



(12) 发明专利申请

(10) 申请公布号 CN 103955537 A

(43) 申请公布日 2014. 07. 30

(21) 申请号 201410207365. 1

H04L 29/06(2006. 01)

(22) 申请日 2014. 05. 16

H04L 9/00(2006. 01)

(71) 申请人 福州大学

地址 350108 福建省福州市闽侯县上街镇大学城学园路 2 号福州大学新区

(72) 发明人 吴阳 林柏钢 杨旸 钟玲  
陈何峰 王淑娥 李宇翔

(74) 专利代理机构 福州元创专利商标代理有限公司 35100

代理人 蔡学俊

(51) Int. Cl.

G06F 17/30(2006. 01)

H04L 29/08(2006. 01)

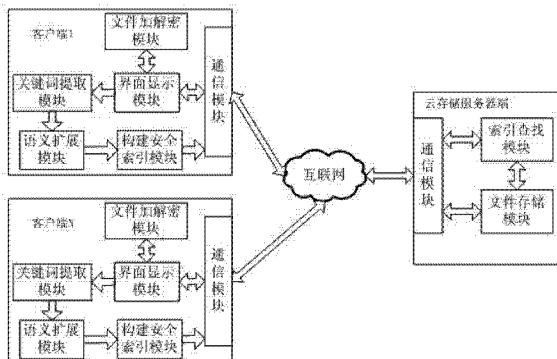
权利要求书1页 说明书4页 附图2页

(54) 发明名称

一种语义模糊可搜索加密云盘设计方法及系统

(57) 摘要

本发明涉及一种语义模糊可搜索加密云盘设计方法及系统。所述方法：首先，通过客户端，输入密钥对用户需上传的文件进行加密；提取所述加密文件的关键词，并进行语义扩展构建同义词集合；然后，通过布隆过滤器对上述构建的同义词集合进行安全索引的构建，并将该安全索引上传至云存储服务器端；客户端根据用户输入密钥生成关键词陷门，并将该陷门提交给云存储服务器端；最后，云存储服务器端根据用户提交的关键词陷门，通过安全索引查找相应文件，并将该文件返回给客户端；客户端通过密钥对云存储服务器端返回的文件进行解密。本发明主要为密文文档的信息检索提供同义词语义模糊搜索功能，并且支持多关键词检索，对文档的更新操作具有较高的效率。



1. 一种语义模糊可搜索加密云盘设计方法,其特征在于:包括如下步骤,

步骤 S01 :通过客户端,输入密钥对用户需上传的文件进行加密;

步骤 S02 :提取上述加密文件的关键词,并进行语义扩展构建同义词集合;

步骤 S03 :通过布隆过滤器对上述构建的同义词集合进行安全索引的构建,并将该安全索引上传至云存储服务器端;

步骤 S04 :客户端根据用户输入密钥生成关键词陷门,并将该陷门提交给云存储服务器端;

步骤 S05 :云存储服务器端根据用户提交的关键词陷门,通过安全索引查找相应文件,并将该文件返回给客户端;

步骤 S06 :客户端通过密钥对云存储服务器端返回的文件进行解密。

2. 根据权利要求 1 所述的一种语义模糊可搜索加密云盘设计方法,其特征在于:所述云存储服务器端根据用户提交的更新请求,修改所述安全索引即可完成文件更新。

3. 根据权利要求 1 所述的一种语义模糊可搜索加密云盘设计方法,其特征在于:所述步骤 S01 的文件加密是通过对称加密算法对文件进行加密。

4. 根据权利要求 3 所述的一种语义模糊可搜索加密云盘设计方法,其特征在于:所述对称加密算法为 AES、DES 或 3DES 加密算法。

5. 一种语义模糊可搜索加密云盘系统,其特征在于:包括一客户端和一云存储服务器端;所述客户端包括一实现文件加解密的文件加解密模块、一用于提取加密文件关键词的关键词提取模块、一用于对提取的关键词进行语义扩展并构建同义词集合的语义扩展模块、一通过布隆过滤器对所述同义词集合进行安全索引构建的构建安全索引模块和一上传安全索引、提交陷门以及上传 / 下载文件的第一通信模块;所述云存储服务器端包括一根据陷门查找安全索引的索引查找模块、一存储客户端上传加密文件的文件存储模块和一发送加密文件至所述客户端并与所述第一通信模块通信的第二通信模块。

