基于属性加密的计算机数据库安全检测工具的设计与运用

陈良英
（四川信息职业技术学院，四川 广元 628017）

摘　要：提出网络域数据密匙共享和准入机制，创建新型数据库密匙架构。设计了基于数据属性加密的计算机数据库安全检测工具，就数据库数据进行属性加密，重组加密文件关键序列架构，在传统 CDM 数据线性规划算法的基础上，引入加密文件序列信息，设计 TRIE 树结构，将数据序列项转换为字母 ID，利用存储容器存储数据项，建立 Apriority 数据序列索引，并生成索引检索数据表，通过数据库扫描，对异常项进行隔离，实现数据库安全检测。仿真实验数据显示，应用上述检测工具，数据库文件误用异常检测率提高了 17%，伪装攻击拦截率提高 20%，可以肯定有效地提高计算机数据库信息的安全性。

关键词：新型密匙架构；属性加密；数据库；误用检测；安全检测
中图分类号：TP309.7　　**文献标识码**：A　　DOI：10.19358/j.issn.2096-5133.2020.04.006
引用格式：陈良英. 基于属性加密的计算机数据库安全检测工具的设计与运用[J]. 信息技术与网络安全，2020，39(4)：30-35.

Design and application of computer database security detection tool based on attribute encryption

Chen Liangying
(Sichuan Vocationl College of Information Technology, Guangyuan 628017, China)

Abstract: In this paper, we propose a data key sharing and access mechanism in network domain, and create a new database key architecture. A computer database security detection tool based on data attribute encryption is designed. The database data is encrypted by attributes and the key sequence structure of encrypted files is reorganized. On the basis of the traditional CDM data linear programming algorithm, the encrypted file sequence information is introduced, the TRIE tree structure is designed, the data sequence items are converted to letter ID. The storage container is used to store data items, the Apriority data sequence index is established, and index retrieval data tables are generated. Through database scanning, anomaly items are isolated to achieve database security detection. The simulation results show that the application of the above-mentioned detection tools can increase the detection rate of database file misuse by 17% and the interception rate of camouflage attack by 20%. It can be confirmed that this method can effectively improve the security of computer database information.

Key words: new key architecture; attribute encryption; database; misuse detection; security detection

0　引言

我国信息技术研究起始于 20 世纪 60 年代。经过数十年的发展，已经融入到我国国民生活的各个角落。现阶段我国互联网相关领域和技术正在不断与社会接轨，我国正处于高速发展的信息时代。信息社会使个人和企业之间以数据信息的形式进行交流，这就让现代互联网数据信息量不断增长。在这样一个数据量激增的时代，企业和个人均会产生巨大的数据资源，这些数据资源的存储离不开数据库。社会中的个人和企业，均与数据库存在紧密的关系，不仅个人的财政信息如银行卡余额、身份证号码等存在于数据库中，大量企业信息甚至是涉及国家高度军事机密或重要数据也存在于数据库中。然而因为现代互联网信息具有巨大的潜在价值，其存储位置相对集聚性较强，不法分子针对数据库的网络攻击往往可以获取最直接有效的攻击效果，数据库已经成为非法人员进行资源窃取的首要攻击目标。2018 年，Virerzon 公司发布了《2018 年度数据库泄露

报告》,该报告是由 Virerzon 公司以及 39 家合作企业共同完成的,其中包括了著名数据网站 Delouti、EMC、Chetk、Intouerce、Froint 公司等。报告内容显示,仅 2018 年上半年,全球就发生了 945 起较大型的数据泄露事件,共计导致 45 亿条数据泄露,与 2017 年相比数量增加了 133%。为了有效降低安全隐患,现代计算机数据库的管理人员必须对数据库进行安全检测,消除安全隐患。传统情况下,针对计算机数据库安全检测工具一般基于数据代码和数据密匙,这种检测手段在初期虽然取得了一定效果,但是随着入侵技术的提高,对数据库安全隐患的隔离率,已经不能满足数据库维护的需要。针对这一情况,设计提出了以属性加密为核心的计算机数据库安全检测工具,维护计算机系统数据库的安全[1]。

1 计算机数据库安全检测工具设计

设计的安全检测工具主要分为三个部分:一是文件属性加密部分,设计通过生成公有和私有密匙进行文件属性加密;二是加密文件数据序列索引,对加密文件进行序列重组;最后利用数据文件建立 TRIE 树结构,根据索引序列进行数据迁移,即完成数据检测。

