

基于细粒度授权的物联网搜索数据隐私保护方案

王佳慧^{1,2}, 刘川意^{2,3}, 方滨兴^{2,3}

(1. 北京邮电大学计算机学院, 北京 100876;
2. 哈尔滨工业大学深圳研究生院, 广东 深圳 518055;
3. 东莞电子科技大学电子信息工程研究院, 广东 东莞 523000)

摘要: 随着物联网和云计算技术的飞速发展和广泛应用, 物联网搜索引擎应运而生。然而, 物联网搜索引擎的开放性, 使在互联网搜索领域就已经存在的数据隐私问题变得更加突出。首先, 阐述了物联网搜索数据隐私保护的研究背景和安全挑战, 并针对安全挑战, 将密文搜索和基于属性的访问控制算法有机结合。然后, 提出了基于细粒度授权的物联网搜索数据隐私保护方案, 并对方案进行了安全分析和性能分析。最后, 总结并指出了未来研究方向。

关键词: 物联网搜索; 隐私保护; 属性加密; 密文搜索; 细粒度授权

中图分类号: TP309

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2017.00127

Data privacy preservation for the search of Internet of things based on fine-grained authorization

WANG Jia-hui^{1,2}, LIU Chuan-yi^{2,3}, FANG Bin-xing^{2,3}

(1. School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China;
2. Harbin Institute of Technology(Shenzhen), Shenzhen 518055, China;
3. Electronic and Information Engineering Institute, Dongguan University of Electronic Science and Technology, Dongguan 523000, China)

Abstract: With the rapid development of the Internet of things (IoT) technology and cloud computing technology, the search engine for Internet of things become a hot research topic. However, because of the openness of the search of IoT, the privacy in traditional Internet search area becomes more prominent and faces more challenges. Firstly, the research background and challenges of data privacy preservation for search of IoT were described. Secondly, the scheme of data privacy preservation for the search of Internet of things based on fine-grained authorization was proposed, which combined the encrypted search algorithm with the attribute based access control algorithm. Thirdly, the security analysis and performance analysis of the scheme were also carried out. Finally, the future research work was summarized and pointed out.

Key words: search for Internet of things, privacy preservation, attribute based encryption, searchable symmetric encryption, fine-grained authorization

收稿日期: 2016-10-14; 修回日期: 2016-12-20。通信作者: 王佳慧, jiahw520@gmail.com

基金项目: 国家高技术研究发展计划(“863”计划)基金资助项目(No.2015AA016001); 广东省产学研合作基金资助项目(No.2016B090921001); 山东省自主创新及成果转化专项基金资助项目(No.2014ZZCX03411); 国家自然科学基金资助项目(No.61370068)

Foundation Items: The National High Technology Research and Development Program of China (863 Program) (No.2015AA016001), Production-Study-Research Cooperation Project in Guangdong Province (No.2016B090921001), The Innovation Project in Shandong Province (No.2014ZZCX03411), The National Natural Science Foundation of China (No.61370068)

1 引言

在拥有海量实体的物联网^[1]中,人们越来越需要准确、及时、智能地搜索现实世界中物理实体的相关信息,如附近是否有适合带婴儿的西餐厅、附近是否有循环租用的车架等,由此使面向物联网的数据搜索引擎应运而生。例如,由微软研究院设计的 SenseWeb^[2]提供了基于静态元数据和基于位置的传感器搜索。由瑞士苏黎世联邦理工大学、德国吕贝克大学以及德国都科摩通信实验室设计的 Dyser^[3]是一个物联网实时搜索引擎,不仅支持物理实体静态信息的搜索,还能根据用户指定的当前状态实时搜索物理实体。Shodan^[4]是一个能提供在线设备的物联网搜索引擎,只要输入搜索关键字,就可以找到全世界在线的网络摄像头、路由器、信号灯、核电站、冰箱、医疗设备等包含信息漏洞的设备。尤斯曼哈克的“Thingful”是能提供智慧城市搜索的物联网搜索引擎,比如人们早上骑车上班时,通过个性化仪表盘查看当地数据,就能获悉污染指数和交通状况,以及附近是否有循环租用的车架。

尽管物联网搜索仍处于萌芽状态,但由此带来的数据隐私问题却已经不容忽视。韩国产业研究院认为,2020年因物联网数据隐私问题导致的经济损失将达到180亿美元^[注1]。物联网搜索的应用领域广泛,包括国民经济和社会发展各个领域,因此其数据隐私泄露不仅包含传统的网络虚拟空间,还将进一步侵犯到实体生活当中,除了经济损失外,个人的生命安全以及国家基础设施也都将受到极大的威胁。同时,物联网搜索对象和搜索空间的广泛性、搜索数据的高度动态性、搜索内容的高度时空性、搜索语言的复杂性等特点,使物联网搜索的数据隐私保护面临更大的安全挑战。可以说,数据隐私保护是制约全球化的物联网搜索发展并实际应用于产业界的关键因素。

首先,物联网搜索服务提供者后台服务器系统存有大量通过摄像头、传感器、可穿戴设备、家庭空调系统等感知设备收集的数据信息,这些

数据信息可能包括大量用户隐私数据,而由于物联网和行业应用的充分融合发展,这些数据信息如果不加以安全隐私的规范,将成为一个庞大的汇集百万电子耳、电子眼和电子鼻的监控系统。类似于“棱镜计划”,通过内部数据搜索和挖掘导致的个人数据隐私泄露和遭受侵犯将变得更加容易,且影响范围更广。因此,需确保在物联网搜索过程中,只有得到授权的数据搜索者才能访问相应权限的数据内容。

其次,物联网搜索服务如果被攻击者恶意利用,则可能对个人的生命安全及国家基础设施构成严重威胁。黑客通过物联网搜索引擎可以搜索家庭中有安全漏洞的智能设备,进而控制其他安全设备,比如控制智能大门,解除监控摄像头的监视功能等。通过物联网搜索有安全漏洞的智能儿童鞋、智能手环等,可以查看孩子地理位置和行为轨迹,这些漏洞一旦被别有用心地利用,则可能对儿童安全构成更大的威胁。而网站的研究者曾使用 Shodan 定位到了核电站的指挥和控制系统及一个粒子回旋加速器。使用 `port:554 has_screenshot: true` 在 Shodan 中就可以搜索到安装在各个地方的摄像头。由于物联网搜索引擎的存在,所有与物联网相连的实体,包括智能家庭、智能汽车、智能城市等显得更加不安全和脆弱。因此,确保物联网搜索服务不被别有用心攻击者利用也是一个很大的挑战。

