(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 104394144 A
(43)申请公布日 2015.03.04

(21)申请号 201410683018.6

(22)申请日 2014.11.24

(71)申请人 蔡志明
地址 518035 广东省深圳市福田区笋岗西路3002号深圳市第二人民医院银华大厦6楼611室
申请人 全筱筱 潘军杰 熊文举 郭岱琦

(72)发明人 蔡志明 全筱筱 潘军杰 熊文举 郭岱琦

(74)专利代理机构 深圳市合道英联专利事务所（普通合伙） 44309
代理人 廉红果

(51)Int.Cl.
*H04L 29/06*(2006.01)
*H04L 1/00*(2006.01)

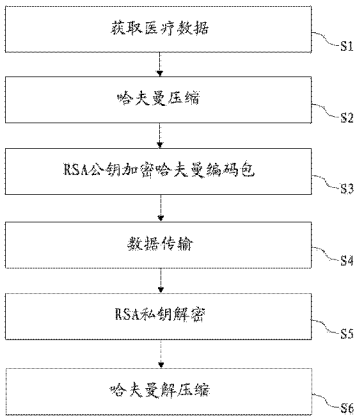
权利要求书2页 说明书4页 附图1页

(54)发明名称

一种云存储医疗数据的安全传输方法

(57)摘要

本发明公开了一种云存储医疗数据的安全传输方法，包括以下步骤：获取医疗数据；对医疗数据进行哈夫曼压缩，生成压缩数据包和哈夫曼编码包；采用RSA算法的公钥对哈夫曼编码包进行加密；将压缩数据包以及加密后的哈夫曼编码包传输到云端服务器；采用RSA算法的私钥对加密后的哈夫曼编码包进行解密；基于哈夫曼编码包，对压缩数据包进行解压缩。本发明减少了数据传输过程中对网络资源的占用，在确保加密和解密的效率前提下，提高了数据传输的安全性。




CN 104394144 A

1. 一种云存储医疗数据的安全传输方法，其特征在于，包括以下步骤：

S1、获取医疗数据；

S2、对医疗数据进行哈夫曼压缩，生成压缩数据包和哈夫曼编码包；

S3、采用 RSA 算法的公钥对哈夫曼编码包进行加密；

S4、将压缩数据包以及加密后的哈夫曼编码包传输到云端服务器；

S5、采用 RSA 算法的私钥对加密后的哈夫曼编码包进行解密；

S6、基于哈夫曼编码包，对压缩数据包进行解压缩。

2. 如权利要求 1 所述的一种云存储医疗数据的安全传输方法，其特征在于，所述公钥和私钥通过以下方法产生：

(1) 取两个互异的大质数 p、q；

(2) 计算 n = p×q；

(3) 随机选取整数 e，且 e 与 (p-1)×(q-1) 互为素数；

(4) 选择数 d，使其满足 (e×d)mod[(p-1)×(q-1)] = 1，从而确定 (n，e) 为公钥，(n，d) 为私钥。

3. 如权利要求 2 所述的一种云存储医疗数据的安全传输方法，其特征在于：所述加密的算法为 C = Me mod n，所述解密的算法为 M = Cd mod n，其中，C 为加密后的哈夫曼编码包，M 为加密前的哈夫曼编码包。

4. 如权利要求 1-3 任一项所述的一种云存储医疗数据的安全传输方法，其特征在于，在步骤 S2 中，所述哈夫曼压缩通过以下方法实现：

(1) 统计医疗数据文件中每种字符出现的概率 $P(S_i)$，i = 1,2,3,...,q，其中，q 为医疗数据中含有的种类数，$S_i$ 为第 i 种字符；

(2) 将 q 种字符作为 q 个信源符号，按出现的概率大小递减排列；

(3) 用字符‘0’和‘1’分别代表概率最小的 2 个信源符号，并将这 2 个概率最小的信源符号合并成 1 个信源符号，从而得到只包含 q-1 个符号的新信源，称为缩减信源 S1；

(4) 把缩减信源 S1 的符号仍按概率大小递减次序排列，再将其最后两个概率最小的信源符号分别用字符‘0’和‘1’表示，并且合并成一个符号，这样又形成了 q-2 个信源符号的缩减信源 S2；

