计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.191000066

基于 TFR 模型的公安云平台数据分级分类安全访问控制模型研究

顾荣杰　吴治平　石　焕
公安部第三研究所　上海 201204
(rongjiegu@vip.163.com)

摘　要　近年来,公安大数据建设不断提速,各地数据中心的统一建设带来敏感数据的高度集中,涉及国家安全和公民个人信息的泄露和违规使用的风险急剧上升。在数据加密存储、角色访问控制等传统方法的基础上,提出了一种新的基于数据治理属性分级分类的访问控制模型。通过对数据敏感性、人员、数据进行分级分类,该模型实现了基于数据表、字段、数据记录级别的分层控制,有利于实现灵活度更高、颗粒度更细的公安敏感数据的分级分类精准访问授权控制,可有效应用于当前智慧公安大数据云平台数据访问安全控制体系的构建。该模型已实际应用于部分地区的智慧公安建设中,并取得了较好的成效。

关键词:大数据;公安云;敏感数据;分级分类访问控制;授权访问
中图法分类号　TP391

New Approach for Graded and Classified Cloud Data Access Control for Public Security Based on TFR Model

GU Rong-jie, WU Zhi-ping and SHI Huan
The Third Research Institute of the Ministry of Public Security, Shanghai 201204, China

Abstract　In recent years, the development of big data for public security is accelerating. The unified construction of public security data centers around the country has brought about high centralization of sensitive data, thus the risk of leakage of information regarding national security and illegal use of personal information is sharply increasing. On the basis of traditional data security protection methods such as data encryption and role-based access control, this paper presents a new access control model based on data grade and classification. Based on the grading and classification of data sensitivity, personnel and data, this model can achieve hierarchical control based on the level of data table, data field and data record, which is helpful to achieve precise access authorization control of grading and classification for sensitive public security data with higher flexibility and finer granularity, and can be effectively applied to the construction of data access security control system of modern big data cloud platform for smart public security. This model has been applied to the construction of smart public security in some areas and has achieved satisfied results.

Keywords　Big data, Public security cloud, Sensitive data, Graded and classified access control, Authorized access

1　研究背景

2018 年 1 月,公安部正式将大数据上升为部级战略,在全国掀起智慧公安建设的新一轮高潮[1-3]。公安大数据建设的目标是建立统一的公安云,具体做法是将原先分散在各手段业务领域的各类关键、敏感数据以及来自科信和社会公开采集的数据,统一汇聚到统一存储、管理、应用的云计算平台上。各类敏感数据的高度集中存储,带来了大数据条件下涉及国家安全和公民个人信息的数据管理和访问控制的新挑战。一旦使用不慎,就会造成严重的泄密和公民隐私侵犯案件。因此必须建立严格的云数据资源授权访问机制,以确保数据访问安全,避免数据的滥用。

数据分级分类是对公安数据资源中的数据访问级别进行限定的基础和依据。通过数据资源的分级和分类,对涉及敏感内容、隐私内容等的记录和字段进行分级别的访问限制,防止敏感信息的扩散。

2　研究与应用现状

访问控制是信息安全领域最热门的研究方向之一,从自主访问控制(Discretionary Access Control, DAC)[4]到强制访问控制(Mandatory Access Control, MAC)[4],再到基于角色的访问控制(Role-based Access Control, RBAC)[5-7]以及基于属性的访问控制(Attribute-based Access Control, ABAC)[8],各类模型均有其特点和不足。早期的公安信息化系统应用 DAC 或者 MAC 实现访问控制。然而,DAC 经常导致不适当授权或权限撤销不及时,形成安全隐患;MAC 在灵活性上的天然缺陷,导致难以根据动态变化的案件研判或情报分析需求授予相关干警必要的访问权限,不利于公安干警开展工作。从 1990 年代后期开始,公安信息化系统广泛应用 RBAC。最近,为了支持更为动态的权限配置,ABAC 在公安大数据的安全保障方案中被广泛讨论。RBAC 和 ABAC 较好地平衡了安全性与灵活性,然而在大数据时代,面对公安海量的数据对象,应用 RBAC 或 ABAC 实现细粒度的访问控制时,角色构建或策略构建相当复杂,极易出现权限误配。