最后,本文针对前述物联网搜索数据隐私保护面临的安全挑战,提出面向物联网搜索的基于细粒度授权的数据隐私保护方案。通过数据拥有者在将数据文件发送给物联网搜索服务提供者之前加密数据文件,确保在物联网搜索过程中,除数据搜索者之外的其他个体都无法获得授权访问的数据内容。并且通过基于属性的访问控制机制来实现物联网搜索服务细粒度的授权使用,确保不被别有用心攻击者利用,允许数据所有者对数据访问权限完全控制,对数据搜索者的权限做了双重限定,在数据搜索者加入到搜索系统时,首先通过属性权威中心对其搜索权限进行授权。随后通过数据文件访问权限认证之后,把和数据搜索者属性相对应的数据文件过滤出来。由此,对于同一个搜索请求,由于访问控制权限不同,

注1: 物联网的致命弱点(经济观察网), <http://www.eeo.com.cn/2016/0317/284248.shtml>。

不同的数据搜索者获得的搜索结果也不完全相同,可以更细粒度地确保数据文件的隐私。本文方案在确保数据隐私的前提下,最大限度地实现了数据的灵活细粒度访问。同时,为了提高效率,使用两级加密方案加密数据文件,即使用密文策略属性加密算法来加密对称密钥,使用对称密钥来加密数据文件,再一起发送给物联网搜索服务提供者。为了获得和数据搜索者匹配的搜索结果,需要验证数据搜索者拥有的权限,这通过基于属性的认证协议完成。

2 相关工作

目前还未检索到针对物联网搜索的数据隐私保护进行研究的相关文献,本文是将基于属性的访问控制及认证机制和密文搜索算法有机结合起来保护物联网搜索过程中的数据隐私。

2.1 密文搜索

密文搜索协议作为一个加密系统,可以在确保数据隐私的基础上,实现对密文的搜索操作。通常根据使用的加密协议分为基于对称加密的密文搜索方案和基于非对称加密的密文搜索方案。Song 等^[5]于 2000 年首次提出了基于对称加密的密文搜索方案,但是该方案仅支持固定大小的关键字搜索,而且由于其不构建关键字索引,采用顺序搜索,搜索效率很低。Goh^[6]通过构建基于安全索引的 SSE 构建方案 Z-IDX,使效率大大提高。Chang 等^[7]的方案避免了 Goh 方案的正向误检概率,基于预先构建的词典,为每个文件构建一个索引,进一步降低了通信开销。Curtmola 等^[8]采用倒排索引代替正向索引,搜索操作只需要 $O(1)$ 时间。Liesdonk 等^[9]提出了和文献[8]性能相当的方案,同时能有效支持属性更新。Kamara 等^[10]基于文献[8]做出改进,效率和安全性有所提高,随后进一步改进,支持动态索引更新。Cash 等^[11]进一步改善性能以支持大规模数据。Boneh 等^[12]提出了 PEKS 方案,并基于双线性对给出了几种构造方案。Abdalla 等^[13]进一步提出 PEKS 方案的完整定义,同时给出了基于身份匿名方案构造 PEKS 的流程。文献[14~16]针对文献[12]方案需要使用安全通道的问题,分别设计了在随机预言模型下^[14,15]和标准模型下^[16]不需要安全通道的

PEKS 方案。文献[16,17]提出了基于身份加密方案的 PEKS 方案。

随后的研究主要集中在以下 3 个方面。

1) 支持搜索方式的扩展:以上所述方案大多只支持单一关键字搜索,而后续的很多方案对搜索方式进一步扩展,包括支持多关键字搜索^[18~23]、子集搜索^[24]、模糊搜索^[20,21,25,26]、排序搜索^[22,23,27]、范围查询^[28]的方案。

2) 支持对服务器存储的密文文件动态地添加、更新或删除^[29~32]。

3) 搜索效率的提升,此方面主要通过减少双线性对使用的频率或使用其他技术替代双线性对来构造密文搜索方案^[33,34]。

2.2 属性加密

基于属性加密 (ABE, attribute based encryption) 方案首先由 Sahai 和 Waters 提出^[35],最初是为了改善基于生物信息的身份加密系统的容错性能,ABE 可以看作是基于身份加密方案的推广,将身份信息拆分成一个更细粒度的能够表示用户身份的信息集合,并将每个子集合称为属性。文献[36]采用 (t, n) 门限访问结构,分别提出了适用于小规模属性全集和大规模属性全集的方案,并给出了在选择身份安全模型下的安全性证明。随后 Goyal 等^[37]依据访问策略的实现方式不同分为两类:基于属性的密钥策略加密方案 (KP-ABE, key policy attribute based encryption) 和基于属性的密文策略加密方案 (CP-ABE, ciphertext policy attribute based encryption),并首次构造了 KP-ABE 方案。两者的主要区别首先在于访问策略关联的载体不同,前者是密钥和访问策略相关联,后者则是密文和访问策略相关联;其次,两者的主体地位不同,前者是解密者处于协议主体地位,解密者负责定义访问策略,而加密者仅负责将属性信息绑定到密文中,解密符合访问策略的加密者的密文;而后者是加密者处于主体地位,加密者负责定义访问策略,只有符合访问策略的解密者才能解密。文献[10]首次提出了 CP-ABE 方案。由于 CP-ABE 中数据拥有者拥有访问策略制定的权力,一个密文可以由多个不同的密钥进行解密,访问控制策略改变时只需根据新的访问控制策略重新加密密文,密钥管理开销

较小等特点,更适用于物联网搜索场景下的基于访问控制的隐私保护。

随后的研究主要集中在以下4个方面。

1) 支持多个密钥授权中心(KDC)的ABE方案^[38-43],与只支持单个KDC的ABE方案相比,其允许多个独立的KDC对属性和分发的密钥进行监督,分散了KDC的权利。