6. 根据权利要求 5 所述的一种语义模糊可搜索加密云盘系统,其特征在于:所述文件加解密模块通过对称加密算法对文件进行加密。

7. 根据权利要求 6 所述的一种语义模糊可搜索加密云盘系统,其特征在于:所述对称加密算法为 AES、DES 或 3DES 加密算法。

8. 根据权利要求 5 所述的一种语义模糊可搜索加密云盘系统,其特征在于:所述客户端还包括一界面显示模块,该界面显示模块用于选取加密文件、显示搜索结果以及输入检索关键词。

9. 根据权利要求 5 或 8 所述的一种语义模糊可搜索加密云盘系统,其特征在于:所述语义扩展模块采用 WordNet 对所述同义词集合进行语义扩展。

## 一种语义模糊可搜索加密云盘设计方法及系统

### 技术领域

[0001] 本发明主要应用于可搜索加密云盘领域,特别是一种语义模糊可搜索加密云盘设计方法及系统。

### 背景技术

[0002] 随着云计算的发展,传统计算机的计算和存储功能转移到云服务器中,有效的减少用户的计算开销和存储成本。360云盘,百度云盘等网络存储服务都是直接将文档存储在云存储服务器中,这样云服务器管理员和外部攻击者都可以直接或间接的接触到用户的数据,特别对于一些敏感数据,用户将数据存储在云服务器中后就失去了对数据的直接控制力,可能导致用户个人隐私数据的泄露和滥用。

[0003] 将数据加密后存储到云存储服务器中,无论是云服务器管理员还是外部攻击者都无法获取数据真实内容,保护了用户的隐私性。但是,这给密文数据的信息检索带来了极大的挑战,可搜索加密技术用于解决密文检索问题。可搜索加密技术最初提出的时候,对文档中的每个单词采用了两层加密的结构,但是全文搜索的效率较低,并且不支持多关键词搜索。相关文献 :Song D X, Wagner D, Perrig A. Practical techniques for searches on encrypted data[C]//Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on. IEEE, 2000: 44-55. 而目前的多关键的检索索引采用向量模型并且通过矩阵加密,这样生成陷门和查询效率较低,并且只支持精确关键词搜索。相关文献 :Cao N, Wang C, Li M, et al. Privacy-preserving multi-keyword ranked search over encrypted cloud data[J]. Parallel and Distributed Systems, IEEE Transactions on, 2014, 25(1): 222-233. Li R, Xu Z, Kang W, et al. Efficient multi-keyword ranked query over encrypted data in cloud computing[J]. Future Generation Computer Systems, 2014, 30: 179-190. 目前提出的模糊搜索方案主要是针对拼写错误的模糊搜索,并不支持语义模糊搜索。相关文献 :Li J, Wang Q, Wang C, et al. Fuzzy keyword search over encrypted data in cloud computing[C]//INFOCOM, 2010 Proceedings IEEE. IEEE, 2010: 1-5. Liu C, Zhu L, Li L, et al. Fuzzy keyword search on encrypted cloud storage data with small index[C]//Cloud Computing and Intelligence Systems (CCIS), 2011 IEEE International Conference on. IEEE, 2011: 269-273. 本发明采用布隆过滤器作为索引结构,具有较高的更新和查找效率,并且支持多关键词检索,对关键词进行同义词扩展后再构建安全索引以完成语义模糊搜索请求。

[0004] 由于用户描述同一事物所用的词汇不同,若不能进行同义词搜索,将严重影响系统的使用性。

### 发明内容

[0005] 本发明的目的在于提供一种解决现有可搜索加密技术无法完成语义模糊搜索,更

新效率低下,多关键词查找效率低下等不足的语义模糊可搜索加密云盘设计方法及系统。

[0006] 为实现上述目的,本发明的技术方案是:一种语义模糊可搜索加密云盘设计方法,包括如下步骤,

步骤 S01:通过客户端,输入密钥对用户需上传的文件进行加密;

步骤 S02:提取上述加密文件的关键词,并进行语义扩展构建同义词集合;

步骤 S03:通过布隆过滤器对上述构建的同义词集合进行安全索引的构建,并将该安全索引上传至云存储服务器端;

步骤 S04:客户端根据用户输入密钥生成关键词陷门,并将该陷门提交给云存储服务器端;

步骤 S05:云存储服务器端根据用户提交的关键词陷门,通过安全索引查找相应文件,并将该文件返回给客户端;