(5) 依次继续下去，直至信源最后只剩下两个信源符号为止，将这最后两个信源符号分别用字符‘0’和‘1’表示；

(6) 从最后一级缩减信源开始，进行回推得到每种字符所对应的由字符‘0’和‘1’组成的字符串序列，作为伪码字；

(7) 基于每种字符对应的伪码字，建立一个映射，$f(S_i) = c_i$，i = 1,2,3,...,q，其中，$S_i$ 代表不同的字符，$c_i$ 代表与字符 $S_i$ 对应的伪码字；

(8) 建立另一个映射，$g(c_i) = \omega_i$，i = 1,2,3,...,q，其中，$c_i$ 代表与字符 $S_i$ 对应的伪码字，$\omega_i$ 代表与伪码字 $c_i$ 对应的码字。利用该映射将每个伪码字转换成对应的二进制的码字，从而获得哈夫曼编码表，即生成了所述哈夫曼编码包；

(9) 对于医疗数据文件中的每个字符，找到在哈夫曼编码表中与其对应的码字，并用该码字对其进行替换，遍历医疗数据文件中的所有字符，从而完成了对医疗数据文件的压缩，即生成了所述压缩数据包。

5. 如权利要求 4 所述的一种云存储医疗数据的安全传输方法，其特征在于，在步骤 S6 中，所述解压缩通过以下方法实现：对于压缩数据包中的每个码字，找到在哈夫曼编码表中与其对应的字符，并用该字符对其进行替换，遍历压缩数据包中的所有码字，即完成了对压缩数据包中的解压缩。

6. 如权利要求 1 所述的一种云存储医疗数据的安全传输方法，其特征在于：所述医疗数据包括医疗影像数据、检验数据以及患者信息数据。

# 一种云存储医疗数据的安全传输方法

## 技术领域

[0001] 本发明涉及数据传输技术领域，特别涉及一种云存储医疗数据的安全传输方法。

## 背景技术

[0002] 随着计算机和通讯技术发展，为数字化影像和传输奠定基础，实现彻底无胶片放射科和数字化医院，已经成为现代化医疗不可阻挡的潮流。目前国内众多医院已完成医院信息化管理，其影像设备逐渐更新为数字化，网络互联和综合图像信息数据库系统建设变得尤为重要，针对海量医疗数据的管理和存储，更是成为重中之重。

[0003] 随着云计算技术的飞速发展，为构建低成本、高可用、高性能的区域医学图像存储协作平台提供了一条有效的途径。通过高性能、大容量云存储系统，为无法单独购买大容量存储设备的医院提供方便快捷的空间服务，满足不断增加的海量医学数据存储和管理服务。

[0004] 为了保障医疗数据的安全性，医院与云存储系统进行数据传输时，通常需要对医疗数据进行加密处理。

[0005] 目前通用的加密算法主要分为对称算法和非对称算法。对称算法采用相同的密钥进行加密和解密，其最大的困难是密钥分发问题，必须通过当面或在公共传送系统中使用安全的方法交换密钥。对称加密由于加密速度快、硬件容易实现，因此仍被广泛用来加密各种信息。但对称加密也存在着固有的缺点：密钥更换困难，经常使用同一密钥进行数据加密，给攻击者提供了攻击密钥的信息和时间，安全性差。非对称算法，采用公钥进行加密而利用私钥进行解密，主要包括 RSA、DH、EC、DSS 等。公钥是可以公开的，任何人都可以获得，数据发送人用公钥将数据加密后再传给数据接收人，接收人用自己的私钥解密。非对称加密的安全性主要依赖难解的数学问题，密钥的长度比对称加密大得多。因此非对称加密算法虽然具有很高的安全性，但其加密效率较低，不适合大量数据的加密传输。

## 发明内容

[0006] 本发明的目的在于提供一种云存储医疗数据的安全传输方法，其减少了数据传输过程中对网络资源的占用，在确保加密和解密的效率前提下，提高了数据传输的安全性。

[0007] 为实现上述目的，本发明采用以下技术方案：

[0008] 一种云存储医疗数据的安全传输方法，包括以下步骤：

[0009] S1、获取医疗数据；

[0010] S2、对医疗数据进行哈夫曼压缩，生成压缩数据包和哈夫曼编码包；

[0011] S3、采用 RSA 算法的公钥对哈夫曼编码包进行加密；

[0012] S4、将压缩数据包以及加密后的哈夫曼编码包传输到云端服务器；

[0013] S5、采用 RSA 算法的私钥对加密后的哈夫曼编码包进行解密；

[0014] S6、基于哈夫曼编码包，对压缩数据包进行解压缩。

[0015] 进一步地，所述公钥和私钥通过以下方法产生：

[0016] (1) 取两个互异的大质数 p、q ;