3　概念定义

公安云平台数据分级分类访问问题的业务需求是在云端数据共享的趋势下,根据不同用户所属部门、业务职责、权限

通信作者:吴治平(zhiping.wu@aliyun.com)

等级的不同，提供不同的资源视图和访问授权，具体的控制颗粒度需求分为数据表、数据记录和数据字段等 3 个级别。

3.1 数据表级别的访问控制

数据表是数据库最重要的组成部分。本文所定义的数据表是数据项(列)和数据记录(行)的集合。数据表级别的控制一般源于对不同的用户群体分类访问授权控制的需要，从而对不同表按照某种逻辑进行归类形成多个集合。

3.2 数据记录级别的访问控制

数据记录是指数据表中的每一行。每一条数据记录包含这一行中的所有信息。数据记录是数据字段实例化后的数值内容的集合。数据记录是动态增长的。数据记录级别的控制一般源于对相关敏感的人员身份、物品属性的访问控制需要，只有拥有足够级别权限且经过授权的人员才能访问。

3.3 数据字段级别的访问控制

数据字段是构成数据表结构最基本的单元。多数情况下，数据表的“列”称为数据项或数据字段。数据字段代表所有数据记录共有的属性。数据字段级别的控制一般源于对同一数据表中部分敏感字段的访问控制需要，如可以访问与某些人员有关的记录，但对其住址、联系方式的访问要进一步授权。

3.4 数据资源分级访问控制

数据资源分级是指根据数据表、数据字段、数据记录的敏感程度不同，对数据表、数据字段、数据记录赋予不同敏感级别的过程。数据表、数据字段根据其属性的敏感程度进行分级，数据记录根据其内容涉及的敏感程度(通常根据对象的影响力级别)进行分级。

按照信息系统通用安全技术要求国家标准 GB/T20271－2006 将资源划分为 5 个等级(见表 1)，或者根据公安部门的业务敏感度划分为敏感、限制和公开 3 个等级(见表 2)。

表 1 GB/T20271－2006 定义的五级敏感等级

Table 1 Five data sensitivity grades defined by GB/T20271－2006

数据信息类别	安全保护等级	分类标准
公开	第一级	受到破坏后，会对公民、法人和其他组织的权益有一定影响，但不会危害国家安全、社会秩序、经济建设和公共利益
内部	第二级	受到破坏后，会对国家安全、社会秩序、经济建设和公共利益造成一定危害
保密	第三级	受到破坏后，会对国家安全、社会秩序、经济建设和公共利益造成较大危害
机密	第四级	受到破坏后，会对国家安全、社会秩序、经济建设和公共利益造成严重危害
绝密	第五级	受到破坏后，会对国家安全、社会秩序、经济建设和公共利益造成特别严重危害

表 2 公安业务三级敏感等级划分

Table 2 Three data sensitivity grades in public security

数据信息类别	安全保护等级	分类标准
敏感	第一级	公安部门按照政策法律获取的可能涉及国家安全和政治稳定的关键信息
限制	第二级	公安部门通过侦察获取的记录信息，以及从外部获取的国家机关、社会单位所掌握的，涉及公民或组织身份、行为依法不能公开的重要信息
公开	第三级	公安部门采集的基础性资料信息，以及从外部获取的国家机关、社会单位所掌握的依法可以面向社会公开的一般信息