2) 支持灵活的细粒度访问控制策略表示的ABE方案^[44],主要研究如何支持策略属性的灵活变更、用户属性的灵活变更,密钥管理如何更好地支持策略属性和用户属性发生变更以及如何设计更为细粒度的访问控制。

3) 支持审计的ABE方案,目的在于避免密钥滥用,包括支持审计的KP-ABE^[45,46]方案和支持审计的CP-ABE^[47-49]方案。

4) 基于属性的访问控制和密文搜索相结合^[50-53]。其都是基于每个文件构建一个索引,即采用正向索引,仅支持索引中包含的关键字搜索,搜索的范围相对于密文内容搜索存在较大的局限性,无法满足密文信息的搜索需求,同时搜索效率大大降低。

3 背景知识

定义1 双线性映射。

令 G_0 和 G_1 为 p 阶乘法循环群,其中, p 为素数。令 g 为 G_0 生成元,映射 $e: G_0 \times G_0 \rightarrow G_1$ 是双线性映射,若满足如下3条性质。

1) 双线性: 对任意 $u, v \in G_0$, 任意的 $a, b \in \mathbb{Z}_p$, 满足 $e(u^a, v^b) = e(u, v)^{ab}$ 。

2) 非退化性: 映射 e 不是把 $G_0 \times G_0$ 所有的元都映射到 G_1 的单位元上。如果 g 是 G_0 的生成元,则 $e(g, g)$ 是 G_1 的生成元。

3) 可计算性: 对任意的 $u, v \in G_0$, 都存在有效的算法计算 $e(u, v)$ 。

则 e 为对称双线性映射,因为 $e(g^a, g^b) = e(g, g)^{ab} = e(g^b, g^a)$ 。

定义2 访问策略树。

树的非叶子节点 x 表示门限门 (k_x, num_x) , 由其第 num_x 子节点和门限值 k_x 描述,其中, $0 < k_x < num_x$ 。当 $k_x = 1$ 时,门限门为 OR 门;当 $k_x = num_x$ 时,门限门为 AND 门。树的每个叶子和一个属

性相关,其中, $k_x = 1$ 。

为了便于操作,定义如下函数。

$parent(x)$: 返回树中节点 x 的父节点(除根节点)。

$att(x)$: 只有当节点 x 是叶子节点时才有定义,返回和叶子节点 x 相关的属性。

$index(x)$: 假设每个节点的子节点用 $1 \sim num$ 的树标识,则返回节点 x 的标识。

定义3 CP-ABE 算法。

一个典型的 CP-ABE 方案通常由以下几个算法组成。

Setup: 初始化算法。输入公开参数和属性全集 U , 输出一个公钥对 $\{PK, MSK\}$ 。其中, PK 为系统公钥, MSK 为主密钥。

Encrypt: 加密算法。输入明文 M 、系统公钥 PK 和访问结构 A , 输出和访问结构相关联的密文 C 。

KeyGen: 私钥生成算法。输入属性集合 S 、主密钥 MSK 和系统公钥 PK , 输出用户私钥 SK 。

Decrypt: 解密算法。输入密文 C 、用户私钥 SK 和系统公钥 PK , 如果 $S \in A$, 解密成功得到明文 M 。

定义4 对称密文搜索算法。

一个典型的对称密文搜索方案通常由以下几个算法组成。

Setup: 系统初始化算法,用于生成对称加密算法和伪随机函数的密钥。

TokenGen: 搜索令牌生成算法,利用用户的关键字生成对应的搜索令牌。

Enc: 加密算法,利用密钥将用户文件和关键字信息加密成对应的密文和加密索引。

Search: 搜索算法,运行在服务器端,利用搜索令牌在加密索引上执行搜索操作。

Dec: 解密算法,运行在客户端,利用密钥解密文件。

4 系统架构和解决方案

4.1 系统架构

基于细粒度授权的物联网搜索数据隐私保护方案及流程如图1所示,主要包含以下6个参与者。

1) 属性权威中心(AAC, attribute authority center)。属性权威中心在本系统中是一个可信机构,可以包含多个属性权威机构,主要负责系统

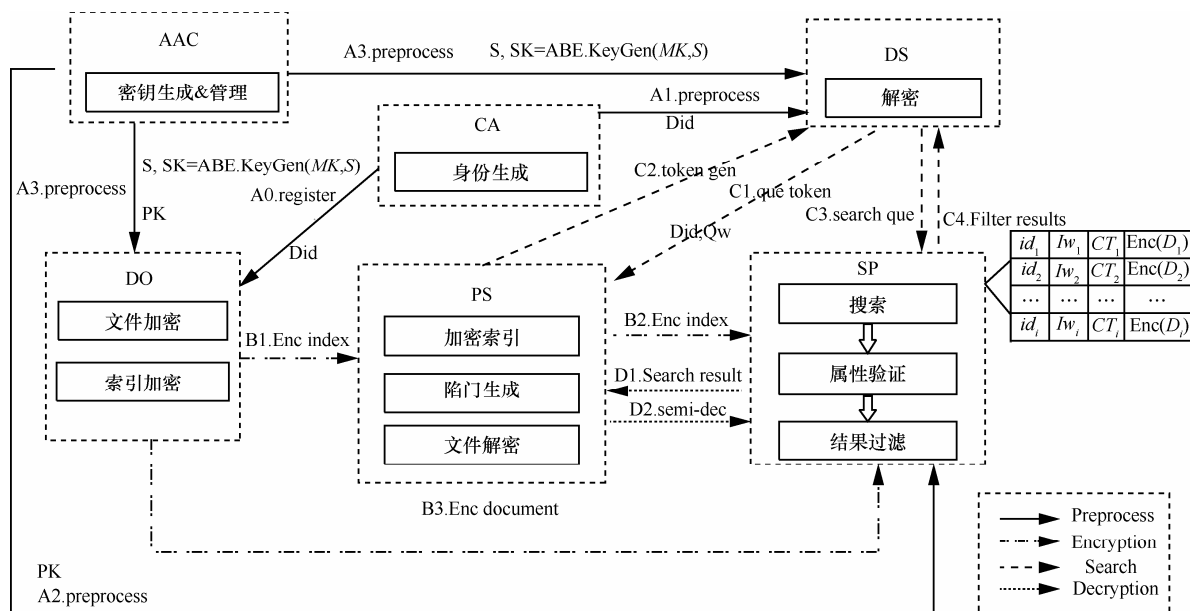


图1 基于授权的物联网搜索数据隐私保护方案

参数初始化、为系统生成公开参数、根据身份和属性为数据搜索者生成和分发密钥。根据数据所有者的请求撤销数据搜索者的访问权限，并向代理服务器提供用户撤销证书。

2) 数据所有者 (DO, data owner)。数据所有者在代理服务器的协助下生成密文索引，上传加密数据文件和密文索引给物联网搜索服务提供者。数据所有者对要上传到物联网搜索服务提供者的数据文件，设置不同的访问策略，并且将访问策略嵌入到密文中。

