请日 2015.01.28

(65)同一申请的已公布的文献号

申请公布号 CN 104657432 A

(43)申请公布日 2015.05.27

(73)专利权人 北京金和软件股份有限公司

地址 100085 北京市海淀区上地东路1号院  
盈创动力大厦A座401室

(72)发明人 杨爱民 龙晓春

(74)专利代理机构 北京远大卓悦知识产权代理  
事务所(普通合伙) 11369

代理人 史霞

(51)Int.Cl.

G06F 17/30(2006.01)

(56)对比文件

CN 102917307 A,2013.02.06,

CN 102651714 A,2012.08.29,

CN 104104649 A,2014.10.15,

KR 20110132849 A,2011.12.09,

US 2009092064 A1,2009.04.09,

AU 2012100308 A4,2012.04.19,

xyz\_lmn.URL短地址压缩算法 微博短地址  
原理解析(Java实现).《CSDN博客》.2012,第1-5  
页.

审查员 单娟

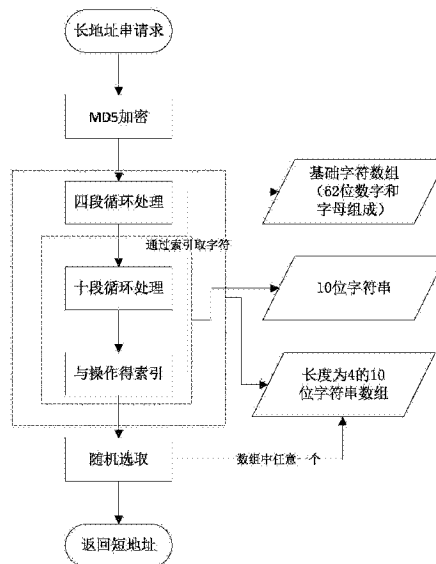
权利要求书1页 说明书3页 附图1页

(54)发明名称

一种降低长地址转换为短地址重复率的方法

(57)摘要

本发明提供一种降低长地址转换为短地址重复率的方法,包括:将长地址通过加密算法生成32位字符串,将所述32位字符串分成四段,每段包括8位字符串;将每段中所述8位字符串均视作16进制串,并转换为32位二进制数字,32位二进制数字全部参与运算;将所述32位二进制数字做十次循环处理,获得一个10位字符串;对所述四段的8位字符串分别十次循环处理,获得4个10位字符串;所述4个10位字符串,组合成一个长度为4的字符串数组,随机取一个短地址即可。本发明的有益效果是提供了一种长地址转换为短地址的方法,该方法可大大降低传统长地址转短地址算法的重复率,在一定时间范围内,重复短地址数/总请求数的比值从1.6%降低至 $5.024e^{-5}$ 。



1. 一种降低长地址转换为短地址重复率的方法,其中,包括:

将长地址通过加密算法生成32位字符串,将所述32位字符串分成四段,每段包括8位字符串;

将每段中所述8位字符串均视作16进制串,并转换为32位二进制数字,32位二进制数字全部参与运算;

将所述32位二进制数字做十次循环处理,获得一个10位字符串;

对所述四段的8位字符串分别十次循环处理,获得4个10位字符串,其中,对每段中所述8位字符串进行十次循环处理具体步骤为:

步骤一、所述32位二进制数字与参数进行与运算,得到第一数,把所述第一数当做索引在基础字符数组中得到对应的第一位字符;将所述32位二进制数字右移三位得到第一个新数字,完成第一次循环;

步骤二、将所述步骤一中的新数字与所述参数进行与运算,得到第二数,把所述第二数当做索引在基础字符数组中得到对应的第二位字符;将所述第一个新数字右移三位得到第二个新数字完成第二次循环;

步骤三、按照所述步骤二的计算方法循环十次,即可获得所述一个10位字符串;

所述4个10位字符串,组合成一个长度为4的字符串数组,随机取一个短地址即可;

