**目录**

[引言 1](#_Toc416274767)

[1.访问控制策略研究现状 2](#_Toc416274768)

[1.1自主访问控制 2](#_Toc416274769)

[1.2强制访问控制 3](#_Toc416274770)

[1.3基于角色的访问控制 3](#_Toc416274770)

[1.4基于属性的访问控制 3](#_Toc416274770)

[1.5总结 4](#_Toc416274770)

[2.多元属性 4](#_Toc416274777)

[2.1主体属性 5](#_Toc416274778)

[2.2客体属性 5](#_Toc416274779)

[2.3环境属性 6](#_Toc416274780)

2.4可视性属性…………………………………………………………………………………………………6

[3.基于多元属性的动态访问控制机制 8](#_Toc416274784)

[3.1访问控制流程 9](#_Toc416274788)

[3.2控制策略表达式 10](#_Toc416274788)

[3.3策略树与LSSS矩阵 11](#_Toc416274788)

[3.4用户权限验证 14](#_Toc416274788)

[4验证实验 1](#_Toc416274784)5

[4.1访问控制模块 15](#_Toc416274791)

[4.1.1访问控制功能模块组成 1](#_Toc416274793)5

[4.1.2访问控制功能模块作用 1](#_Toc416274793)6

[4.2实验结果 1](#_Toc416274791)6

[4.2.1主体属性测试 16](#_Toc416274793)

[4.2.2客体属性测试 18](#_Toc416274793)

[4.2.3环境属性测试 20](#_Toc416274793)

[4.2.4可视性属性测试 22](#_Toc416274793)

[5.总结 2](#_Toc416274784)3

# 引言

随着全球互联网的建立以及计算机技术的飞速发展，基于信息技术的政务、商务、军事以及科研活动的应用领域逐渐扩大，形成了庞大的信息网络。在此基础上，各个领域和部门之间的协作日益增多，地域分散的组织通过信息网络实现了资源共享，也促使了信息系统从集中式向多域分布式的进化，如何保证多域分布式环境下资源的安全共享与互操作，已成为亟待解决的重要议题。

在联合信息环境中，动态有效的访问控制是保障共享信息平台安的全重要技术措施，访问控制实际上是通过制定访问控制规则或策略来允许或限制用户对数据或服务的访问，由于联合信息环境中的资源和服务的分布性、动态性、匿名性以及异构性等原因，要实现细粒度的访问控制变得更复杂，虽然传统IT运行环境下有很多模型和方法来实现细粒度访问控制，但这些模型和方法都要求拥有数据的所有者和存储服务提供者属于同一信任域中，但这一条件在联合信息环境下不再成立，数据提供者和数据访问者多数不属于同一信任安全域，数据服务提供商不能完全了解数据提供者和数据访问者的全部信息，特别是在跨域访问的情况下，更对访问数据的用户知之甚少，所以跨域访问中很难给数据访问者一个安全有效的数据访问授权。因此，传统的访问控制技术并不适用于联合信息环境，在联合信息环境中实施安全有效的访问控制将会面临一系列的挑战。

（1）用户与环境动态复杂性

数据提供者、数据访问者、信息资源和互联网环境都是动态的，是不断变化的，比如用户的身份及行为等都具有很大的不确定性风险，数据提供者针对数据访问的授权也是不断变化的，因此传统的基于身份的访问控制技术主要针对传统IT环境，无法联合信息环境的动态性需求和安全性需求。

（2）多域环境

联合信息环境具有多域特征，数据拥有者、数据服务提供者和数据服务访问者很可能位于不同的管理域，当用户跨域访问资源时，需要能够跨越不同域的复杂的授权认证系统，因此如何应对联合信息环境中的跨域访问控制问题非常重要，不同域要想实现资源共享，不仅需要制定域内访问授权策略，而且还需要处理跨域的访问控制授权策略问题，以实现联合信息环境中全方位的多域访问控制问题。

（3）细粒度授权机制

现有的解决细粒度访问控制问题的方法基本是为每一文件引入文件访问控制列表(ACL)用于细粒度访问控制，或将文件分成一些文件组来提高效率。然而随着系统用户规模的不断扩大，基于文件访问控制列表方法的复杂度也随之增大。另一方面，基于文件组的方法只能够提供粗粒度的数据访问控制。在联合信息环境下，现在最常采用基于角色的访问控制及其扩展模型技术来实现细粒度访问控制，但这种技术应用起来也遭遇瓶颈，一方面原因是在联合信息环境中需要定义大量的角色，这会给角色的分配和管理带来困难，另一方面是不同的数据服务提供者处于不同逻辑域，很难建立这种角色映射关系。因此目前的研究多集中在使用证书或基于属性的访问控制技术来满足大量数据的细粒度访问控制需求，然而这方面的研究大多都还处于起步阶段，尚未有成熟稳定的技术方案。因此对云终端用户进行细粒度的授权和控制是当前联合信息环境应用研究中挑战之一。

本课题面向联合信息环境，引入属性访问控制技术，研究设计一种适应于联合信息环境下基于多元属性的访问控制策略，针对允许访问的用户，并进一步根据其属性确定数据提供方式（见科技报告3），从而进行细粒度的访问控制。

# 1.访问控制策略研究现状

访问控制一直都是信息安全研究的重点问题，不论是传统的IT、互联网领域，还是物联网领域，访问控制在其中都扮演着重要的角色，对于保证服务器软硬件、资源和用户数据的安全性起着举足轻重的作用。

国内外研究者在数据与服务的访问控制领域进行了很多相关研究，也产生了很多的访问控制方式，如自主访问控制、强制访问控制、基于角色的访问控制、基于属性的访问控制等。这些访问控制机制也在传统的web应用、管理平台等方面得到了大量的应用，并在这些系统中发挥着重要的作用，为系统的安全做出了重要的贡献。