3) 数据搜索者 (DS, data searcher)。数据搜索者拥有一些指定的属性集合以及相应的密钥。可以提出搜索请求，并获得和其访问权限匹配的搜索结果。解密加密的搜索结果，本方案中数据搜索者只下载其能解密的文件，使通信开销大大减少。并且，物联网搜索者通常使用资源有限的设备，而本方案中由代理服务器执行部分解密操作来降低数据搜索者的开销，同时确保代理服务器无法获得明文搜索结果。

4) 代理服务器 (PS, proxy server)。代理服务器负责协助数据搜索者生成密文索引和搜索陷门，并负责为数据搜索者执行数据文件部分解密操作。

5) 物联网搜索服务提供者 (SP, IoT search provider)。物联网搜索服务提供者是诚实、好奇

的服务提供商。主要负责数据的存储和搜索，并负责属性验证和数据过滤，会诚实地根据数据搜索者的搜索请求返回相应访问权限的加密数据给代理服务器，但是也可能在通信过程中试图获取数据内容。

6) 认证中心 (CA, certificate authority)。认证中心负责为数据所有者和数据搜索者颁发身份证书。

4.2 方案流程

如图1所示，在属性权威中心注册后，数据搜索者使用加密的搜索关键词发起搜索请求，由代理服务器协助生成搜索陷门，再通过与物联网搜索服务提供者交互，获得过滤后的数据搜索者有权访问的加密文件，最后由代理服务器部分解密加密数据文件返回给数据搜索者最终解密，获得明文数据。

具体包括以下几个阶段。

1) 系统初始化阶段

属性权威中心运行系统初始化算法。记为 $\text{ABE.Setup}() \rightarrow (PK, MSK, K_{msk})$ 。输出主密钥 MSK 、系统公开参数 PK 和主搜索密钥 K_{msk} 。通过安全信道发布 PK 给数据所有者和物联网搜索服务提供者，保存 MSK 和 K_{msk} 私有。随机选择主搜索密钥 $K_{msk}=k(k \in Z_p)$ 。随机选择 $\alpha, \beta \in Z_p$ ，并计算 PK 和 MSK 。其中， $PK = \{G_1, g, g^\beta, e(g, g)^\alpha\}$ ， $MSK = \{\beta, g^\alpha\}$ ， e 为定义1的双线性

映射。

2) 认证中心颁发证书

认证中心为数据所有者和数据搜索者颁发身份证书 D_{id} 。此证书具有唯一性，可以用来标识不同的数据所有者和数据搜索者。将身份标识和属性信息分离，避免在数据搜索者请求搜索时需要提交身份标识从而泄露身份信息。

3) 为数据搜索者和数据所有者分发密钥

当数据搜索者和数据所有者第一次搜索数据或第一次上传数据时，需要到属性权威中心注册，获得相应的访问权限，根据数据搜索者的 D_{id} 分配相应属性集 S ，调用 $ABE.KeyGen(MK, S)$ 生成并通过安全信道分配秘密密钥 SK ，记为 $ABE.KeyGen(MSK, S) \rightarrow SK$ 。其中， r 和 r_j 为属于 Z_p 的随机数，

$$SK = \{D = g^{\frac{\alpha+r}{\beta}}, \forall j \in S: D_j = g^r H(j)^{r_j}, D_j^i = g^{r_j^i}\}。$$

随机选择 $\mu \in Z_p$ ，令 $U_{D_{id}} = \mu, P_{D_{id}} = g^{\frac{k}{\mu}}$ 。发送 $(U_{D_{id}}, SK)$ 给数据搜索者和数据所有者，同时发送 $(D_{id}, P_{D_{id}})$ 给代理服务器，使代理服务器可以协助数据所有者生成加密索引，协助数据搜索者生成搜索陷门。

4) 关键字索引构建

为了使数据搜索者可以在加密后的数据中完成搜索，必须构建加密的关键字索引。物联网各实体信息通常以文本形式存储在物联网搜索服务提供商。假设关键字信息已经从文本信息中提取，记为 $W = \{w_1, w_2, \dots, w_n\}$ 。

索引构建方法如下。

对于每一个关键字 $w_i \in W$ ，数据所有者计算 $A_i = h(w_i)^{r_i}$ ，其中， h 是将关键字映射到 G_0 中的随机数的散列函数， $r_i \in Z_p$ 是一个随机数。然后数据所有者将 $(A_1, A_2, \dots, A_n)$ 和 D_{id} 发送给代理服务器。代理服务器根据 D_{id} 查询到数据所有者相应的 $P_{D_{id}}$ ，接着对于每一个 A_i ，计算 $B_i = e(A_i, P_{D_{id}})$ 。代理服务器发送 $\{B_i, 1 \leq i \leq n\}$ 给数据所有者。接着数据所有者对于每一个关键字 w_i ，计算 $k_i = h\left(\frac{U_{D_{id}}}{V_i^{r_i}}\right)$ ，令 $I_{w_i} = (R_i, [R_i]_{k_i})$ ，其中， $[R_i]_{k_i}$ 表示随机数使用安全的对称密钥算法（如 AES），密钥为 k_i 加密。最终，数据所有者为关键字集 W 生成的加密索引记为 $I_w = \{I_{w_1}, I_{w_2}, \dots, I_{w_n}\}$ 。