3.5 数据资源分类访问控制

数据资源分类是指根据数据来源、业务手段、敏感等级、访问对象等各种因素，将数据资源表划分成不同数据资源集合的过程。

4 公安资源分级分类访问控制的 TFR 模型及方法

4.1 公安数据资源的分类原则和体系

4.1.1 数据资源分类原则

数据资源分类是根据业务需要动态进行，为权限分配提供基础和依据，其作用类似授权模型中的“角色”。

4.1.2 数据资源分类体系

对数据资源的分类可以参考以下方法，但以实际应用需求为准：(1)依据信息来源的不同，将数据资源分为公安核心信息、公安管控信息、社会管理信息和社会公开信息；(2)依据信息所属手段业务不同，可以分为若干类别；(3)根据信息对象属性不同，可以分为身份信息、行踪信息、生活信息、通信信息、交易信息和其他信息(见表 3)。

表 3 基于信息属性的分类

Table 3 Classification based on information attributes

信息类别	分类标准
身份信息	与真实身份相关的身份信息，如身身份证号、社保卡号、医保卡号、银行账号、电子邮件地址、各类网络身份信息等
轨迹信息	与地理位置相关的信息，如 GPS 信息、出入境信息、出行信息、住址信息、上网服务场所编码、营业场所编码、地址、基站、经纬度等
生活信息	与社会生活相关的信息，如医疗信息、资产信息、法律信息、交通信息、交际信息等
通信信息	指与通信相关的信息，如手机号码、固话号码等
交易信息	指与交易相关的信息，如金融交易信息、财务信息、网络购物信息等

4.2 公安数据资源分级原则和体系

数据资源分级是指根据数据表、数据记录、数据字段的敏感程度不同，对数据表、数据字段、数据记录赋予不同敏感级别的过程。数据分级是实施数据隔离和数据保护的第一步。

4.2.1 数据资源分级原则

数据资源的分级应依据信息的敏感程度进行，分为数据表 T、数据字段 F 和数据记录 R 3 个维度。这 3 个维度各自独立控制，又相互交叉，共同控制是否允许用户访问相应的资源。针对公安数据资源的查询、布控、分析等操作，须先对用户数据资源访问权限进行鉴权，即根据用户 T-F-R 三元组的信息决定用户是否具备访问权限。

对用户访问等级的授权和鉴权都通过统一授权审批管理模块进行。在数据分类已授权的情况下，数据分级的判断优先顺序为：数据表＞数据字段＞数据记录。数据访问服务应根据用户授权三元组对数据分类、数据表、数据项记录、数据记录分级等条件按照先后顺序分别检查。

4.2.2 数据资源分级方法

数据资源 3 个维度(数据表 T、数据字段 F 和数据记录 R)的分级均以量化数值表示，对应的敏感等级量化值从低至高表示为 1－9(具体数值可以配置)。对于敏感度赋值为 0 的数据表和数据字段视为完全公开信息，不需要采取任何访问控制措施；对于敏感度赋值为 0 的数据记录，同样不作任何过滤操作。

(1)数据表分级配置操作。初始配置操作方法:根据表 2 定义的 3 个等级对应的敏感度量化等级范围,分别为(7－9)、(4－6)、(1－3)。对于完全不具有敏感性的数据表,将其数据表敏感等级赋值为 0。当用户数据资源权限等级三元组中的数据表权限 T 值大于对应目标数据表的敏感等级时,用户的访问行为才会被允许,反之其访问行为将被拒绝。数据表的分级由管理员事先配置完成,在授权管理中心进行管理。

(2)数据字段分级配置操作。数据字段对应的敏感等级量化值从低至高表示为 1－9。数据字段的敏感度分级可以依据自身需求进行定义。

字段分级使用原则如下:当用户数据资源权限等级三元组的数据表权限值 F 大于或等于对应目标数据字段的敏感等级时,用户的访问行为才会被允许,反之其访问行为将被拒绝或部分受限。数据字段的分级由管理员事先配置完成。