访问控制是每个成熟的系统都不可缺少的机制，它能有效的保护系统中的资源、程序以及设备，并能通过合理设置安全策略，达到控制访问、隔离访问者和数据的目的。同时，由于数据属性、用户属性和环境属性的多样化，在不同的环境下针对不同用户应为不同数据动态地选择数据提供模式，从而提高保护机制的安全性。



图1-1 经典访问控制模型

## 1.1自主访问控制

Discretionary Access Control (DAC，自主访问控制)，其基本思想是系统中的主体（用户或者资源的所有者）可以根据自己的意愿，自主地将其拥有的对客体（资源或者服务）的访问权限全部或部分地授予给其它主体。自主访问控制是一种有效的访问控制机制，能有效保护客体不被非法或者越权访问。

虽然自主访问控制是一种有效的访问控制机制，也已经在不少的系统中实现了，但是该机制也有自己的弱点。首先，自主访问控制的客体访问权是掌握在客体的所有者手中的，这样使得访问权非常的分散，不利于集中管理。其次，访问权所有者可能不是专业人员，对访问控制不了解，因而不能有效的对访问权进行管理。最后，也是最重要的一点，自主访问控制中，访问权可以传递，即A将自己的访问权授予给B以后，B同样可以将从A那里获得的访问权授予给其他人。若访问权被传递出去，管理就会非常困难，并且会带来严重的安全问题。

## 1.2强制访问控制

Mandatory Access Control (MAC，强制访问控制)，其基本思想是系统中的主体和客体都会被管理员事先分配一个固定的安全属性，而这个属性是不能被除了管理员以外的人修改的。这个安全属性决定了一个主体是否可以访问一个客体。每次访问中，系统检测安全属性以便确定一个主体是否有权访问该客体。文献中提到，强制访问控制最先使用于军事应用中，并且常与自主访问控制相结合使用。

强制访问控制在自主访问控制的基础之上增加了对客体属性的划分，并通过访问之前对访问权限和安全属性的匹配判断是否允许访问的进行。某种程度上讲，强制访问控制提高了访问控制的安全性，可以防止非法访问。但是这种访问控制过于严格，过于复杂，由于每个主体对每个客体的安全属性都会有一个匹配值，会产生N\*N的一个对应关系，造成管理上的不便，影响系统的灵活性。

## 1.3基于角色的访问控制

Role based Access Control (RBAC，基于角色的访问控制)，其基本思想是，建立一个角色集合，并赋予角色一定的访问权限。一旦用户被分配了某个角色后，该用户就拥有了此角色的所有访问权限。基于角色的访问控制在主体和客体之间加入了角色层，将主体和客体分隔开来。使得新建用户时只要分配相应的角色即可，而不需要赋予多种访问权限，而且角色的修改相比于用户的修改要少很多，有利于简化权限管理，减少系统开销。

基于角色的访问控制是一种普遍适用、更为有效的访问控制方式，也是访问控制领域的研究热点，其基于访问者身份进行授权，既符合常理，又方便快捷[4]。文献[5]在RBAC的基础上，加入了属性，使得能够较好的进行更细致的访问控制，但将其仍有一些不足之处，不能支持开放环境的访问控制。例如：若非法访问者，获取了合法的身份信息，便可获得对应的全部权限；若获得了某个角色，便可以无限制地使用指派给该角色的所有权限，这样容易造成权限滥用，不符合最小权限原则。总之，RBAC不适用于开放的环境，也很难做到细粒度的访问控制。

## 1.4基于属性的访问控制

Attribute Based Access Control (ABAC，基于属性的访问控制)，其主体思想是利用主体、客体以及所处环境等相关的属性来进行访问控制和授权，通过属性来对服务访问进行限制。由于ABAC不是直接利用主客体之间的关系，因而可以适应开放的网络环境。

与传统的访问控制不同，基于属性的访问控制是基于请求者和资源的属性来进行授权，因而其具备灵活性、可扩展性的特点，能够有效支持匿名访问。文献[6]提出了提升ABAC安全和性能的方法，使得ABAC的效率和安全性有了进一步的提升。

基于属性的访问控制(Attribute-Based Access Control ABAC)思想源于信任管理及其扩展，文献[72]和[73]分别提出基于主客体属性的访问控制矩阵模型ABAM和用有限集合论描述主、客体属性的访问控制逻辑框架LABAC，二者对属性描述均为粗粒度，也没有给出访问控制具体实现方法。ABAC模型考虑用户、资源、操作、执行动态请求时的运行上下文与授权相关的各种属性，细化了访问控制的粒度，ABAC模型及其扩展模型[74\_79]能够根据用户、资源、动作和运行上下文的属性动态的进行授权，具有更好的灵活性和扩展性。

## 1.5总结

本节对上述访问控制机制的优越点进行总结，以期为基于多元属性的动态数据服务访问控制技术的设计与实现打下基础。

表1-1 常见访问控制机制优缺点对比

|  |  |  |
| --- | --- | --- |
| 访问控制策略 | 优点 | 缺点 |
| 自主访问控制 | 1、简单直观 | 1. 管理分散； 2. 授权困难； 3. 效率低下； |
| 强制访问控制 | 1. 管理策略简单 2. 安全性高 | 1. 缺乏灵活度； 2. 应用领域窄 |
| 基于角色的访问控制 | 1. 安全性高； 2. 管理方便； 3. 系统开销小； | 1. 不适合开放式环境； 2. 控制粒度粗； 3. 缺乏动态性； |
| 基于属性的访问控制 | 1. 适合开放式环境； 2. 灵活性； 3. 可扩展性； | 1. 动态性； 2. 细粒度； |

表1-1对各种访问控制机制做了对比。可以看出，相比于自主访问控制、强制访问控制、基于角色的访问控制机制，基于属性的访问控制策略可以使服务开放系统针对不同的场景提供有效的服务，同时支持提供匿名访问。