5) 数据文件加密上传

在构造索引 I_w 后，数据所有者加密数据文件内容。首先，随机选择对称密钥 sk 并使用其加密数据文件内容，可以使用任意安全的对称加密算法（如 AES）。然后，为数据文件指定访问策略树 $\mathcal{T}$ ，使用基于属性加密的算法 $ABE_Enc(PK, sk)$ 加密，记为 CT 。以自上向下的方式处理访问策略树 $\mathcal{T}$ ，对树中每一个节点，选择一个度为 $k_x - 1$ 的多项式 $q_x(\cdot)$ (k_x 是门限值)，且 $q_x(\cdot)$ 必须满足如下条件：如果 x 是访问策略树的根节点，则 $q_x(0) = s$ ，其中， s 是 Z_p 中的随机数。对于其他任意节点，令 $q_x(0) = q_{parent(x)}(index(x))$ 并随机选择 d_x 个其他节点来定义 q_x 。令 ψ 是访问策略树的叶子节点集合。 H 是映射属性到 G_0 中的随机数的散列函数。 $CT = \{\mathcal{T}, \tilde{C} = sk \cdot e(g, g)^{\alpha s}, C = g^{\beta s}, \forall x \in \psi: C_x = g^{q_x(0)}, C'_x = H(att(x))^{q_x(0)}\}$ 。最终，数据所有者上传索引 I_w ，加密后的数据文件和 CT 给物联网搜索服务提供商。由物联网搜索服务提供商者为数据文件指派唯一的数据文件标识 id 。物联网搜索服务提供商者中文件的数据存储格式如图 2 所示。

文件 id	索引 I_w	CT	$Enc(DataFuke)$
---------	----------	------	-----------------

图2 数据存储格式

6) 搜索算法

在属性权威中心注册后，就可以使用物联网搜索服务搜索其感兴趣的内容。搜索阶段包含以下 2 个过程。

① 陷门生成算法

数据搜索者要想搜索关键字为 w 的数据文件，首先要在代理服务器的协助之下生成陷门。数据搜索者计算 $Q_w = h(w)^{U_{D_{id}}}$ ，然后发送他的身份 D_{id} 和 Q_w 给代理服务器。当代理服务器收到陷门生成请求，代理服务器通过查询获得数据搜索者相应的 $P_{D_{id}}$ 。然后，返回搜索陷门 $t = h(e(Q_w, P_{D_{id}}))$ 给数据搜索者。

② 物联网搜索服务提供商者搜索数据

当数据搜索者接收到搜索陷门 t 之后，发送 $(t, S = \{A_1, A_2, \dots, A_m\})$ 给物联网搜索服务提供商者，其中， S 是其秘密密钥 SK 关联的属性集合。当物联网搜索服务提供商者接收到搜索请求 $(t, S = \{A_1, A_2, \dots, A_m\})$ 之后，运行以下的搜索算法开始搜索。

输入：陷门和可搜索加密文件库。

输出：数据搜索者有权访问的加密文件集合

和 CT 。

- 1) $F = \emptyset$;
- 2) for 每一个文件 F_{id}
- 3) 获得相应文件索引; I_w
- 4) for $i=1; i \leq n; i++$
- 5) get the I_{w_i} from the index I_w ;
- 6) if $[R_i]_{k_i} = [R_i]_l$
- 7) $F = F \cup \{F_{id}\}$;
- 8) end if
- 9) end for
- 10) end for
- 11) 随机选择 $p_m \in E_p$
- 12) Set $\mathcal{T} = A_1 \wedge A_2 \wedge \dots \wedge A_m$
- 13) $CT = \text{ABE_Enc}(PK, p_m, \mathcal{T})$
- 14) if $p'_m = p_m$
- 15) return (success) ;
- 16) end if
- 17) for 文件集 F 中的每一个文件 F_{id}
- 18) if DS 属性不满足 F_{id} 属性
- 19) 从文件集 F 中删除 F_{id} ;
- 20) end if
- 21) for 文件集 F 中的每一个文件 F_{id}
- 22) 输出相应的 CT 和加密文件;
- 23) end for

此搜索算法主要分为 3 个部分。第一部分对应伪代码 1)~10)行, 对于每一个文件的加密索引, 比较使用搜索陷门加密的随机数和分解后的索引中相应部分是否相同, 如果相同, 则将其加入到结果集中。第二部分对应 11)~16)行, 主要是属性验证。物联网搜索服务提供者选择随机消息, 调用加密算法 $\text{ABE_Enc}(PK, p_m, \mathcal{T})$ 加密 p_m , 发送密文给数据搜索者。其中, 访问策略树构造遵从格式 $\mathcal{T} = A_1 \wedge A_2 \wedge \dots \wedge A_m$ 。在收到密文之后, 数据搜索者调用 ABE_Dec 算法解密, 发送解密后的结果 p'_m 给物联网服务提供者。如果解密后的结果等于 p_m , 物联网搜索服务提供商就可以确信其拥有来自属性权威中心指派的属性集 S 。因为只有属性集相关的密钥才可以正确解密消息 p_m 。第三部分依据数据搜索者拥有的属性集, 进一步过滤搜索结果列表, 将数据搜索者无权访问的文件过滤掉, 对应 17)~22)行。最后将过滤后的结果返回给数据搜索者, 只有数据搜索者有权访问的数据才会返回, 从而实现物联网搜索服务细粒度的