步骤 S06:客户端通过密钥对云存储服务器端返回的文件进行解密。

[0007] 在本发明实施例中,所述云存储服务器端根据用户提交的更新请求,修改所述安全索引即可完成文件更新。

[0008] 在本发明实施例中,所述步骤 S01 的文件加密是通过对称加密算法对文件进行加密。

[0009] 在本发明实施例中,所述对称加密算法为 AES、DES 或 3DES 加密算法。

[0010] 本发明还提供了一种语义模糊可搜索加密云盘系统,包括一客户端和一云存储服务器端;所述客户端包括一实现文件加解密的文件加解密模块、一用于提取加密文件关键词的关键词提取模块、一用于对提取的关键词进行语义扩展并构建同义词集合的语义扩展模块、一通过布隆过滤器对所述同义词集合进行安全索引构建的构建安全索引模块和一上传安全索引、提交陷门以及上传 / 下载文件的第一通信模块;所述云存储服务器端包括一根据陷门查找安全索引的索引查找模块、一存储客户端上传加密文件的文件存储模块和一发送加密文件至所述客户端并与所述第一通信模块通信的第二通信模块。

[0011] 在本发明实施例中,所述文件加解密模块通过对称加密算法对文件进行加密。

[0012] 在本发明实施例中,所述对称加密算法为 AES、DES 或 3DES 加密算法。

[0013] 在本发明实施例中,所述客户端还包括一界面显示模块,该界面显示模块用于选取加密文件、显示搜索结果以及输入检索关键词。

[0014] 在本发明实施例中,所述语义扩展模块采用 WordNet 对所述同义词集合进行语义扩展。

[0015] 相较于现有技术,本发明具有以下有益效果:

1、本发明主要为密文文档的信息检索提供同义词语义模糊搜索功能,即使用户对某一关键词描述不同也能返回相关文档,并且支持多关键词检索,基于布隆过滤器的索引结构对文档的增加、删除和更新操作具有较高的效率;

2、用户将敏感数据存储到云服务器中,可以先将文档加密后上传,之后通过密文检索返回所需文档,保证文档的隐私性。

## 附图说明

[0016] 图 1 为本发明网络系统架构图。

[0017] 图 2 为本发明客户端结构图。

[0018] 图 3 为本发明云存储服务器端结构图。

## 具体实施方式

[0019] 下面结合附图,对本发明的技术方案进行具体说明。

[0020] 本发明一种语义模糊可搜索加密云盘设计方法,包括如下步骤,

步骤 S01 :通过客户端,输入密钥对用户需上传的文件进行加密 ;

步骤 S02 :提取上述加密文件的关键词,并进行语义扩展构建同义词集合 ;

步骤 S03 :通过布隆过滤器对上述构建的同义词集合进行安全索引的构建,并将该安全索引上传至云存储服务器端 ;

步骤 S04 :客户端根据用户输入密钥生成关键词陷门,并将该陷门提交给云存储服务器端 ;

步骤 S05 :云存储服务器端根据用户提交的关键词陷门,通过安全索引查找相应文件,并将该文件返回给客户端 ;

步骤 S06 :客户端通过密钥对云存储服务器端返回的文件进行解密。

[0021] 在本发明实施例中,所述云存储服务器端根据用户提交的更新请求,修改所述安全索引即可完成文件更新。

[0022] 所述步骤 S01 的文件加密是通过对称加密算法对文件进行加密;所述对称加密算法为 AES、DES 或 3DES 加密算法。

[0023] 本发明还提供了一种语义模糊可搜索加密云盘系统,包括一客户端和一云存储服务器端;所述客户端包括一实现文件加解密的文件加解密模块、一用于提取加密文件关键词的关键词提取模块、一用于对提取的关键词进行语义扩展并构建同义词集合的语义扩展模块(所述语义扩展模块采用 WordNet 对所述同义词集合进行语义扩展)、一通过布隆过滤器对所述同义词集合进行安全索引构建的构建安全索引模块和一上传安全索引、提交陷门以及上传 / 下载文件的第一通信模块;所述云存储服务器端包括一根据陷门查找安全索引的索引查找模块、一存储客户端上传加密文件的文件存储模块和一发送加密文件至所述客户端并与所述第一通信模块通信的第二通信模块;所述客户端还包括一界面显示模块,该界面显示模块用于选取加密文件、显示搜索结果以及输入检索关键词。