[0017] (2) 计算 n = p×q ;

[0018] (3) 随机选取整数 e,且 e 与 (p-1)×(q-1) 互为素数 ;

[0019] (4) 选择数 d,使其满足 (e×d)mod[(p-1)×(q-1)] = 1,从而确定 (n,e) 为公钥,(n,d) 为私钥。

[0020] 更进一步地,所述加密的算法为 C = Me mod n,所述解密的算法为 M = Cdmod n,其中,C 为加密后的哈夫曼编码包,M 为加密前的哈夫曼编码包。

[0021] 优选地,在步骤 S2 中,所述哈夫曼压缩通过以下方法实现 :

[0022] (1) 统计医疗数据文件中每种字符出现的概率 $P(S_i)$,i = 1,2,3,...,q,其中,q 为医疗数据中含有的种类数,$S_i$ 为第 i 种字符 ;

[0023] (2) 将 q 种字符作为 q 个信源符号,按出现的概率大小递减排列 ;

[0024] (3) 用字符‘0’和‘1’分别代表概率最小的 2 个信源符号,并将这 2 个概率最小的信源符号合并成 1 个信源符号,从而得到只包含 q-1 个符号的新信源,称为缩减信源 S1 ;

[0025] (4) 把缩减信源 S1 的符号仍按概率大小递减次序排列,再将其最后两个概率最小的信源符号分别用字符‘0’和‘1’表示,并且合并成一个符号,这样又形成了 q-2 个信源符号的缩减信源 S2 ;

[0026] (5) 依次继续下去,直至信源最后只剩下两个信源符号为止,将这最后两个信源符号分别用字符‘0’和‘1’表示 ;

[0027] (6) 从最后一级缩减信源开始,进行回推得到每种字符所对应的由字符‘0’和‘1’组成的字符串序列,作为伪码字 ;

[0028] (7) 基于每种字符对应的伪码字,建立一个映射,$f(S_i) = c_i$,i = 1,2,3,...,q,其中,$S_i$ 代表不同的字符,$c_i$ 代表与字符 $S_i$ 对应的伪码字 ;

[0029] (8) 建立另一个映射,$g(c_i) = \omega_i$,i = 1,2,3,...,q,其中,$c_i$ 代表与字符 $S_i$ 对应的伪码字,$\omega_i$ 代表与伪码字 $c_i$ 对应的码字。利用该映射将每个伪码字转换成对应的二进制的码字,从而获得哈夫曼编码表,即生成了所述哈夫曼编码包 ;

[0030] (9) 对于医疗数据文件中的每个字符,找到在哈夫曼编码表中与其对应的码字,并用该码字对其进行替换,遍历医疗数据文件中的所有字符,从而完成了对医疗数据文件的压缩,即生成了所述压缩数据包。

[0031] 优选地,在步骤 S6 中,所述解压缩通过以下方法实现 :对于压缩数据包中的每个码字,找到在哈夫曼编码表中与其对应的字符,并用该字符对其进行替换,遍历压缩数据包中的所有码字,即完成了对压缩数据包中的解压缩。

[0032] 优选地,所述医疗数据包括医疗影像数据、检验数据以及患者信息数据。

[0033] 采用上述技术方案后,本发明与背景技术相比,具有如下优点 :

[0034] 1. 通过对医疗数据进行哈夫曼压缩,减少了数据传输过程中对网络资源的占用。

[0035] 2. 由于只对哈夫曼编码树进行加密,加密和解密需要处理的数据量将大大减少,克服了不对称加密算法难以处理大量数据的限制,在确保加密和解密的效率前提下,提高了数据传输的安全性。

## 附图说明

[0036] 图 1 为本发明的工作流程示意图。

## 具体实施方式

[0037] 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。

[0038] 实施例

[0039] 请参阅图 1，本发明公开了一种云存储医疗数据的安全传输方法，包括以下步骤：

[0040] S1、获取医疗数据

[0041] 医疗数据包括医疗影像数据、检验数据、患者信息数据等医疗数据，医疗数据的可利用医院现有的 PACS 系统、检验管理系统、HIS 系统等医疗信息系统进行获取。