授权访问。

7) 文件解密

数据搜索者接收到物联网搜索服务提供者发送的搜索结果后, 在代理服务器的协助下, 完成文件解密。这是因为在物联网搜索场景中, 数据搜索者的资源通常都有限, 交由运算能力更为强大的代理服务器执行, 如此做可以减少数据搜索者的运算开销。算法如下。数据搜索者将其私钥 $SK = \{D, D_j, D'_j, j \in S\}$ 分解成 2 个部分, 记为 $SK_1 = \{D\}$, $SK_2 = \{D_j, D'_j, j \in S\}$ 。并且将搜索结果中所有的 CT 和 SK_2 发送代至理服务器, 代理服务器运行如下解密算法得到 $A = e(g, g)^{rs}$ 。对于每一个叶子节点 x , 属性 $i = \text{att}(x)$, 如果 $i \in S$, 计算 $\text{Decnode}(CT, SK_2, x) = \frac{e(D_i, C_x)}{e(D'_i, C'_x)} = \frac{e(g^r H(i)^{r_i}, g^{qx(0)})}{e(g^r i, H(i)^{qx(0)})} = e(g, g)^{r_{i,x}(0)}$, 否则输出 $\perp$ 。对每一个中间节点, 使用拉格朗日插值法对至少 k_x 个如其子节点 $\{z_j\}$ 的 $e(g, g)^{r_{z_j}(0)}$ 形式来计算 $e(g, g)^{r_{x,x}(0)}$ 。最后, 对于访问策略树中的根节点 R , 令 $A = e(g, g)^{r_{R,x}(0)} = e(g, g)^{rs}$, 发送给数据搜索者。最后, 数据搜索者使用 SK_1 来继续解密得到对称密钥 $sk = \frac{\tilde{C}}{e(C, D)}$, 再使用对称密钥 sk 解密对应的数据文件内容。

8) 搜索权限撤销

假设要撤销的数据搜索者的证书为 D_{id} , 仅需要属性权威中心通知代理服务器删除元组 $(D_{id}, P_{D_{id}})$, 若无法构建搜索陷门, 则无法再使用物联网搜索服务。

5 安全分析和性能分析

5.1 安全分析

本文所设计的方案假设物联网搜索服务提供者是诚实好奇的 (半可信模型), 即对加密的数据内容或者收到的消息好奇, 但会正确执行搜索服务, 并且假设属性权威中心是可信第三方。因此安全分析主要关注数据文件内容、关键字的隐私性、身份信息隐私性、合谋攻击。

1) 数据文件隐私性

数据文件隐私性即需要数据拥有者的数据对

代理服务器、物联网搜索服务提供者、属性权威中心、未授权用户保持机密性。在本文所提方案中,数据文件采用2层加密方案,数据文件使用对称密钥算法加密,然后再使用CP-ABE算法加密对称密钥。假设对称密钥算法(如AES)是安全的,则其数据机密性依赖于CP-ABE算法。而在文献[10]中证明了CP-ABE在通用双线性群模型下是可证安全的,因此本文方案也是可证安全的,可以保证数据文件的隐私性。

在解密阶段,物联网搜索服务提供者将数据搜索者有权访问的加密数据文件返回给数据搜索者,因为没有私钥,其无法获得加密数据文件的明文。而数据搜索者未授权的数据文件,由物联网搜索服务提供者过滤掉,因此数据搜索者无法访问未授权数据文件。代理服务器和属性权威中心不参与具体数据搜索过程,也就无法获得数据文件明文。因此可以确保在物联网搜索过程中,除数据搜索者外的其他个体都无法获得搜索的数据内容,就保证了数据拥有者的数据只能被具有权限的数据搜索者访问。

2) 关键字隐私性

关键字隐私性需要确保物联网搜索服务提供者以及攻击者无法获得索引或者搜索陷门中的关键字。发送给物联网搜索服务提供者的安全索引和陷门中的关键字经过数据所有者和数据搜索者分别加密,而加密密钥保存在数据所有者和数据搜索者中,因此物联网搜索服务提供者以及攻击者都无法获得关键字信息,保证了关键字的隐私性。代理服务器虽然协助生成安全索引和陷门,但是也无法获得关键字信息。

3) 身份信息的隐私性

由证书中心颁发身份证书之后,数据搜索者在使用物联网搜索服务时,无需向物联网搜索服务提供者出示身份证书,从而确保身份信息的隐私性。

4) 合谋攻击

基于属性加密机制的授权访问方案,数据搜索者的密钥SK与随机数相关,因而可以防止数据搜索者合谋攻击。

5.2 性能分析

由于物联网搜索服务提供者后台通常是云计算平台,平台计算能力强大,所以本文主要关注有数据隐私保证的搜索系统对数据所有者和数据搜索者造成的计算开销影响。对数据所有者而言,主

要关注索引建立时间和加密对称密钥时间,对数据搜索者而言,开销最大的CP-ABE解密操作大部分由代理服务器完成,因此引入的额外开销可以忽略。

基于调用GNU高精度数学库(GMP, GNU MPBignum library)^{注2}的双线性对密码库(PBC, pairing-based cryptography library)^{注3}(GMP库提供任意精度的算法API供PBC调用以构建双线性对密码系统)实现本文所提方案。加密的内容为128 bit AES密钥。系统环境如下。数据所有者:内存8 GB;CPU 2.3 GHz。数据搜索者:内存8 GB;CPU 2.3 GHz。

数据所有者相对无隐私保护的搜索过程来说,额外开销主要包括加密数据文件的时间和索引建立时间,其中,加密数据文件的时间具体包括使用对称密钥加密数据文件的时间、使用CP-ABE加密对称密钥的时间。图3表明数据所有者加密128 bit AES对称密钥的时间随属性增加大体呈线性增加趋势。在物联网搜索实际场景中,可能并不需要30个属性来做访问控制,当属性个数小于10时,加密的时间小于150 ms,以此开销来换取数据隐私的保护是值得的。图4表明在属性个数小于10时,索引建立时间小于12 ms。