其中,还包括预先给出构成短地址的基础字符数组。

2. 如权利要求1所述的降低长地址转换为短地址重复率的方法,其中,所述加密算法为MD5算法。

3. 如权利要求1所述的降低长地址转换为短地址重复率的方法,其中,所述长地址为请求的整个地址。

4. 如权利要求3所述的降低长地址转换为短地址重复率的方法,其中,所述参数为0x0000003D。

5. 如权利要求1所述的降低长地址转换为短地址重复率的方法,其中,所述基础字符数组的长度为62,所述基础字符数组包括0至9的数字与大小写字母。

## 一种降低长地址转换为短地址重复率的方法

### 技术领域

[0001] 本发明涉及长地址转换为短地址的方法。更具体地说,本发明涉及一种降低长地址转换为短地址重复率的方法。

### 背景技术

[0002] 短地址的兴起源于微博,由于主流微博140字限制,所以短地址渐渐流行起来,且应用范围越来越广,这是因为地址本身对用户来说没有任何意义,用户只关心它指向的内容,所以地址越简单越好。

[0003] 短地址好处很多,例如它可以缩短链接长度,使得手机、微博分享比较方便等,但正是短地址需要跳转到原来长地址,也导致短地址存在着缺点,即无法查看原来的长地址。

[0004] 传统的长地址转换为短地址的算法为:1) 将长地址MD5生成32位签名串,分为4段,每段8个字节;2) 对这四段循环处理,取8个字节,将他看成 16进制串与0x3fffffff (30位1) 与操作,即超过30位的忽略处理;3) 这30位分成6段,每5位的数字作为字母表的索引取得特定字符,依次进行获得6 位字符串;4) 总的MD5串可以获得4个6位串;取里面的任意一个就可作为这个长地址的短地址。

[0005] 在大并发大数据量的情况下,传统的长地址转换为短地址算法会产生重复的数据,即不同的长地址生成了相同的短地址,导致通过短地址不能正确的定位到原始长地址。

### 发明内容

[0006] 本发明针对上述问题,提供了一种降低长地址转换为短地址重复率的方法,该方法可大大降低传统长地址转换为短地址算法的重复率,在一定时间范围内,重复短地址数/总请求数的比值从1.6%降低至 $5.024e^{-5}$

[0007] 本发明的一个目的是研究了长地址转换为短地址的方法。

[0008] 本发明还有一个目的是该方法解决了传统长地址转换为短地址时重复率较高的问题,在一定的时间范围内,使得重复短地址数/总请求数的比值由传统算法的1.6%降低为 $5.024e^{-5}$ 。

[0009] 本发明还有一个目的是通过对传统算法的步骤2和步骤3进行改进,大大降低了长地址转换为短地址的重复率。

[0010] 为此,本发明提供了一种降低长地址转换为短地址重复率的方法,包括:

[0011] 将长地址通过加密算法生成32位字符串,将所述32位字符串分成四段,每段包括8位字符串;

[0012] 将每段中所述8位字符串均视作16进制串,并转换为32位二进制数字, 32位二进制数字全部参与运算;

[0013] 将所述32位二进制数字做十次循环处理,获得一个10位字符串;

[0014] 对所述四段的8位字符串分别十次循环处理,获得4个10位字符串;