(3)数据记录分级配置操作。数据记录对应的敏感等级量化值从低至高表示为 1－9。数据记录的分级通常是由对象身份的敏感性引发。数据记录分级控制的使用是以全局敏感对象分级名单的创建为前提的。敏感对象分级量化可综合行政级别等因素进行统一考虑,这里不再赘述。数据记录的分级标识是在数据清洗入库过程中完成的。数据治理单元分析入库数据资源的每一条记录,如果系统发现记录中的某些字段出现了敏感对象名单中包括的特定实体身份标识或虚拟身份标识,例如性别、身份证号码、出生日期、微信号码、手机(固话)号码、住址等,将根据敏感对象名单中对应目标的级别,自动对记录进行分级标识。

数据记录分级访问授权过程如下:当用户对上述数据记录所在数据表提出访问请求时,先完成数据表、字段的鉴权,数据访问接口自动从授权管理模块读取用户数据记录权值 R,并将其作为条件合并到访问操作中。当用户数据记录权值 R 大于数据记录敏感等级时,该记录才会被返回,反之将会被过滤。

4.3 基于 TFR 模型的鉴权判别优先级顺序

数据资源分级和分类是数据资源访问控制的决策依据。一般情况下,对特定数据资源的访问首先要判断数据资源所在分类是否被授权。在数据分类已授权的情况下,数据分级的判断优先顺序为:数据表>数据字段>数据记录,即将用户授权三元组(数据表 T、数据字段 F 和数据记录 R)先后与目标资源的数据表权值、数据字段权值、数据记录权值一一比较检查,最终确定用户是否具有访问权限。

算法 1 表及字段鉴权算法

Input:(U,〈T,R,F〉,{TABLE$_i$.FIELD$_j$},{T$_i$ for TABLE$_i$},{F$_j$ for TABLEi.FIELD$_j$})
Output:(U,〈R〉,{TABLE$_m$.FIELD$_n$})
for each TABLE$_i$
if (T>=T$_i$ for TABLE$_i$)
  for each TABLE$_i$.FIELD$_j$
    if(F>=F$_j$ for TABLE$_i$.FIELD$_j$)
      add (TABLE$_i$.FIELD$_j$,{TABLE$_m$.FIELD$_n$})

其中,U 为用户,〈T,R,F〉为用户三元组,{$TABLE_i.FIELD_j$}为用户请求访问数据字段(请求访问整个数据表视为请求访问该表中的所有字段),T_i for $TABLE_i$ 为数据表 $TABLE_i$ 的敏感等级,F_j for $TABLE_i.FIELD_j$ 为数据字段 $TABLE_i.FIELD_j$ 的敏感等级。

算法 2 数据记录鉴权算法

Input:(U,R,{R$_i$ for RECORD$_i$})
Output:(U,{B$_i$ for RECORD$_i$})
for each R$_i$ for RECORD$_i$
  if (R >= R$_i$ for RECORD$_i$)
    add (TRUE,{B$_i$ for RECORD$_i$})
  else
    add (FALSE,{B$_i$ for RECORD$_i$})

其中,U 为用户,R 为用户三元组〈T,R,F〉中的 R 值,R_i for $RECORD_i$ 为数据记录 $RECORD_i$ 对应的敏感等级,{B_i for $RECORD_i$}为用户 U 针对每条记录可访问或不可访问的布尔值集合。

5 TFR 访问控制模型应用操作场景

5.1 应用初始化操作步骤

数据资源分级分类功能的配置和使用分为以下 3 个步骤(如图 1 所示,其中{O}为数据对象集,{U}为用户集,{〈T,R,F〉}为用户的〈T,R,F〉三元组集)。