本课题将在基于属性的访问控制机制的基础上，梳理数据服务上下文，进而提炼多元属性集，基于多元属性研究并实现动态和细粒度的访问控制。

# 2.多元属性

网络环境下，我们希望数据的提供是动态的，对于某一数据，可以根据数据访问者个人信息的不同，或是当前访问环境的优劣情况，计算出不同的数据提供策略，从而提供不同的数据服务。细粒度，则是说对于不同类型的数据或者不同级别的数据，甚至是同一数据内不同字段的数据，按照不同的方式来呈现，将数据服务细化到数据字段级别。

传统的访问控制，如自主访问控制控制，是针对单个客体授权的，并且授权是可以传递的，从而使授权的分配变得复杂，且不利于管理；强制访问控制对于每个访问者和资源都需要设定权限，灵活性太差，因此其并不适合于大数据环境中数据多、且关系复杂的情况；基于角色的访问控制不能对数据的访问进行动态的细粒度划分，也不能很好的支持开放的环境。通过将访问者的行为信任应用到基于角色的访问控制中动态指派访问者角色，使其能动态分配访问者权限，但仍不能动态地对数据进行细粒度划分。

不同于传统的访问控制模型（DAC，MAC以及RBAC等），基于多元属性的访问控制（ABAC）策略是由主客体等多方面的属性来决定的。因此，需要对数据服务访问上下文属性进行划分和定义。

属性在这里指可以描述用户、物品等性质的特征，包括：主体属性、客体属性、环境属性和可视性属性。各种属性按如下方式定义。一般来说，主体都是指用户，即发起数据服务请求的人。用户都是有自己特定的一系列属性，可能包括用户ID，用户名，邮箱等等，在本课题的研究中，各部门间的数据交流与共享是个重要的部分，因此，主体属性还包含部门属性和职务属性。客体属性是指数据相关的属性，例如，数据的上传者和隶属部门，数据的安全级别、数据类型、数据大小等。环境属性是指用户发出数据访问请求时的环境，包括访问的发起地点（经纬度）、IP地址、接入网络类型。可视性属性是指数据在数据链中可视范围，包括全程可见性、上游可见性、下游可见性和全程不可见性，可用于确定操作的作用范围。多元属性的具体划分情况如下。

## 2.1主体属性

主体属性Subject：用于描述访问者自身信息，主要包括用户id、所属部门、职务、是否是超级用户、登录次数、用户行为记录等。

在大数据平台上，我们希望服务和数据资源能够被尽量多的人访问，但同时也希望对不同的访问者进行限制，区别其能获取的服务和数据。访问者多种多样，从系统的角度看，存在管理员和其他用户两类访问者，其他用户又可以根据其信任度的不同，划分为游客、普通用户、超级用户等，不同级别的用户具有不同的行为权限。从而实现对访问者身份的验证。

## 2.2客体属性

客体属性Object：用于描述被访问资源的信息。是对被访问信息（主要是各种文本、图片数据）的重要性进行的数据化的抽象。它主要用于在对外提供数据的时候进行限制，以使得被访问信息得到不同程度的保护。包括数据dataid、数据类型、数据发布者id、归属部门、安全级别等。

平台上的数据主要分为两大类，一类是以表格的形式存储在数据库内的数据，另一类则是以文件的形式存储在服务器端，包括图片、word和pdf文档、excel表格等。对于不同类型的数据，具有不同的数据提供方式。选择合理有效的数据提供方式既能保证用户的高效的细粒度访问，又能保证大数据平台下数据的安全性和数据链整体的鲁棒性与一致性。

不同的数据，其重要程度可能是不一样的，针对不同数据，需要考虑其在系统和数据链中的重要程度，从而为其设定一个合理的安全级别，以区分重要程度不同的数据。特别地，对于第一类表格类型的数据，其安全级别甚至细化到表格内的各个字段。

用户或管理员在上传发布数据时，可根据实际情况和自身需求对数据和字段的安全级别进行划分，设置不同的安全级别，如表2-2所示。从而实现细粒度的访问控制和数据提供。

表2-1 数据安全级别划分

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **安全级别** | 0 | 1 | 2 | 3 |
| **数据** | 低级数据 | 普通数据 | 中级数据 | 高级数据 |

表2-1 数据安全级别划分

## 2.3环境属性

环境属性Environment：是对访问者发起服务访问时所处的环境的信息进行的数据化的抽象，用于描述访问者所处环境的基本信息，，包括访问者的IP地址、网络环境（有线/无线）、经纬度等。

对于大数据平台而言，用户对服务的访问可能处于任何环境，这里的环境包括访问者所处的网络环境、访问时的IP地址、地理位置和访问时间等环境属性，且这些属性的安全程度对于不同的数据所有者而言也会不一样。

例如，某个数据发布者在发布数据时，为自己发布的数据设置ip安全策略，策略认为在某个特定的IP范围内发起的针对自己数据的访问是安全的。当用户发出数据服务请求时，系统会提取其环境属性内的ip属性值，并与安全策略进行比较，从而判断数据请求者的环境优劣情况，并根据环境情况选择不同的数据提供方式。这就需要提取访问者所在环境的信息，构建环境属性集，用于基于属性的访问控制的计算。

## 2.4可视性属性

数据可视性属性Visibility，用于描述用户之间的关系，其主要体现在大数据平台中各种数据链上存在的上下游关系。

在大数据平台中，存在着各式各样不同的数据链，每条数据链的各个节点上的数据所有者和用户都可能是不一样的，这些数据所有者和用户之间可能存在不同的关系，他们可能会倾向于通过这种关系来进行一些访问限制。基于数据链模型，可将用户分为上游用户和下游用户，因而可以按照全程可见策略、上游可见策略、下游可见策略及全程不可见策略四种模式研究用户对数据的访问权限。基于这四种权限，针对数据链存在分链和合链的复杂情况，进而研究多链并存的情况下权限管理策略的优化，以将数据链的可见性关系用于访问控制。