1.1 文件属性加密

基于属性的加密机制是现代云计算网络以及分布式等环境下,对数据中心数据库进行访问控制的主体研究热点之一。因为现代计算机数据库一般均采用分布式处理结构,而且在多分部机构 ABE 中,一般不会采用中央授权的管理方式,所以在对数据库全局身份标识进行共谋攻击时,身份全局管理会出现一定的安全隐患。而设计的计算机数据库安全检测攻击引入了一种新型的文件属性加密方案,用户在进行多项密匙申请时,可以根据用户自身的指令协要,实现用户标识不公开,提高属性加密效果。

设计提出的文件属性加密主要基于新型的密匙架构,加密方案将密匙主要分为两个部分即域内密匙和域外密匙。域内密匙的生成,主要是通过具有域内密匙属性的授权加密机构(AA)进行密匙授权。AA 会为当前加密用户提供一个特质的属性密匙。而域外密匙则需要域内属性密匙原本数据代码完成,详细的域内和域外密匙分发方案如表 1 所示。

表 1 密匙组成数据表

密匙	数据表述	密匙功能
CK_a	授权属性公共密匙	建立信任关系
SK_a	授权属性私有密匙	创建加密文件属性公共密匙
Pki	属性共有密匙	加密文件
SkIu	用户属性密匙	解密文件

每一个数据库授权结构均有一个相应代码下的加密公共密匙和一个私有密匙。具有通信链路的不同授权书信机构之间,可以在密匙建立的初期,通过密匙信息共享建立公共密匙的信任链路。数据库每一个网络域的属性授权机构,如果要建立加密,需要具有信任域中所有区域的属性授权机构公共密匙信息。详细密匙授权架构如图 1 所示[2]。

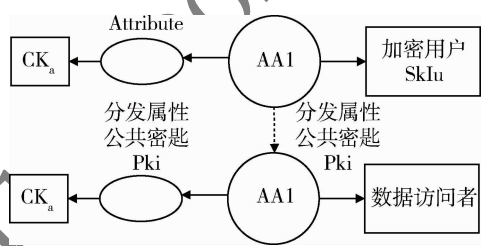

图 1 网络域内属性密匙分布架构图

根据图 1 可以看出,AA1 具有自身的公共密匙和私有密匙,同时获取了其他域内的私有密匙,可以直接用于生成具有管理属性的公共密匙,以及文件加密的私有密匙。对于每一个文件网络域属性,都有一个公共密匙,用于加密文件。一个文件数据的管理者可以根据访问策略,利用不同域内属性的公共密匙进行数据加密。具体方案为:

令数据 $N = p_1p_2p_3$,其中 $p_1p_2p_3$ 为素数,G 和 G_T 分别为数据 N 的双线数据群,用 $e:G \times G \rightarrow G_T$ 表示加密数据未知结构下数据线性映射。在 AA 中生成用户数据文件标识,映射到 G 文件中的群元素中,另外制定数据外定义下的有限数据文件哈希函数集合,每一个 AA 生成的私有密匙,均需要存在于且仅存在于哈希函数集合中,对文件属性进行处理[3]。需要着重说明的是,该方案主要基于文件属性多数字阶层的群构造,但是所有的 G 文件均需要存在于子群中。文件属性加密操作如下。

(1)Globat Setup(GP),输入安全性参数 λ,根据参数生成全局性公共参数 GP,选择一阶 GP 为数据 N 的实际下述线性数据集群。在 GP 中,主要包括数据 N,G_{p1} 的生成源 g_1,以及随机属性加密函

数的 $H1$ 和 $T1$ 表述。

(2) Authery Setup(AP):根据属性授权机构密匙执行下的初始化进程,每一个授权单位均需要选择一个随机指数;从上述建立的哈希函数集中生成一个具有一致性的随机函数,将这个函数作为属性授权机构的索引,也就是授权机构的加密密匙,则属性授权机构的私有密匙表示为 SK_a,公有密匙表示为 CK_a,而 IDA 则表示为当前属性授权机构的身份编码。其表达公式为:

$$SK_a = [S_a, x_a, \forall_a] \tag{1}$$

$$CK_a = [g_1^{s_1 H_1}(IDA), \forall_a] \tag{2}$$

(3) Auto-attacker:因为属性公共密匙的授权机构需要根据网络域属性进行公参,属性标识以及授权属性私有密匙生成,其中:

$$ID_j = IDA\pi ID \tag{3}$$

则 Pki 的生成公式可以表示为:

$$Pki = \{Pk_j e(g_1, g_2)^{s_x(H_{xx}(ID_j))}, PK_g iH_{xa}(id_j)\} \tag{4}$$

公式(4)中的属性身份标识 ID_j 主要是由 AA 的身份表示属性和对应编码生成,而 IDA 则主要表示为授权机构内的身份编码。文件属性的多样化全局加密标识,使得属性标识是整个局域网信任域中唯一的验证性标识,所以每个公共属性标识在信任域中,也具有唯一性[4]。根据全局加密表示,设计为加密用户提供了全局身份标识,此标识由当前数据库和使用用户所在域的 AA 标识组成。在 IDA 下,使用用户需要申请属性标识,而且设计为了实现使用用户标识的全网唯一性,会为使用用户提供一个不同属性编码。在加密过程中,加密工具会首先检查使用者身份是否具有属性授权机构的负责项,然后验证使用者的实际身份,检测是否具备属性授权的可能性,如果不可以则输出结果为空,反之则生成数据密匙。

(4) Decrypt:针对属性密匙的机密算法为 CT 输出密文,全局参数。同一个身份标识,用户的机密属性密匙集具有外部相似性。如果使用者的属性密匙满足加密时的访问控制要求,则可以完成属性解密。

1.2 建立 Apriority 数据序列索引架构

利用上述方式对数据库文件进行加密后,原本的序列架构会被打乱,在进行安全检测过程中,无法进行有效识别。设计针对加密文件的密匙序列[5],对其进行文件关键序列架构重组。因为加密属性文件具有明显的封闭开源性,所以设计提出了 Apriority 封闭性数据索引架构。整个架构采用集中式数据管理,包括序列程序检索、序列存储以及序列引源的建立。这种索引数据架构的相对配置资源较为简单,配置体系和操作整体较为便捷,其架构如图 2 所示。

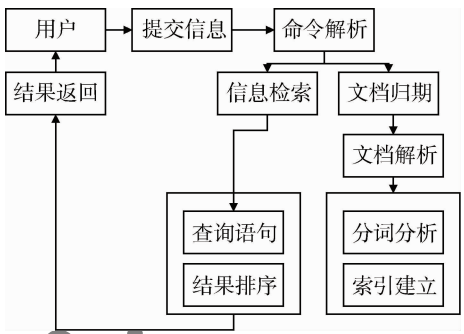

图 2 序列索引架构示意图

设计的索引序列架构从索引功能上来说包括四个索引单元分别为:Apriority 底层数据管理单元、Apriority 数据序列交换单元、Apriority 并行序列解析单元、Apriority 文档数据解析单元[6]。

Apriority 底层数据管理单元主要负责帮助安全检测管理工具的使用用户,建立正常的加密数据文件序列索引库,将使用者提交的加密数据文档和文档序列号,利用镜像索引结构进行数据准压缩、数据合并以及数据优化。另外 Apriority 底层数据管理单元还可以针对使用用户提交的各类加密文件索引语句进行密匙解读。

Apriority 数据序列交换单元主要负责对加密文件序列索引用户的各类处理结果进行回馈,包括数据输入文档的检索信息以及数据高亮分页处理等。该序列交换命令的解析包括整个索引库的数据分配过程,使用用户在进行操作时,用户界面和后台服务界面会进行数据资源管理器之间的数据串联,同步传递状态命令质量检测等。此外数据序列交换单元还包括对当前加密信息的综合判断,并与后台相关联[7]。

而 Apriority 并行序列解析单元和 Apriority 文档数据解析单元,则主要负责将各类加密文件的序列文档进行集聚汇总,通过提取文档解析分析内

容,获取明确的文件标识码,进而生成解析序列。在进行序列解析过程中,需要对加密数据进行准确的数据描述和数据整理,利用适当的数据索引源进行有序数据文件标记。这些标记后的特征数据可以作为检测工具中所有特征数据索引字段的数据集合,根据数据集合中的索引项,可以将索引序列进行割裂操作,然后生成数据序列文档。序列内容包括序列用户名称、序列存储量、序列数据量、序列生成日期、序列格式文档 ID、序列存储位置等,如图 3 所示。

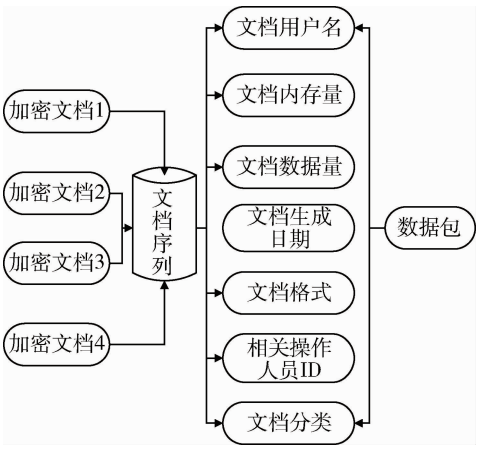

图 3　索引数据集合

在建立 Apriority 数据序列索引时,需要针对不同的文档索引格式提出相应的属性要求。

是否可以进行有效存储:数据索引是否完全包含加密数据序列,即数据系列数据集是否存在于数据索引存储中。这种存储方式需要完全符合小型文本文档内容的域名格式,否则很容易引起索引序列的索引量过载。