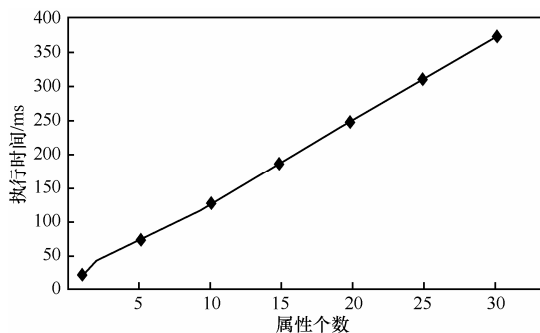


图3 128 bit AES 密钥加密时间

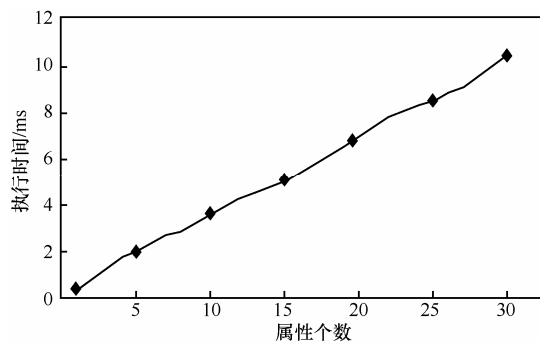


图4 索引建立时间

注2: <http://crypto.stanford.edu/pbc/>。

注3: <http://gmplib.org/>。

数据搜索者在获得搜索权限之后, 数据搜索过程对数据搜索者是透明的, 相对无隐私保护的搜索过程来说, 仅需在搜索结果返回后, 从加密文件中分离出加密使用的对称密钥, 并使用其解密数据文件。增加的计算开销主要包括 CP-ABE 解密对称密钥时间和对称密钥解密数据文件时间开销。而本文方案将数据解密的大部分工作外包给代理服务器, 从而使数据搜索者解密的时间开销大大降低, 对数据搜索者的性能影响较小。

6 结束语

本文首次提出了基于细粒度授权的物联网搜索数据隐私保护方案, 在尽可能不影响搜索服务质量的同时, 确保数据拥有者的数据文件和关键字隐私。通过建立数据文件安全索引和密文搜索方案来确保物联网搜索的数据文件和关键字的隐私性, 通过基于属性的访问控制及认证算法来实现物联网搜索的细粒度灵活的访问控制, 并对方案的安全性和性能进行了分析。本文方案通过物联网搜索服务提供者来过滤掉数据搜索者无权访问的数据, 而不需要将所有匹配数据结果下载到本地, 大大减少了数据搜索者的通信开销; 且本文对数据文件使用两级加密, 仅使用 CP-ABE 算法来加密对称密钥, 降低了 CP-ABE 算法高开销对隐私方案的影响。另外, 通过将部分解密工作外包给代理服务器, 使数据搜索者的计算开销大大降低。

本文未对数据搜索的属性和密钥管理做进一步分类和研究, 且本文面向物联网搜索的隐私保护技术, 其中, 搜索的条件主要是基于关键字的搜索, 研究支持物联网复杂搜索条件的隐私保护模型和应用也十分必要。这也是下一步主要的工作内容。

参考文献:

- [1] ATZORI L, IERA A, MORABITO G. The Internet of things: a survey[J]. *Computer Networks*, 2010, 54(15): 2787-2805.
- [2] KANSAL A, NATH S, LIU J, et al. BSenseWeb: an infrastructure for shared sensing[J]. *IEEE Multimedia*, 2007, 14(4): 8-13.
- [3] OSTERMAIER B, ROMER K, MATTERN, et al. A real-time search engine for the Web of things[C]//*Internet of Things*, Tokyo. 2010.
- [4] BODENHEIM R, BUTTS J, DUNLAP S. Evaluation of the ability of the Shodan search engine to identify Internet-facing industrial control devices[J]. *International Journal of Critical Infrastructure Protection*, 2014, 7(2): 114-123.
- [5] SONG X D, WAGNER D, PERRIG A. Practical techniques for searches on encrypted data[C]//*The IEEE Symposium on Security and Privacy*. 2000.
- [6] GOH E. Secure indexes[R]. IACR ePrint Cryptography Archive, 2003.
- [7] CHANG Y, MITZENMACHER M. Privacy preserving keyword searches on remote encrypted data[C]//*The Applied Cryptography and Network Security*. 2005.
- [8] CURTMOLA R, GARAY J, KAMARA S, et al. Searchable symmetric encryption: improved definitions and efficient constructions[C]//*The 13th ACM Conference on Computer and Communications Security (CCS 2006)*. 2006.
- [9] LIESDONK P V, SEDGHI S, DOUMEN J, et al. Computationally efficient searchable symmetric encryption[C]//*Secure Data Management, VLDB Workshop (SDM 2010)*. 2010: 87-100.
- [10] KAMARA S, PAPAMANTHOU C, ROEDER T, et al. Dynamic searchable symmetric encryption[C]//*2012 ACM Conference on Computer and Communications Security*. 2012: 965-976.
- [11] CASH D, JAEGER J, JARECKI S, et al. Dynamic searchable encryption in very-large databases: data structures and implementation[C]//*Network and Distributed System Security Symposium (NDSS'14)*. 2014.
- [12] BONEH D, CRESCENZO G D, OSTROVSKY R. Public key encryption with keyword search[C]//*Eurocrypt'04*. 2004.
- [13] ABDALLA M, BELLARE M, CATALANO D. Searchable encryption revisited: consistency properties, relation to anonymous IBE, and extensions[C]//*Crypto'05*. 2005.
- [14] BAEK J, SAFARI-NAINI R, SUSILO W. Public key encryption with keyword search revisited[C]//*The International Conference on Computational Science and Applications (ICCSA 2008)*. 2008.
- [15] RHEE H S, PARK J H, SUSILO W, et al. Improved searchable public key encryption with designated tester[C]//*The ACM Symposium on Information, Computer and Communications Security (ASIACCS 2009)*. 2009.
- [16] FANG L, SUSILO W, GE C, et al. A secure channel free public key encryption with keyword search scheme without random oracle[C]//*The International Conference on Cryptology and Network Security*. 2009.
- [17] KERSCHBAUM F, SORNIOTTI A. Searchable encryption for outsourced data analytics[C]//*The 7th European Conference on Public Key Infrastructures, Services and Applications (EuroPKI'10)*. 2010.
- [18] CAO N, WANG C, REN K. Privacy-preserving multi-keyword ranked search over encrypted cloud data[C]//*IEEE Infocom*. 2011.
- [19] SUN W, WANG B, CAO N. Privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking[C]//*ACM Symposium on Information, Computer and Communications Security*. 2013: 71-82.
- [20] CHUAH M, HU W. Privacy-aware bedtree based solution for fuzzy multi-keyword search over encrypted data[C]//*The International Conference on Distributed Computing Systems Workshops*. 2011.
- [21] WANG B, YU S, LOU W. Privacy-preserving multi-keyword fuzzy search over encrypted data in the cloud[C]//*IEEE Infocom*. 2014.
- [22] MASHAURI D, LI R, HAN H. Adaptive multi-keyword ranked search over encrypted cloud data[C]//*The International Conference on Collaborate Computing*. 2015.