全程可视策略，即数据发布者发布的数据可以被数据链中的所有数据使用者访问。如图2-1所示，用户R1发布的数据可以允许数据链中的所有用户访问。



图2-1 基于全程可视策略的数据访问控制

上游可视策略，即数据发布者发布的数据只允许数据链上游的数据使用者所访问。如图2-2所示，R1发布的数据只允许被数据链上游的的P1和Q1所访问。



图2-2 基于上游可视策略的数据访问控制

下游可视策略，即数据发布者发布的数据只允许数据链下游的数据使用者所访问。如图2-3所示，R1发布的数据只允许被数据链下游的S1所访问。



图2-3基于下游可视策略的数据访问控制

相邻可视策略，即数据发布者发布的数据只允许数据链上相邻的数据使用者所访问。如图2-4所示，R1发布的数据只允许数据链上相邻用户访问。



图2-4基于相邻可视策略的数据访问控制

在数据服务多元属性划分与定义的基础之上，本课题将访问控制的数据服务模型定义为如下四元组形式：

Sub表示主体属性；Obj表示客体属性；Env表示环境属性；Vis表示可视性属性。

图2-5 基于多元属性的数据服务本体

如图2-5所示，是本课题建立的基于多元属性的访问控制数据服务模型。在该模型描述中数据服务由主体属性、客体属性、环境属性、可视性属性四部分组成，这四个方面分别代表四个不同的属性侧重方面，主体属性侧重于对服务用户身份的描述，客体属性侧重于对数据属性的描述，环境属性侧重于对数据访问环境的描述，可视性侧重于对数据的服务范围的描述。四类属性共同构成了参与访问控制的属性集，并以此来建立访问控制数据服务模型，为平台上访问控制机制的设计实现奠定了基础。

# 3.基于多元属性的动态访问控制机制

基于多元属性的动态访问控制机制是大数据平台下提供动态数据服务的基础。本章将具体介绍该机制的设计和计算过程，包括访问控制总体流程、访问控制策略树和LSSS属性矩阵、用户属性子矩阵验证算法等。

## 3.1访问控制流程

用户登录平台并发出访问请求，向访问控制模块提供自身的uid，所请求数据的dataid，以及当前所处的环境属性集合。访问控制功能模块将提取用户本次访问请求时的各类属性，构建多元属性集合，根据访问控制属性集合提取访问控制策略，匹配策略得到用户属性子矩阵，计算其是否有解，从而判断用户是否满足访问控制策略。在本系统内，我们采用黑名单策略，若用户满足访问控制策略，则表明用户为黑名单用户，拒绝其本次访问请求，否则允许服务，跳转至数据提供功能模块，计算数据提供策略，向用户返回安全、合理、细粒度的数据服务。总体流程图如下。

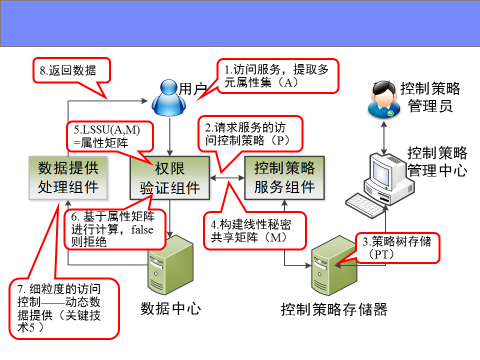


图3-1基于多元属性的访问控制总体流程图

接下来对于图中的各个步骤进行说明。

Step1用户发出数据服务访问请求，系统提取用户自身的相关属性、请求的数据服务的相关属性和当前访问请求的环境属性等多种属性，构建访问控制多元属性集A。

属性集A包含了主体属性、客体属性、环境属性和可视性属性等四部分属性，是访问控制机制的基础模型。系统对用户进行访问控制，首先就需要从属性集A内提取所需要的相关属性，然后才能进行安全合理细粒度的权限验证。

Step2根据数据服务的客体属性，请求访问控制策略P。

Step3从控制策略存储器内提取访问控制策略树PT。

用户在大数据平台上发布数据时，可根据数据属性和自身需求设置访问控制策略，或者可由系统的管理员来设置符合需求的访问控制策略。控制策略管理中心调用相关的功能模块对控制策略进行规范化描述，并以布尔表达式的形式存储在控制策略管理中心内，作为该数据访问控制的依据。

Step4将访问控制策略树PT转化成LSSS属性矩阵M。

根据LSSS矩阵构建，将访问控制策略P生成的策略树PT转化成LSSS共享矩阵M，用于和属性集A内的属性进行验证计算。

Step5采用LSSU（A,M）算法，将多元属性集A和LSSS属性矩阵M进行匹配，构建用户属性子矩阵M’。

LSSU（A,M）算法是访问控制的核心算法，在step4中，权限验证组件得到了由控制策略服务组件内返回的属性矩阵M和属性集π(x)。以属性集π(x)为索引表，从访问控制模型属性集A内提取索引表π(x)中包含的属性A’，A’内可能包括索引表π(x)内的全部属性，也可能只包含某几个属性，甚至可能为空的属性集。根据索引表π(x)，从属性矩阵M内提取A’内各个属性的向量值，并将向量值按顺序进行组合，构建一个新的矩阵M’，M’即为访问控制的用户属性子矩阵。显然，M’为属性矩阵M的子矩阵。

Step6验证子矩阵M’，若满足访问控制策略，则拒绝服务，否则跳转至数据提供功能模块。

调用用户权限验证算法，判断用户是否满足当前的访问控制策略。在本系统内，由于采用的是黑名单策略，因而当用户满足访问控制策略，表明用户为黑名单用户，因而将拒绝他的访问控制请求。否则，将允许用户进行访问，跳转至数据提供处理组件，计算数据的安全提供策略。