是否可以进行有效索引:如果当前加密数据的索引序列无法被正确索引,那么数据库安全检测工具的使用人员就需要在该数据源上进行分配数据检测,例如加密文档存储位置检索,文档数据分类检索,生成额外的检索条件。

是否出现明确的检索分词:检索分词主要表现为按照某种特定的分词计划进行数据切取后,可以从数据中匹配到的属性关键词。这种关键词可以有效提高搜索结果的精确度。部分加密数据检索属性无法作为关键词使用,如数据文档生成和归档日期等。

根据当前加密数据的索引序列和具体检索需要,对当前计算机数据库内属性加密信息制备检索数据,如表 2 所示。

表 2　Apriority 数据序列索引检索数据表

名称	是否存储	是否索引	有无分词
文档用户名	1.00	1.00	1.00
文档内存量	1.00	0.00	1.00
文档数据量	1.00	0.00	0.00
文档生成日期	1.00	1.00	0.00
文档格式	1.00	0.00	1.00
相关操作人员 ID	0.00	1.00	1.00
文档分类	0.00	1.00	1.00
文档存储位置	1.00	0.00	0.00

1.3　检测工具完成

目前数据库安全检测算法一般采用 CDM 数据线性规划算法[8],设计基于上述数据序列索引架构和加密文件域名架构,对其进行改进,提出新的 TRIE 树算法,完成数据库的检测。TRIE 树结构最早被用于高效检索字典结构。其根深度被定义为 0。结构的第 N 层节点会指向第 $N+1$ 层的数据节点,其指向针被称之为检索数据边。每一个数据边都用对应的字母或者 ID 表示。例如节点 A 指向节点 B,则认为 A 节点是 B 节点的母系节点,B 节点是 A 节点的子系节点。因为 TRIE 树结构不仅可以有效存放检索单次项,还可以存放各种序列集。具体方法就是将 TRIE 树的节点或者边作为有序序列集的子项,此时 TRIE 树各个路径就可以与序列重合。根据上述操作,设计将生成的 Apriority 数据序列作为序列项,生成序列网格。在进行数据项集发掘时,利用字母代替加密数据序列项集,这样就可以将加密数据序列表示成由字母构成的单次项。图 4 为存放了选项集的一个 TRIE 树结构。

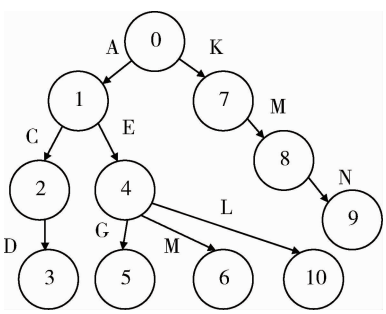

图 4　五个选项集下的 TRIE 树结构

根据 TRIE 树结构,对当前加密数据文件序列事物进行存储,存储容器为 CECTOR(……),此时选择安全精简事物记录,该记录需要默认程序选择。精简记录存储容器为 map < CECTOR < …… > ,unsigneder long > 。

进行多次数据库扫描后,数据库本体数据项会存在于存储容器中。这里的精简事物即只存在于数据库中的频繁数据项。此时重复数据将被计数,以此建立数据公式:

$$\begin{cases} \text{shellelement}: \begin{cases} \sigma_{eq} \leqslant [\sigma_{eq}]_1 \\ T_{max} \leqslant [T_{max}]_1 \end{cases} \\ \text{Beamelement}: \sigma_b \leqslant [\sigma_b]_1 \end{cases} \tag{5}$$

公式中 σ_{eq}、$[\sigma_{eq}]_1$ 分别为当前数据库数据最高项序列和最低项序列;T_{max} 和 $[T_{max}]_1$ 分别为数据库扫描后的最大存储和最小存储;σ_b 和 $[\sigma_b]_1$ 数据边的两个极值。根据公式(5)数据库本体文件序列会完全迁移到数据容器中,当出现安全问题项或者未加密数据项时,该项无法生成加密序列,同时无法被迁移,通过文件基数,即可为探知,检测宣告完成。