[0024] 所述文件加解密模块通过对称加密算法对文件进行加密;所述对称加密算法为 AES、DES 或 3DES 加密算法。

[0025] 以下为本发明的实施例。

[0026] 如图 1-3 所示,由于用户描述同一事物所用的词汇不同,若不能进行同义词搜索,将严重影响系统的使用性。针对现有可搜索加密技术无法完成语义模糊搜索,更新效率低下,多关键词查找效率低下等不足。本发明提出了一种基于布隆过滤器的多关键词语义模糊可搜索加密云盘设计方法及系统,属于软件开发领域,主要包括客户端和服务器端。

[0027] 客户端:用户通过客户端,输入密钥对文档进行加密操作,客户端获取文档关键词并进行语义扩展构建同义词集合,并对扩展后的同义词集合通过布隆过滤器构建安全索引,并将安全索引上传至服务器端,当用户提交搜索请求时,客户端根据密钥生成关键词陷门,并将陷门提交给服务器端完成搜索请求。对服务器返回的文档使用密钥进行解密。

[0028] 服务器端：存储密文文件和索引文件，对于用户提交的关键词陷门，通过安全索引找到相应文档，并将该文档返回给客户端。当用户需要更新时，根据用户提交的更新请求，增加、删除和更改操作只需修改服务器端相应的布隆过滤器索引即可，时间效率为 $O(1)$ 。

[0029] 本发明主要应用于可搜索加密云盘领域，提出了一种基于密文存储检索的应用软件，本软件主要包括：(1)若干客户端，每个客户端主要负责文档加解密，构建文档安全索引和提交搜索请求等功能，A. 文件加解密模块：主要通过调用现有的对称加密算法对文档进行加密，如AES, DES, 3DES等传统对称加密算法，B. 界面显示模块：主要负责选取加密文件、显示搜索结果以及输入检索关键词等功能，C. 关键词提取模块：主要通过提取文档标题的关键词并过滤掉a, the, of等停用词，D. 语义扩展模块：通过WordNet对提取的关键词进行语义扩展，构造同义词集合，E. 构建安全索引模块：通过带密钥的哈希函数构建布隆过滤器索引，F. 通信模块：主要负责提交用户搜索陷门，上传/下载密文文档集以及上传安全索引等操作；(2)云存储服务器端，主要负责文件存储和索引查找等功能，A. 索引查找模块：主要通过用户提交的关键词陷门查找索引并返回搜索结果，B. 文件存储模块：主要负责存储客户端上传的密文文档，通过索引查找结果，将检索到的文档返回给用户，C. 通信模块：主要负责将检索到的密文文档返回给用户以及接收客户端的发送请求。

[0030] 本文客户端和服务器端均采用Windows平台下的java jdk1.7环境进行开发。客户端运行在Windows 7系统下，对PDF文档的处理采用PDFBox的java类库，语义扩展采用WordNet 2.1版本并使用麻省理工学院提供的JWI WordNet java类库；云存储服务器端运行在Windows server 2003服务器下。

[0031] 本发明的使用方法如下：

用户先通过客户端对文档进行加密，并对文档提取关键词形成安全索引，将加密后的文档和安全索引上传到云存储服务器中，需要文档时，在搜索框中输入关键词，客户端对关键词构造陷门并将陷门提交至云存储服务器，服务器通过陷门对索引进行查找，将查找到的文档返回给客户端，客户端对文档进行解密。