Step7计算数据提供策略。

用户通过访问控制权限验证，进行数据访问时，系统并不是简单的讲数据返回给用户，而是会根据当前的访问环境优劣情况、访问者与数据的隶属关系、当前用户的信任度、数据的可视性属性等多元属性，计算数据服务等级，服务等级不同，则数据提供方式不同。当服务等级较高时，数据将采用直接返回给用户，采用明文不加密的方式；当服务等级较低时，数据将进行保留格式加密或是隐藏在模版内，返回给用户的是加密隐藏后的数据，用户可在更为安全的环境下，通过自身的密钥对数据进行解密和恢复；当服务等级特别低时，数据可能进行高级的隐藏或是AES加密的方式进行处理，加密后的数据无法恢复，从而避免数据的泄露。详细内容见科技报告3。

Step8向用户返回安全、合理、细粒度的数据服务。

## 3.2控制策略表达式

访问控制策略是访问控制的基础和核心，对于访问控制功能模块，合理安全的访问控制策略是不可或缺的内容。为了实现动态、安全、细粒度的访问控制，本系统对访问控制策略进行规范化的描述，采用布尔表达式的形式将包含多个属性的访问控制条件组合成策略表达式。

布尔表达式是[布尔运算](http://baike.baidu.com/view/638530.htm)量和逻辑运算符按一定[语法规则](http://baike.baidu.com/view/4269308.htm)组成的式子。逻辑运算符通常有∧（与）、∨（或）、﹃（非）三种（在某些语言中，还有≡（等价）及→（蕴含）等等）；[布尔运算](http://baike.baidu.com/view/397415.htm)对象可以是[逻辑值](http://baike.baidu.com/view/379053.htm)（True 或False）、[布尔变量](http://baike.baidu.com/view/2792492.htm)、关系表达式以及由括号括起来的布尔表达式。

不论是[布尔变量](http://baike.baidu.com/view/2792492.htm)还是布尔表达式，都只能取[逻辑值](http://baike.baidu.com/view/379053.htm)True或False。在最基本的层次来说，所有的布尔表达式，不论表达式的长短、逻辑符的复杂程度大小和布尔运算对象的数量多少，其最终的输出值只能是true或false。在计算机内通常用1（或非零整数）表示真值（True），用0表示[假值](http://baike.baidu.com/view/1958571.htm)(False)。

对于访问控制机制，用户与策略之间无论有怎样复杂的关系和计算过程，最终结果只有允许或拒绝两种情况，即true或false。

因而，本系统内的访问控制策略统一采用布尔表达式进行设计，属性条件作为布尔运算对象，与逻辑运算符∧（与）、∨（或）、﹃（非）进行组合，规范化描述访问控制策略。

系统的访问控制功能模块内预先划分了多个属性集，所有属性集共同构成了属性域Z：

Z=：…………………………………………….(3-1)

注：Z为属性域，为属性集，为单个属性

设计访问控制策略时，数据发布者或管理员从多个属性集内提取所需的单个属性条件，通过逻辑属性条件组合成布尔表达式，即为访问控制策略，形式示例如下：

P：……………………………………………(3-2)

## 3.3策略树与LSSS矩阵

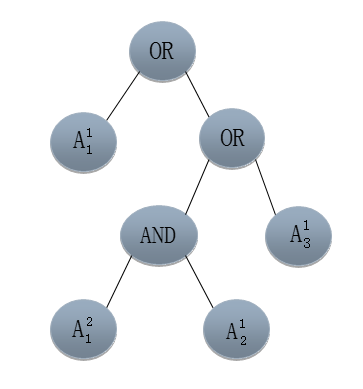
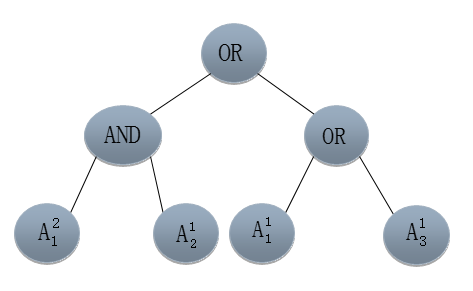
基于多元属性的访问控制方案的关键是如何将访问控制策略内的多元属性进行拆分，并将每一个属性映射到属性集上，其研究的理论基础来自于（n、t）门限秘密共享，门限秘密共享在与访问策略相结合时有两种基本形式，一种是基于拉格朗日插值法的秘密共享树，另一种是基于线性秘密共享机制的共享矩阵。基于LSSS( Linear Secret-Sharing Schemes) 的共享矩阵将访问策略和共享秘密转化到一个跟参与方属性关联的矩阵中，通过矩阵运算减少了大量秘密恢复时的计算量，比秘密共享树的解密效率高，因此得到了广泛应用。

3.2.1 策略树构建

一般而言，门限秘密共享在解秘时存在三种情况，分别为t<n,t=n和t>n。t表示当前拥有的子秘密数，n表示要恢复秘密所需的最小子秘密个数。在基于多元属性的访问控制机制内，控制策略的形式为布尔表达式，将单个属性视为子秘密，策略属性组合只存在And和Or两种情况，因而可将门限策略简化为t=n和t=1，并据此来构建访问控制策略树。

访问控制策略树为二叉树，每一个叶子结点都是一个属性，非叶子结点为门限条件，有And和Or两种情况。对于一条访问控制策略，可根据策略布尔表达式内的逻辑符号对其进行拆分，每一个∨(或)运算连接两条子策略，每个∧(与)运算连接一条子策略内的两个条件。每一个Or门限的子树可视为子策略，每一个And门限的子树视为同一子策略内的条件，构建访问控制策略树。

对于同一访问控制策略，其策略树可能有多种形式，如图3-2、3-3所示的两个策略树均由策略(3-2)转换而成，二者在逻辑上是等价的。

  图3-2策略树(1) 图3-3策略树(2)