2 仿真实验

为验证本文设计的基于属性加密的计算机数据库安全检测工具的实际检测效果,进行有效性试验检测。通过对预设目标进行多范围多角度数据攻击,获取检测数据,进而证明该工具的实际有效性。在试验过程中,模拟样本数据建立数据库。使用 CTM 数据检测工具作为对比组,进行检测效果对比。检测包括误用异常检测和伪装攻击检测两部分,具体实验数据如下。

2.1 误用异常检测

误用异常检测时,主要检测对象为每条用户日志以及聚簇规则是否完全匹配,以下为实验检测定义。

定义 1 如果当前数据库文件的一条日志记录与当前聚簇规则的两侧规则均具有统一性,则认定其具有匹配规则。

定义 2 一个加密数据文件如果其加密序列不符合当前聚簇规则,则其可信度加权值作为异常检测值出现。

根据定义 1 和定义 2,以实验数据库为目标,进行误用异常数据输出,并利用实验对比的两种工具进行检测,其检测结果如图 5 所示。

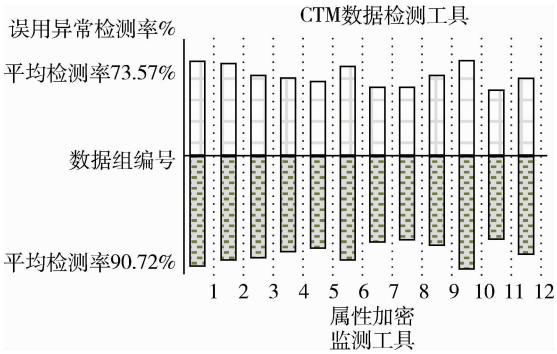

图 5 误用异常检测率对比

根据图 5 数据可以看出,应用设计的属性加密计算机数据库安全检测工具,对误用异常检测具有较好的检测效果。根据数据统计可以确定,其数据库文件误用异常检测率提高了 17%,可以证明其有效性。

2.2 伪装攻击检测

伪装攻击是对当前数据库进行主动攻击活动的总称。伪装攻击检测,主要验证检测工具的鲁棒性。实验共进行了 10 组有效的伪装攻击,其拦截率如表 3 所示。

表 3 伪装攻击拦截率

序号	属性加密/%	CTM/%
1	96	72
2	95	80
3	97	76
4	91.3	77
5	89	70
6	93	75
7	91	72
8	93	81
9	97	80
10	92	76

根据以往的数据拦截经验,当拦截率高于 90% 时,可以认定为绝对拦截,此时处于绝对安全状态;当拦截率为 75% ~90% 时,为次拦截属于基本安全状态;当拦截率低于 70% 时,数据库为危险状态。根据实验数据可以看出,应用属性机密的安全检测工具,拦截率均高于 90%,数据库处在绝对安全状

态,比传统 CTM 检测工具拦截率提高了 20% 以上,进一步证明了其有效性。

3 结论

数据库信息安全,是数据库建设和维护的核心工作之一。对国民生产和经济保障具有重要意义。本文基于属性加密,提出了新型数据库数据安全检测工具,可以有效提高数据库安全保障,具有实际应用价值。

参考文献

[1] 黄保华,贾丰玮,王添晶. 云存储平台下基于属性的数据库访问控制策略[J]. 计算机科学,2016,43(3):167-173.

[2] 罗云锋,熊伟,许庆光. 一种基于属性加密的数据保护与访问控制模型[J]. 网络安全技术与应用,2017,2(9):53-56.

[3] 杜朝晖,朱文耀. 云存储中利用属性基加密技术的安全数据检索方案[J]. 计算机应用研究,2016,33(3):860-865.

[4] 王乐,王芳. 数据库异常数据的检测仿真研究[J]. 计算机仿真,2016,33(1):430-433.

[5] 林素青. 支持访问更新的可验证外包属性加密方案[J]. 网络与信息安全学报,2019,5(1):37-49.

[6] 主艳姣,李艳平,鲁来凤,等. 云存储环境下支持用户动态撤销的属性加密方案[J]. 陕西师范大学学报(自然科学版),2018,46(5):1-8.

[7] 刘熙,胡志勇. 基于 Docker 容器的 Web 集群设计与实现[J]. 电子设计工程,2016,24(8):117-119.

[8] 邓宇乔,杨波,唐春明,等. 基于密文策略的流程加密研究[J]. 计算机学报,2019,42(5):1063-1075.

(收稿日期:2019-11-07)

作者简介:

陈良英(1981 –),女,硕士,副教授,主要研究方向:计算机应用。

版权声明