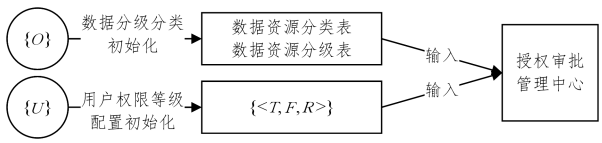


图 1 TFR 访问控制体系的初始化示意图

Fig. 1 Schematic diagram of TFR access control system initialization

步骤 1 数据分级分类配置初始化

统一授权审批管理中心的系统管理员对所有公安数据资源定义的数据表,按照上级建议或自身需要,设定数据表访问等级和每个表对应字段的访问等级,默认情况下所有等级为 0。各地根据自身需要,建立特定的敏感对象名单,并配置相应的级别。系统管理员可以根据业务需要对数据资源表进行分类。

步骤 2 用户权限等级配置初始化

统一授权审批管理中心的系统管理员对每一用户的授权三元组〈数据表 T、数据字段 F 和数据记录 R〉进行初始化配置,分别赋予 1－9 之间的某一数值。若未进行赋值,系统可以设定一个默认权限。

步骤 3 分级分类访问鉴权调用过程

数据资源分级分类主要针对数据资源的查询访问授权。当前的公安大数据建设中,涉及平台数据查询访问的主要为通用查询等服务。以通用查询服务过程为例,具体过程为:1)上层业务系统如公安信息综合应用系统按照原有接口,正常调用通用查询服务,请求中包括用户身份信息(数字证书 ID 等)以及需要访问的目标资源表和相关字段信息;2)通用查询服务接到请求后首先向授权审批管理中心发起鉴权请求,鉴权请求携带用户信息及数据资源信息;3)授权审批管理中心收到请求后,根据用户的授权三元组〈T,R,F〉中的〈T,F〉二元组进行鉴权,鉴权后向通用查询服务返回鉴权结果;4)通用查询服务根据返回的鉴权结果向数据中心发起查询,数据中心执行查询命令并根据用户〈T,R,F〉中的 R 值向通用查询服务返回用户有权访问的数据记录;5)公安信息综合应用系统收到通用查询服务返回的查询结果,用户可根据实际情况在通用查询服务里配置是否返回鉴权结果。鉴权调用

过程中模块关系图如图 2 所示，其中，两部分授权审批管理中心可考虑合并部署。

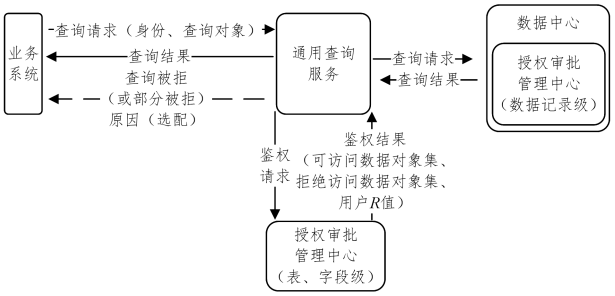


图 2　云平台访问授权体系的模块关系图

Fig. 2　Relations among cloud platform access authorization system modules

一般情况下，在系统实现时数据分级分类控制过程对上层业务应用是透明的，不应增加、改变当前业务的接口和调用方式。但数据访问服务层应在内部实现中增加对授权审批管理服务接口的调用。

5.2　典型应用的场景举例

一般在分类授权的情况下，系统将根据用户的〈T，F，R〉三元组匹配与目标资源权限进行分级鉴权。例如，用户的权限三元组为〈6，5，4〉，说明该用户能访问敏感度为 1－6 等级的数据表、1－5 等级的数据字段和 1－4 等级的数据记录。

下面结合一些典型的应用场景对各类情形进行说明。

场景 1　用户查询请求中的数据表未授权。用户发起查询请求，查询条件中包含的数据表未授权（因数据所在分类未授权或用户数据表权值 T 低于数据表敏感级别），则查询请求将被直接拒绝，同时返回拒绝原因。

场景 2　用户查询请求中的数据表已授权，但数据字段全部未授权。用户发起查询请求，查询条件中包含的数据表已授权，但所查询的数据字段全部未授权（用户数据字段权值 F 低于所有数据字段敏感级别），则查询请求将被拒绝，同时返回拒绝原因。