3.2.2 策略树转换为LSSS矩阵

对于一个访问控制策略树，叶子结点为策略内包含的属性，遍历树的叶子节点，记录每一个叶子结点所代表的属性，得到策略树的属性集π(x)。

计算属性向量时，我们将根结点的向量值预先设为（1），任一非叶子结点向量值为v[x]，如果结点的门限策略为Or（t=1），则该结点的左右两个结点均直接继承父结点的向量值v[x]；如果结点的门限策略为And（t=n），则该结点的左子结点向量值为(v[x]|1)，右子结点的向量值为(0,0 . . . ,−1)。从根结点（1）开始，自顶向下遍历整个策略树，得到每一个叶子结点的向量值。最后，根据最大的向量长度，在其余向量后补0，使所有向量具有相同的长度。

以属性集π(x)作为索引表，根据π(x)内属性的顺序构建一个x\*L的矩阵M，矩阵的行数x表示属性个数，列数L表示属性向量的长度。矩阵M就是我们由访问控制策略的布尔表达式转换而来的LSSS矩阵。

具体计算过程如算法1所示：

|  |
| --- |
| 算法1：LSSS矩阵构建算法，根据策略树PT计算属性向量值，构建LSSS矩阵。 |
| 输入：策略树T（每个结点有0或2个子结点，非叶子结点值为and或or，叶子结点为属性）  输出：属性集π(x)，属性矩阵M |
| 第1步：遍历叶子结点，按顺序构建属性集π(x)  第2步：遍历策略树，计算属性向量，伪码如下所示：  T 是一个树结点，包括T.node(结点值: and, or),T.left, T.right, T.vector ,T.length  // so v(x)=T.vector  {  L=0 //向量长度  maxL=0 //最大向量长度  IF T is a leaf //T.left=NIL; x=T;  return T.vector= empty// v(x)=T.vector  ELSE  ChildV(T, maxL);  Padding(T, maxL)  }  Function ChildV(T, maxL) {  IF T.node=Or{  OrChild(T.left, T.vector, maxL);  OrChild(T.right, T.vector, maxL)  }  IF T.node=And{  AndLeftChild(T.left, T.vector, maxL);  AndRightChild(T.left, T.vector, maxL)  }  }  Function OrChild(X:tree, Y, maxL) {  X.vector=Y // 继承父结点向量  padding for the root node  ChildV(X, maxL)  }  Function AndChildLeft(X:tree, Y, maxL) {  X.L= X.L+1 //向量长度+1  IF X.L> maxL{  maxL=X.L //increases maxL if needed  X.vector = Y | 1 //左子结点向量=（父节点向量|1）  the end of vector and returns it  }  ChildV(X, maxL)  }  Function AndChildRight(X;tree, Y, maxL) {  X.L= X.L+1;  IF X.L> maxL  maxL=X.L;  X.vector=(0ˆ{X.L-1},-1)//右子结点向量=（0,0,0…-1）  ChildV(X, maxL)  }  Function Padding(T, maxL) {  IF L< maxL  { add maxL-L 0’s at the end of v(x); L=maxL};  IF x is not a leaf{  Padding(T.left, maxL);  Padding(T.right, maxL)  }  }  第3步：以π(x)为索引表，由第2步得到的属性向量生成一个x\*maxL的矩阵，x为属性个数，maxL为最大向量长度。该矩阵即为LSSS属性矩阵M。  第4步：输出属性矩阵M。 |

以3.2.2小节策略(3-2)构建的策略树(1)为例，通过计算计算得到的各个属性向量值如图3-4所示，属性集π(x)如图3-5所示：

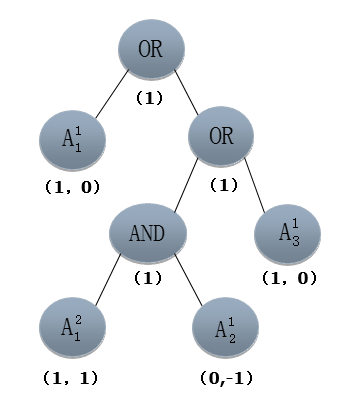


图3-4属性向量

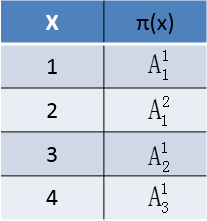


图3-5属性集π(x)

以属性集π(x)为索引，构建属性矩阵M，如图3-6所示：

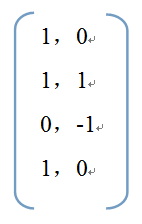


图3-6属性矩阵

## 3.4用户权限验证

验证用户是否满足访问控制策略，本质上就是判断当前的属性集内用户各个属性与是否满足控制策略内的属性条件，也就是对用户的属性向量进行计算后，能否得到单位向量（1），即策略树中根结点的向量。而求解的本质就是是否能找到一组常量值，使得线性方程组AX =ε有解。

通过LSSS(A,M)算法，匹配多元属性集A和访问控制策略矩阵M，得到属性子矩阵M’。求解线性方程组X=ε，为属性子矩阵M’的转置矩阵，ε为单位矩阵。根据线性代数的相关知识可知，对于非齐次线性方程组X=ε，如果方程的增广矩阵(, ε)的秩R和矩阵的秩R’相等，则线性方程组有解。

具体计算过程如算法2所示：

|  |
| --- |
| 算法2：访问权限验证算法，根据LSSS(A,M)算法得到用户属性子矩阵M’，求解线性方程组X=ε，判断用户是否符合访问控制策略。 |
| 输入：多元属性集A，属性矩阵M，策略属性集π(x)。  输出：true或false（用户是否符合访问控制策略） |
| 第1步：以属性集π(x)为索引，建立多元属性集A和属性矩阵M的映射关系，从A内提取π(x)内包括的属性，其余属性则不参与访问控制策略的计算。  第2步：根据第2步建立的映射关系，从矩阵内提取属性对应的行，即属性的向量值，生成新的矩阵M’，M’为属性矩阵M的子矩阵。  第3步：求解线性方程组X=ε，通过判断矩阵M’和增广矩阵（M’|ε）的秩是否相等来判断方程组是否有解。  第4步：如果方程组有解，则用户符合访问控制策略，返回true。否则，返回false |