[0015] 所述4个10位字符串,组合成一个长度为4的字符串数组,随机取一个短地址即可。

- [0016] 优选的是,其中,还包括预先给出构成短地址的基础字符数组。
- [0017] 优选的是,其中,所述加密算法为MD5算法。
- [0018] 优选的是,其中,所述长地址为请求的整个地址。
- [0019] 优选的是,其中,对每段中所述8位字符串进行十次循环处理具体步骤为:
- [0020] 步骤一、所述32位二进制数字与参数进行与运算,得到第一数,把所述第一数当做索引在基础字符数组中得到对应的第一位字符;将所述32位二进制数字右移三位得到第一个新数字,完成第一次循环;
- [0021] 步骤二、将所述步骤一中的新数字与所述参数进行与运算,得到第二数,把所述第二数当做索引在基础字符数组中得到对应的第二位字符;将所述第一个新数字右移三位得到第二个新数字完成第二次循环;
- [0022] 步骤三,按照所述步骤二的计算方法循环十次,即可获得所述一个10位字符串。
- [0023] 优选的是,其中,所述参数为0x0000003D。
- [0024] 优选的是,其中,所述基础字符数组的长度为62,所述基础字符数组包括0至9的数字与大小写字母。
- [0025] 本发明至少包括以下有益效果:本发明提供了一种降低长地址转换为短地址的重复率的方法,该方法中对传统长地址转换为短地址方法中的步骤2 和步骤3进行了改进,传统方法经过步骤2得到30位(即超过30位的忽略),然后进行步骤3的操作,而本发明的步骤2保留全部,并参与步骤3的运算;本发明将请求传递的长地址经过MD5加密算法生成32位字符串,32位字符串经过十次内循环和四次外循环得到长度为4的10位字符串,通过本发明的方法将一定的时间范围内的重复短地址数/总请求数的比值由传统算法的 1.6%降低为  $5.024e^{-5}$ ,解决了传统长地址转换为短地址时重复率较高的问题,大大降低了长地址转换为短地址的重复率。
- [0026] 本发明的其它优点、目标和特征将部分通过下面的说明体现,部分还将通过对本发明的研究和实践而为本领域的技术人员所理解。

## 附图说明

- [0027] 图1为本发明的一个实施例中降低长地址转换为短地址重复率方法的流程图。

## 具体实施方式

- [0028] 下面结合附图对本发明做进一步的详细说明,以令本领域技术人员参照说明书文字能够据以实施。
- [0029] 图1示出了根据本发明的一种降低长地址转换为短地址重复率的方法实现形式,其中包括:对长地址经MD5算法生成的32位字符串进行10次内循环和4次外循环生成四个不同的10位短地址,具体为:
- [0030] (1) 将长地址通过加密算法生成32位字符串,将32位字符串分成四段,每段包括8位字符串;例如长地址为: <http://XXXX.XXXX.com/app/GetAppDetail?appId=d8a701a0-067b-4e3a-abba-6d0bb40214ca>,经过加密算法生成32位字符串,例如所述32位字符串为数字和字母。
- [0031] (2) 将每段中包含的8位字符串均视作16进制串,并将16进制的8位字符串转换为

32位二进制数字,32位二进制数字全部参与下一步的运算;

[0032] (3)将(2)中的32位二进制数字做十次循环处理,获得一个10位字符串;例如生成的短地址为:7mFnr6IZJf;

[0033] (4)对(1)中四段的8位字符串分别十次循环处理,获得4个10位字符串;

[0034] (5)所述4个10位字符串,组合成一个长度为4的字符串数组,随机取一个短地址即可。

[0035] 在另一种实例中,还包括预先给出构成短地址的基础字符数组。

[0036] 在另一种实例中,经过加密算法生成32位字符串的加密算法为MD5算法。

[0037] 在另一种实例中,(1)中的长地址为请求的整个长地址。

[0038] 在另一种实例中,对每段中包含的16进制的8位字符串进行十次循环处理具体步骤为:

[0039] 步骤一、所述32位二进制数字与参数进行与运算,得到第一数,把第一数当做索引在基础字符数组中得到对应的第一位字符;将所述32位二进制数字右移三位得到第一个新数字,完成第一次循环;

[0040] 步骤二、将所述步骤一中的新数字与所述参数进行与运算,得到第二数,把所述第二数当做索引在基础字符数组中得到对应的第二位字符;将所述第一个新数字右移三位得到第二个新数字完成第二次循环;

[0041] 步骤三,按照所述步骤二的计算方法循环十次,即可获得所述一个10位字符串。

[0042] 在另一种实例中,所述参数为0x0000003D。

[0043] 在另一种实例中,构成短地址的基础字符数组的个数为62,包括0至9 的数字与A-Z,a-z。

[0044] 尽管本发明的实施方案已公开如上,但其并不仅仅限于说明书和实施方式中所列运用。它完全可以被适用于各种适合本发明的领域。对于熟悉本领域的人员而言,可容易地实现另外的修改。因此在不背离权利要求及等同范围所限定的一般概念下,本发明并不限于特定的细节和这里示出与描述的图例。