[0042] S2、哈夫曼压缩

[0043] 对医疗数据进行哈夫曼压缩，生成压缩数据包和哈夫曼编码包。该步骤具体通过以下方法实现：

[0044] (1) 统计医疗数据文件中每种字符出现的概率 $P(S_i)$，$i = 1,2,3,\ldots,q$，其中，q 为医疗数据中含有的种类数，$S_i$ 为第 i 种字符；

[0045] (2) 将 q 种字符作为 q 个信源符号，按出现的概率大小递减排列；

[0046] (3) 用字符‘0’和‘1’分别代表概率最小的 2 个信源符号，并将这 2 个概率最小的信源符号合并成 1 个信源符号，从而得到只包含 q-1 个符号的新信源，称为缩减信源 S1；

[0047] (4) 把缩减信源 S1 的符号仍按概率大小递减次序排列，再将其最后两个概率最小的信源符号分别用字符‘0’和‘1’表示，并且合并成一个符号，这样又形成了 q-2 个信源符号的缩减信源 S2；

[0048] (5) 依次继续下去，直至信源最后只剩下两个信源符号为止，将这最后两个信源符号分别用字符‘0’和‘1’表示；

[0049] (6) 从最后一级缩减信源开始，进行回推得到每种字符所对应的由字符‘0’和‘1’组成的字符串序列，作为伪码字；

[0050] (7) 基于每种字符对应的伪码字，建立一个映射，$f(S_i) = c_i$，$i = 1,2,3,\ldots,q$，其中，$S_i$ 代表不同的字符，$c_i$ 代表与字符 $S_i$ 对应的伪码字；

[0051] (8) 建立另一个映射，$g(c_i) = \omega_i$，$i = 1,2,3,\ldots,q$，其中，$c_i$ 代表与字符 $S_i$ 对应的伪码字，$\omega_i$ 代表与伪码字 $c_i$ 对应的码字。利用该映射将每个伪码字转换成对应的二进制的码字，从而获得哈夫曼编码表，即生成了哈夫曼编码包；

[0052] (9) 对于医疗数据文件中的每个字符，找到在哈夫曼编码表中与其对应的码字，并用该码字对其进行替换，遍历医疗数据文件中的所有字符，从而完成了对医疗数据文件的压缩，即生成了压缩数据包。

[0053] S3、RSA 公钥加密哈夫曼编码包

[0054] 采用 RSA 算法的公钥对哈夫曼编码包进行加密，加密算法为 $C = Me \bmod n$，其中，C 为加密后的哈夫曼编码包，M 为加密前的哈夫曼编码包。

[0055] 本领域技术人员应该理解的是，RSA 算法的公钥和私钥都是预先生成的，其具体通过以下方法产生：

[0056] (1) 取两个互异的大质数 p、q ;

[0057] (2) 计算 n = p×q ;

[0058] (3) 随机选取整数 e,且 e 与 (p-1)×(q-1) 互为素数 ;

[0059] (4) 选择数 d,使其满足 (e×d)mod[(p-1)×(q-1)] = 1,从而确定 (n,e) 为公钥,(n,d) 为私钥。

[0060] S4、数据传输

[0061] 将压缩数据包以及加密后的哈夫曼编码包传输到云端服务器。由于医疗数据是以压缩包的形式进行数据传输,从而减少了数据传输过程中对网络资源的占用,提高了数据传输的效率。

[0062] S5、RSA 私钥解密

[0063] 采用 RSA 算法的私钥对加密后的哈夫曼编码包进行解密,解密算法为 M = Cd mod n,其中,C 为加密后的哈夫曼编码包,M 为加密前的哈夫曼编码包。

[0064] S6、哈夫曼解压缩

[0065] 基于解密后的哈夫曼编码包,对压缩数据包进行解压缩。解压缩的具体方法为 :对于压缩数据包中的每个码字,找到在哈夫曼编码表中与其对应的字符,并用该字符对其进行替换,遍历压缩数据包中的所有码字,即完成了对压缩数据包中的解压缩,获得医疗数据。