# 4.验证实验

为了验证上述访问控制策略的正确性和有效性，我们搭建了大数据平台，平台上存储了多种类型的数据，用户登录平台，发送服务请求和数据访问请求，在这个过程中，平台内的访问控制功能模块将保证用户访问的有效性和数据的安全性。

## 4.1访问控制模块

### 4.1.1访问控制功能模块组成

大数据平台的访问控制模块主要由以下几部分组成：

1、数据服务属性集构建模块。数据服务属性集是访问控制模块的核心模块，当用户登录平台，提出访问请求时，提取主体属性、客体属性、环境属性和可视性属性等多元属性，构建访问控制服务模型，并用于之后的访问控制计算。

2、控制策略存储器。逻辑上以二叉树的形式存储由数据发布者或系统管理员根据需求设置的访问控制策略。

3、控制策略服务组件。从存储器内提取访问控制策略，将策略转化为LSSS共享矩阵，得到矩阵M和属性集π(x)。

4、权限验证组件。验证用户访问权限，设计LSSU(A,M)匹配算法，将数据服务属性集A与矩阵M进行匹配，得到用户属性子矩阵M’。通过判断线性方程组X=ε是否有解来判断用户是否符合访问控制策略。

5、数据服务中心。计算数据提供策略，并返回给用户安全、合理、细粒度的数据服务。

### 4.1.2访问控制功能模块作用

平台通过访问控制服务模块，提取多元属性，构建访问控制服务模型，设计和计算基于多元属性的动态数据服务访问控制策略，实现了以下三个目的：

1、对访问者开放服务，屏蔽内部具体实现。

2、重视数据安全，实现数据动态、细粒度访问控制。

3、系统功能模块化，解耦合，保证系统健壮性、可维护性和扩展性。

## 4.2实验结果

本节将对设计、实现的权限控制信息服务进行测试，以期发现错误，提高性能。在任何系统的开发中，测试都是其中很重要的一部分。测试有助于发现系统的漏洞和不完善的地方，有助于改进系统功能、提高系统性能。本系统主要从功能方面进行测试。对于本系统的测试，将基于下述的用户进行。

用户userA-2，属性包括{总装备部、士兵、会员、ip地址、网络环境、地理位置}等。

用户userA-3，属性包括{总后勤部、参谋、会员、ip地址、网络环境、地理位置}等。

数据表cpi，数据来源为总装备部，访问控制策略为“禁止非总装备部用户访问，禁止通过无线网络访问”。

数据表estate，数据来源为大数据平台总部，访问控制策略为“禁止士兵访问”。

图片image2.bmp，数据来源为总装备部，访问控制策略为“禁止非总装备部用户访问”，数据可视性属性为下游可见。

### 4.2.1主体属性测试

在同样的访问环境下，用户userA-2和用户userA-3同时访问同一数据表cpi，访问结果如下所示。



图4-1 userA-2



图4-2 userA-1访问数据表cpi测试结果



图4-3 userA-3

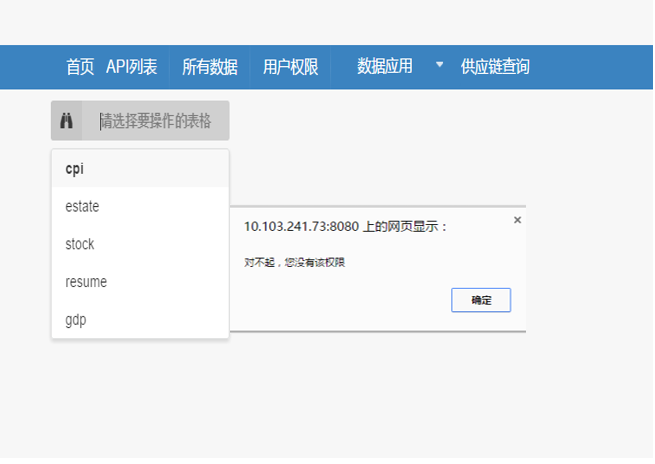


图4-4 userA-3访问数据表cpi测试结果

userA-2访问数据表cpi时，通过访问权限验证，系统允许其访问请求，返回数据表cpi。

userA-3访问数据表cpi时，无法通过访问权限验证，系统拒绝其访问请求。

### 4.2.2客体属性测试

在同样的访问环境下，用户userA-2分别访问不同的数据表cpi和estate，访问结果如下所示。



图4-5 userA-2访问数据表cpi测试结果



图4-6 userA-2访问数据表estate测试结果

userA-2访问数据表cpi时，通过访问权限验证，系统允许其访问请求，返回数据表cpi。

访问数据表estate时,无法通过访问权限验证，系统拒绝其访问请求。

### 4.2.3环境属性测试

在有线网络无线网络环境下，用户userA-2分别访问同一的数据表cpi，访问结果如下所示。



图4-7 访问环境（1）



图4-8 环境（1）下userA-2访问数据表cpi测试结果



图4-9 访问环境（2）

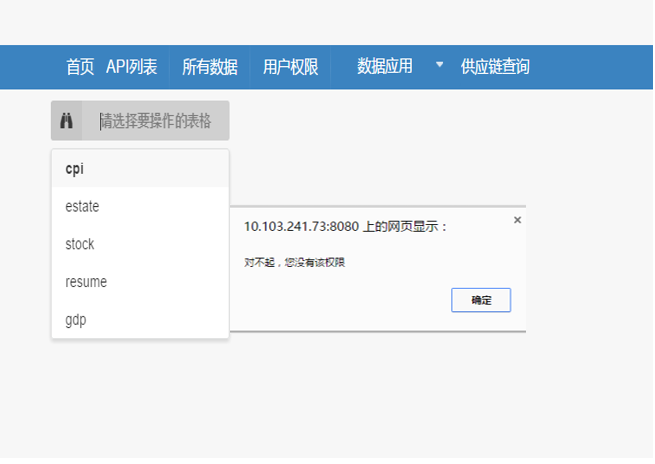


图4-10 环境（2）下userA-2访问数据表cpi测试结果

用户userA-2通过环境（1）有线网络访问数据表cpi，通过访问权限验证，系统允许其访问请求，返回数据表cpi。。

用户userA-2通过环境（2）无线网络访问数据表cpi，无法通过访问权限验证，系统拒绝其访问请求。

### 4.2.4可视性属性测试

图片image2.bmp在数据链上的流向为平台A->平台B->平台C->平台D，在数据平台内的编号为platA-12。用户userA-2在有线网络环境下请求查询图片image2.bmp的数据链流向和使用情况，访问结果如下所示。