场景 3　用户查询请求中的数据表已授权，但数据字段部分未授权。用户发起查询请求，查询条件中包含的数据表已授权，但所查询的数据字段部分未授权（用户数据字段权值 F 低于部分数据字段敏感级别），则只返回部分授权字段，同时返回原因说明。

场景 4　用户查询请求中的数据表已授权，查询请求未指定具体字段。用户发起查询请求，查询条件中包含的数据表已授权，但未指定查询的数据字段，则系统只返回敏感级别小于或等于用户权限等级的全部数据字段。

场景 5　用户查询请求中的数据表、数据字段已授权，但部分数据记录未授权。用户发起查询请求，查询条件中包含的数据表、数据字段已授权，但该数据表中的部分数据记录未授权（用户数据字段 R 值低于部分数据记录的敏感级别），则系统只返回敏感级别小于或等于用户权限等级的数据记录。

场景 6　用户提请越级查看高敏感等级的数据。在某些应用场景下，应用系统可以使用户在分级分类访问机制查询返回后，看到系统存在其他高敏感性未授权的字段，但无法直接访问字段对应的记录内容。应用系统可以提供审批机制，允许用户对某一行记录直接提出越级访问审批请求，根据有关工作办法规定的负责人或授权内设机构负责人等进行分级审批并通过后，系统赋予其更高权限访问有关内容信息。

结束语　本文针对公安大数据建设中云平台数据资源访问控制，对分级分类访问控制策略进行了研究，提出了针对敏感业务资源分级分类实现访问控制的原则、思路和具体操作方法，对有序、安全推进公安敏感信息资源的汇聚和共享提供了系统化的思路和解决方案，在有关警种的大数据建设中得到了实战应用，取得了良好的效果。本文提出的方法有助于公民隐私信息的保护，推动全国公安大数据建设安全、有序、可控地开展。下一步，将推动本文提出的访问控制模型更为广泛地应用于各地智慧公安的建设，并针对公安大数据环境下的细粒度分级分类访问控制的效率、负面授权（negative permission）、权限冲突等问题展开研究。

参 考 文 献

[1] Sohu News. The Ministry of Public Security established the Leading Group for the National Big Data Work in Public Security[OL]. http://news.cyol.com/yuanchuang/2018-01/25/content_16901261.htm.

[2] Sohu News. The National Meeting of Directors of Public Security was convened: focus on the development of big data in public security[OL]. http://www.sohu.com/a/291349459_653639.

[3] Sohu News. Hot news interpretation | The Ministry of Public Security: Six Key Tasks in the Construction of 'Smart Public Security'[OL]. http://www.sohu.com/a/291349459_653639.

[4] HONG F. Introduction to Access Control [D]. Wuhan: Huazhong University of Science and Technology Press, 2010.

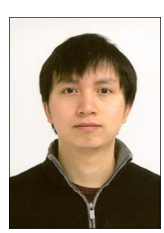
[5] SANDHU R S, COYNE E J, FEINSTEIN H L, et al. Role-Based Access Control Models[J]. IEEE Computer, 1996, 29(2): 38-47.

[6] SANDHU R S, FERRAIOLO D F, KUHN D R. The NIST model for role-based access control: towards a unified standard [C]// ACM Workshop on Role-Based Access Control. ACM SIGSAC, 2000: 47-63.

[7] OSBORN S L. Role-based access control: past, present and future[C]// International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services. ACM, 2006: 4.

[8] SERVOS D, OSBORN S L. Current Research and Open Problems in Attribute-Based Access Control[J]. ACM Computing Survey, 2017, 49(4): 65: 1-65: 45.

GU Rong-jie, born in 1977, Ph.D, professor. His main research interests include network security, massive information processing and police informatization.



WU Zhi-ping, born in 1985, postgraduate, assistant professor. His main research interests include police informatization, access control, and database security.