- [23] SUN X, WANG X, XIA Z. Dynamic multi-keyword top- k ranked search over encrypted cloud data[J]. The International Journal of Security and Its Applications, 2014, 8(1): 319-332.
- [24] BONEH D, WATERS B. Conjunctive, subset, and range queries on encrypted data[C]/TCC. 2007.
- [25] SUN W, WANG B, CAO N. Verifiable privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking[J]. IEEE Transactions on Parallel and Distributed Systems, 2014, 25(11): 3025-3035.
- [26] LI J, WANG Q, WANG C. Fuzzy keyword search over encrypted data in cloud computing[C]/IEEE Infocom. 2010.
- [27] WANG C, CAO N, REN K. Enabling secure and efficient ranked keyword search over outsourced cloud data[J]. IEEE Transactions on Parallel and Distributed Systems (TPDS), 2011, 23(8): 1467-1749.
- [28] SHI E, BETHENCOURT V, CHAN H. Multi-dimensional range query over encrypted data[C]/IEEE Symposium on Security and Privacy. 2007.
- [29] KAMARA S, PAPAMANTHOU C. Parallel and dynamic searchable symmetric encryption[M]/Financial Cryptography and Data Security. Berlin: Springer, 2013: 258-274.
- [30] KAMARA S, PAPAMANTHOU C. Parallel and dynamic searchable symmetric encryption[C]/CCS. 2012.
- [31] STEFANOV E, PAPAMANTHOU C, SHI E. Practical dynamic searchable encryption with small leakage[C]/NDSS. 2014.
- [32] CASH D, JAEGER J, JARECKI S, et al. Dynamic searchable encryption in very large databases: data structures and implementation[C]/Network and Distributed System Security Symposium (NDSS). 2014.
- [33] DI CRESCENZO G, SARASWAT V. Public key encryption with searchable keywords based on Jacobi symbols[C]/Indocrypt 2007. 2007.
- [34] LAI X, LU R, FOXTON K. An efficient searchable encryption scheme and its application in network forensics[C]/E-Forensics. 2010.
- [35] NAVEED M, PRABHAKARAN M, GUNTER C. Dynamic searchable encryption via blind storage[C]/IEEE Symposium on Security and Privacy. 2014.
- [36] SAHAI A, WATERS B. Fuzzy identity based encryption[C]/Eurocrypt. 2005.
- [37] GOYAL V, PANDEY O, SAHAI A. Attribute-based encryption for fine-grained access control of encrypted data[C]/The 13th ACM Conference on Computer and Communications Security(CCS). 2006.
- [38] CHASE M. Multi-authority attribute based encryption[C]/TCC. 2007.
- [39] BOŽOVIC V, SOCEK D, STEINWANDT R. Multiauthority attribute-based encryption with honest-but-curious central authority[J]. International Journal of Computer Mathematics. 2012, 89(3): 268-283.
- [40] CHASE M, CHOW S. Improving privacy and security in multi-authority attribute-based encryption[C]/The 16th ACM Conference on Computer and Communications Security (CCS). 2009.
- [41] LEWKO A, WATERS B. Decentralizing attribute-based encryption[C]/Eurocrypt 2011. 2011.
- [42] LIU Z, CAO Z, HUANG Q. Fully secure multi-authority ciphertext-policy attribute-based encryption without random oracles[C]/The European Symposium on Research in Computer Security (ESORICS). 2011.
- [43] HAN J, SUSILO W, MU Y. Privacy-preserving decentralized key-policy attribute-based encryption[J]. IEEE Transactions on Parallel and Distributed Systems. 2012, 23(11): 2150-2162.
- [44] YU S, WANG C, REN K. Achieving secure, scalable, and fine-grained data access control in cloud computing[C]/IEEE Infocom. 2010.
- [45] YU S C, REN K, LOU W J. Defending against key abuse attacks in KP-ABE enabled broadcast systems[C]/Security and Privacy in Communication Networks. 2009.
- [46] WANG Y, CHEN K, LONG Y. Accountable authority key policy attribute-based encryption[J]. Science China: Information Sciences, 2012, 55(7): 1631-1638.
- [47] LI J, REN K, KIM K. A2BE: Accountable attribute-based encryption for abuse free access control[R]. Cryptology ePrintArchive. 2009.
- [48] LI J, REN K, ZHU B. Privacy-aware attribute based encryption with user accountability[C]/The 12th International Conference. 2009.
- [49] LI J, HUANG Q, CHEN X. Multi-authority ciphertext-policy attribute-based encryption with accountability[C]/The 6th International Symposium on Information, Computer and Communications Security (ASIACCS). 2011.
- [50] SUN W, YU S, LOU V. Protecting your right: attribute-based keyword with fine-grained owner enforced search authorization in the cloud[C]/IEEE Infocom. 2014.
- [51] HAN F, QIN J, ZHAO H. A general transformation from KP-ABE to searchable encryption [J]. Future Generation Computer Systems, 2014, 30: 107-115.
- [52] BOUABANATEBIBEL T, KACI A. Parallel search over encrypted data under attribute based encryption on the cloud computing[J]. Computers & Security, 2015.
- [53] KACI A, BOUABANA-TEBIBEL T. Access control reinforcement over searchable encryption[C]/The 15th IEEE International Conference on Information Reuse And Integration. 2014.
- [54] CASH D, JARECKI S, JUTLA C. Highly-scalable searchable symmetric encryption with support for boolean queries[C]/Crypto, 2013.
- [55] SADAKANE K. Fast algorithms for k -word proximity search[J]. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2001, 84(9): 2311-2318.

作者简介:



王佳慧 (1983-), 女, 山西大同人, 北京邮电大学博士生, 主要研究方向为云计算与云安全、数据安全与数据保护。

刘川意 (1982-), 男, 四川乐山人, 北京邮电大学副教授, 主要研究方向为云计算、网络存储、可信计算。

方滨兴 (1960-), 男, 江西万年人, 中国工程院院士, 主要研究方向为信息与网络安全、内容安全。