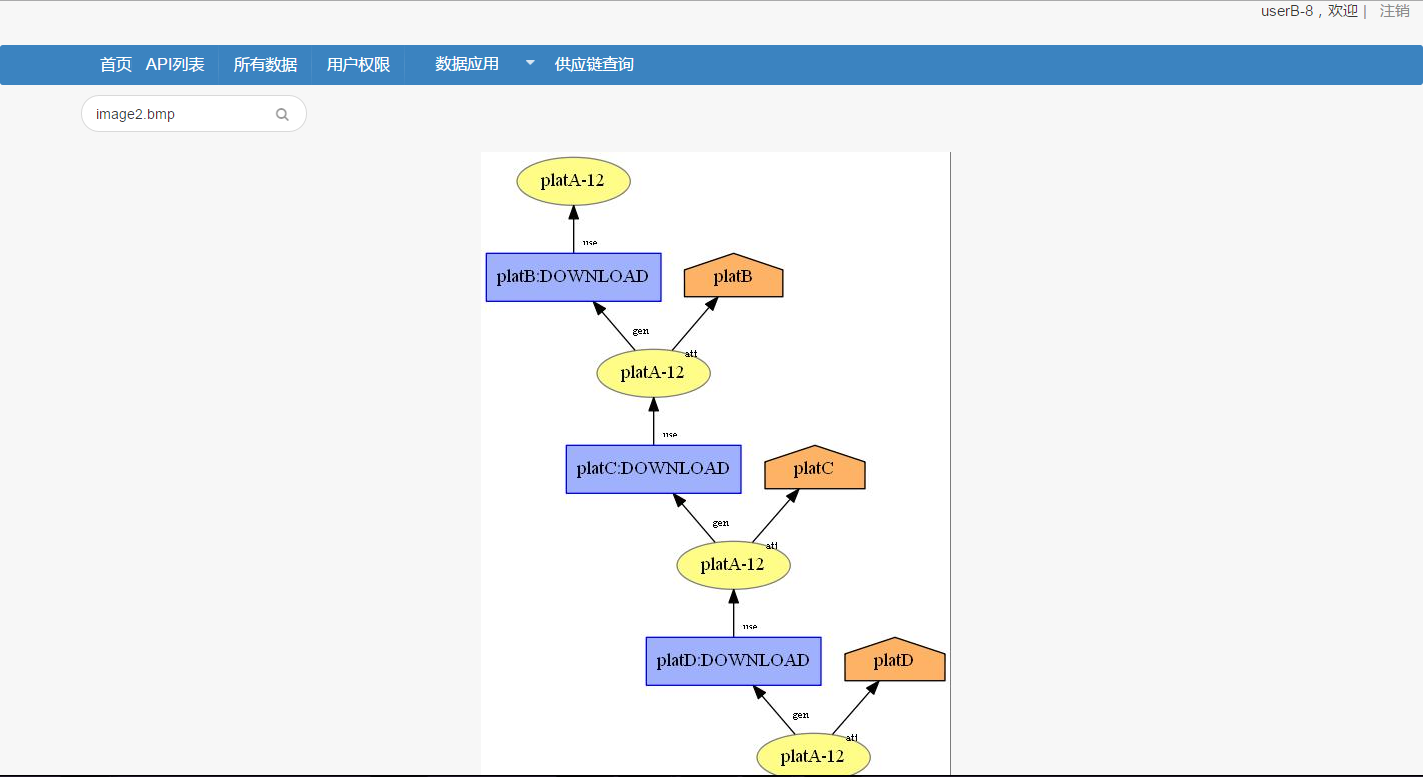


图4-11 userA-2查询图片image2.bmp数据链测试结果

用户userA-2查询图片image2.bmp数据链流向和使用情况，通过访问控制权限验证，系统向用户返回数据下游流向及使用情况。

# 5.总结

随着各个领域和部门之间的协作日益增多，地域分散的组织通过信息网络实现了资源共享，如何保证多域分布式环境下资源的安全共享与互操作已成为一个亟待解决的问题。针对该问题，本课题将大数据平台作为研究和开发背景，认真研读和分析了现有的数据与服务的访问控制方法，设计了一种新的基于多元属性的访问控制策略，并进行了以下总结和研究。

1.首先介绍了基于属性的访问控制机制中属性的划分以及相关的概念与定义。将属性按照所属关系不同划分为环境属性、主体属性、客体属性以及数据可视性属性，在此基础上构建访问控制服务模型。

2.基于逻辑运算规则，对访问控制策略进行规范化处理，转化为布尔表达式P。并根据门限秘密共享机制，将访问控制策略转化为二叉树的形式，得到策略树PT。然后根据LSSS矩阵构建算法计算属性向量，构建属性矩阵M。

3.最后，用属性集表示用户访问请求的属性模型，根据当前访问控制模型内多元属性的值，提取属性矩阵M和策略属性及π(x)。根据LSSU（A,M）匹配算法得到用户的属性子矩阵M’，判断线性方程组X=ε是否有解，以此来判断用户是否符合控制策略。在系统中，我们采用了黑名单策略进行权限验证，如果用户符合控制策略，表明用户为黑名单用户，系统将拒绝其访问请求。否则，通过验证，允许该请求，选择合适的数据提供方式向用户返回安全合理细粒度的数据服务。