[0032] 以上是本发明的较佳实施例，凡依本发明技术方案所作的改变，所产生的功能作用未超出本发明技术方案的范围时，均属于本发明的保护范围。

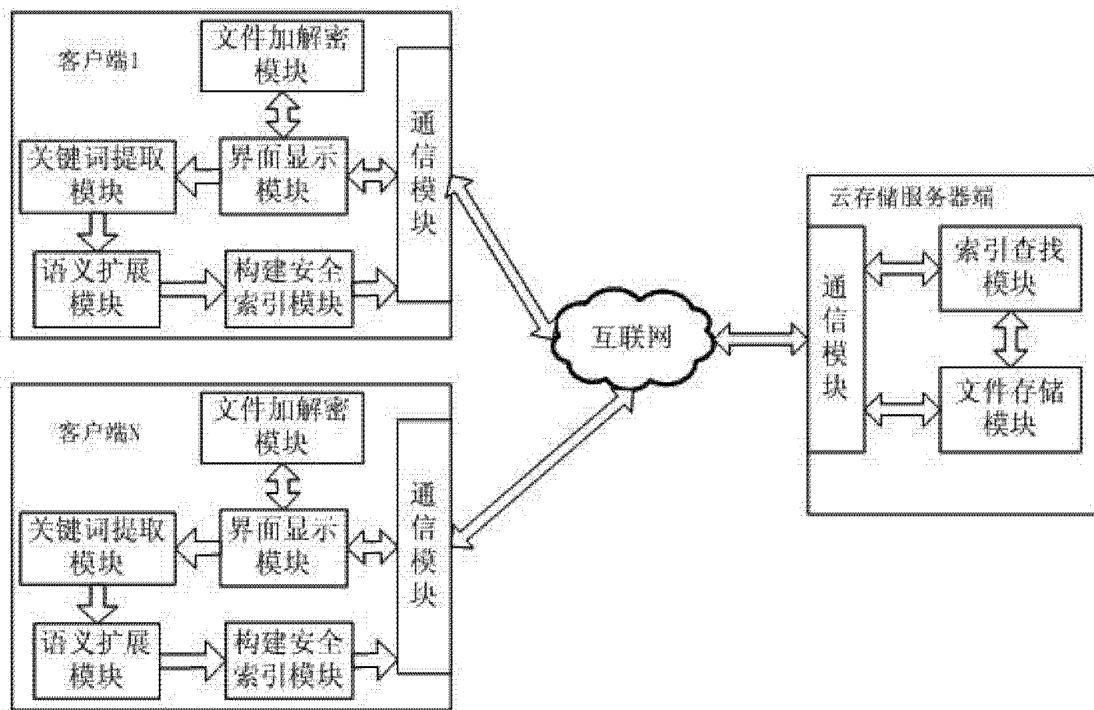


图 1

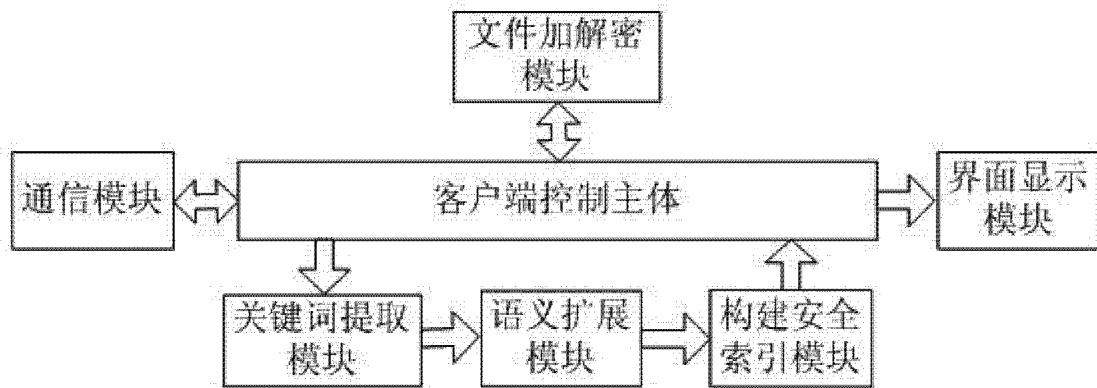


图 2

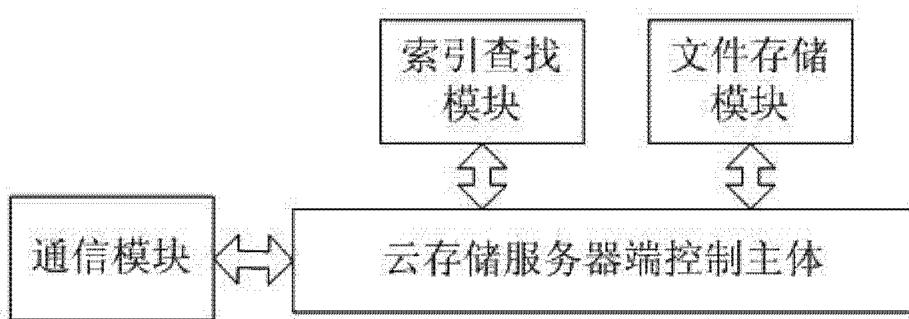


图 3