[0066] 通过哈夫曼压缩、解压缩的具体过程可以看出,经过压缩的医疗数据必须使用压缩过程中形成的哈夫曼编码包(即哈夫曼编码树)才能解压缩。对于不同的医疗数据文件,由于文件内容的不同,形成的哈夫曼编码包也不同。数据传输的过程中需要同时传输压缩数据包和相应的哈夫曼编码包。相对于压缩数据包,哈夫曼编码树的节点数大大小于数据文件的数据量,如果只对哈夫曼编码树进行加密,加密和解密需要处理的数据量将大大减少,对于不对称加密算法无法处理大量数据的限制也可被克服。在数据传输中需要传输的数据量比压缩之前需要传输的数据量大大降低,可以节省大量的网络资源。在大规模的数据安全传输中,可以提高数据传输的效率和安全性。

[0067] 以上所述,仅为本发明较佳的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到的变化或替换,都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应该以权利要求的保护范围为准。

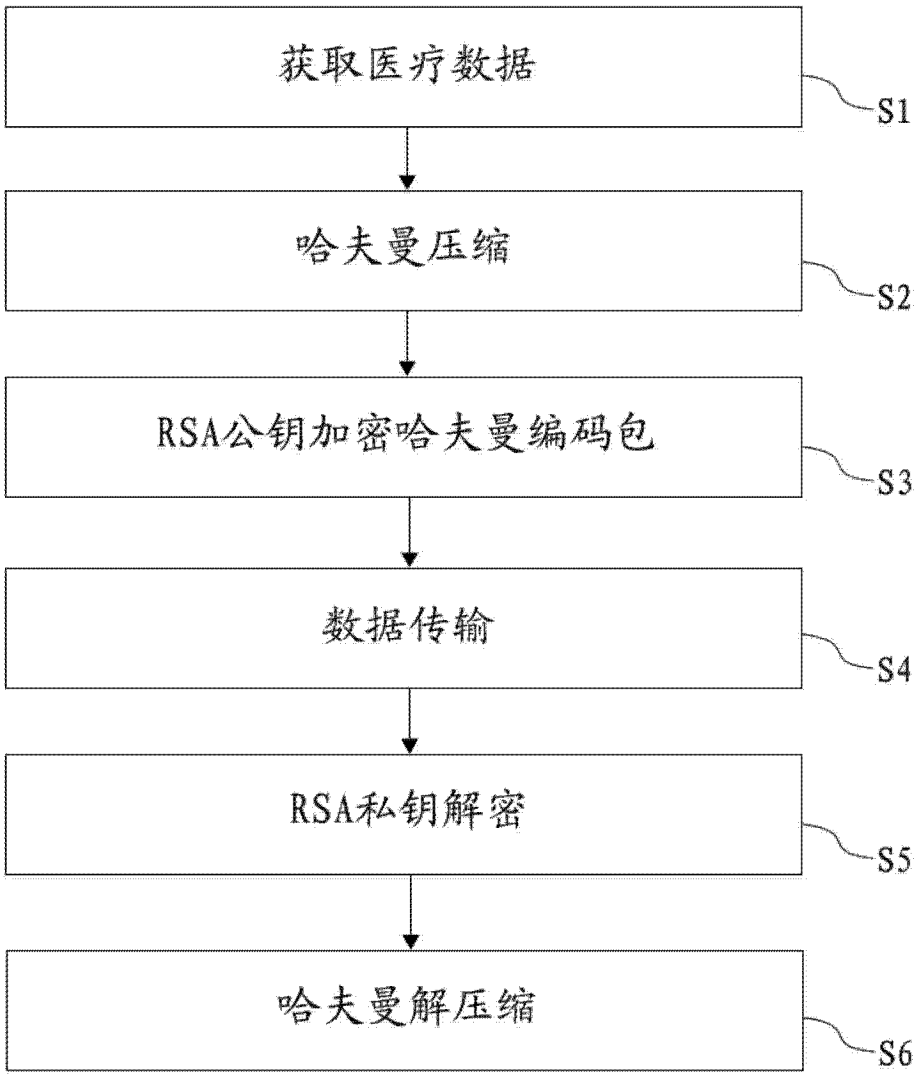

获取医疗数据
S1
哈夫曼压缩
S2
RSA公钥加密哈夫曼编码包
S3
数据传输
S4
RSA私钥解密
S5
哈夫曼解压缩
S6

图 1