基于密钥分存的关系数据库数字水印技术*

胡　斌[1], 施荣华[1], 彭沛夫[2]
(1. 中南大学 信息科学与工程学院, 长沙 410083; 2. 湖南师范大学 物理与信息学院, 长沙 410081)

摘　要: 根据关系数据库的特殊性,结合现有数字水印技术,将可信中心和密钥分存的思想引入关系数据库水印算法。首先争取对水印信息分存,然后根据分存值序列,分成三元组序列,对数据库中有效位数较多的数值型字段值,修改其最低有效位的值。该算法隐蔽性好,可基于部分内容恢复水印信息;可嵌入信息量大,每个字段值可以插入 3 bit;水印恢复不需要原始数据库;可以抵抗各种子集攻击和添加攻击。

关键词: 数字水印; 数据库数字水印; 密钥分存; 可信中心
中图分类号: TP391　　**文献标志码:** A　　**文章编号:** 1001-3695(2007)08-0144-02

Digital watermark technology for databases based on secret sharing with trusted party

HU Bin[1], SHI Rong-hua[1], PENG Pei-fu[2]
(1. *College of Information Science & Engineering, Central South University, Changsha* 410083, *China*; 2. *College of Physics & Information, Hunan Normal University, Changsha* 410081, *China*)

Abstract: A digital watermarking for relational databases algorithm which was based on secret sharing with a trusted party was proposed. The author cut the shared information of digital watermarking into3-bit-set, and update the value of date's LSB, which was of more effective bits in database. This algorithm is good at hiding, embedding much more information and only need part shares without original database in the recover progress and this algorithm is safe and effective to prevent set-attach and add-attach.

Key words: digital watermarking; digital watermarking for database; secret sharing; trusted party

0　引言

数字水印技术[1,2]是网络环境下保护版权的新型技术,可以确立版权所有者,识别购买者或提供关于数字内容的其他附加信息,并将这些信息以人眼不可见的形式嵌入在数字图像、数字音频和视频序列中,用于确认所有权和跟踪行为。数字水印作为一种新兴的信息安全技术已经被许多应用领域所采用。

关系数据库数据不同于多媒体载体数据,有其自身特点,直接将多媒体水印技术应用到关系数据库中是不适合的,因此就必须研究一种适应关系数据库自身特点的水印。自 2002 年, R. Agrawal 等人[3]首次提出关系数据库水印,并提出一种适应关系数据库特性的水印标记策略;2003 年, R. Sion 等人[4]在关系数据库中对数值型属性进行标记的方法实现水印来保护关系数据库的版权;2003 年,牛夏牧等人[5]对关系数据库数字水印进一步研究加入小量有实际意义的水印技术,国内其他水印工作者为推进数据库水印研究发展作了许多努力[6,7]。它们均有一个共同的缺点:嵌入的是没有实际意义的水印,而嵌入水印量较少。在关系数据库中加水印应该充分考虑到关系数据库自身的特殊性以及各种攻击方式,基本要求就是水印算法要在鲁棒性、不可见性之间折中考虑[3,4]。

1) 鲁棒性要好　加入的水印就越能够经受住各种恶意的攻击和善意的数据更新。

2) 不可见性要高　不可见性越高说明加入的水印方法越好,水印越不被用户察觉,越不会因为加了水印而影响关系数据的使用。

3) 嵌入的信息要多　加入的信息量越多水印就越不容易被破坏,误判的概率越小。

4) 误判率要小　误判有两种:一种是有水印时不能够检测提取出来;另一种是无水印时却检测到水印存在。

5) 盲测　由于数据库数据需要经常更新维护,关系数据库的数字水印技术最好能够在水印提取时实现盲测,即水印检测既不需要原始数据库数据(加水印前数据),也不需要水印。该特征关键是可以在关系数据库数据拷贝中检测到水印,而不用考虑原始关系数据库的更新。

1　算法基本思想

1.1　水印嵌入算法(图 1)

a) 将水印信息转换为十进制的 ASCII 码形式(用 k 表示);

b) 按照 Blakley 秘密共享门限体制[8],选取合适的参数值

收稿日期: 2006-06-22; 修返日期: 2006-08-19　　基金项目: 国家自然科学基金资助项目(60173041);湖南省自然科学基金资助项目(021TY2094)

作者简介: 胡斌(1983-),男,浙江金华人,硕士研究生,研究方向为信号与信息处理、数字水印(cubahu@163.com);施荣华(1964-),男,湖南常德人,教授,研究方向为计算机网络与计算机通信保密;彭沛夫(1955-),男,湖南岳阳人,副教授,研究方向为参数优化控制及电子信息.

r, n, 将秘密分存, 得到分存值 $k_i(i=1, 2, \cdots, n)$, 将 k_i 转换为二进制序列;

c) 将得到的二进制序列按每 3 bit 分为一组[7], 如果最后一位不足 3 bit, 则可添上一个或两个 0 补足;

d) 选取数据库中满足有效位数要求的数值型字段;

e) 在第一个满足条件的字段中, 选取字段值 $r_i.A_1$, 计算编号 Index = Hash(key, Primary, $r_i.A_1$) ;

f) 根据 Index 值, 将三元组分组序列按照表 1 顺序替换满足条件的 $r_i.A_1$ 值的最低有效位;

表 1　三元组分组序列

根据倒数第二位数的取值, 修改末位数的值										
要嵌入的信息	0	1	2	3	4	5	6	7	8	9
000	00	11	22	33	44	55	66	77	88	99
001	01	12	23	34	45	56	67	78	89	90
010	02	13	24	35	46	57	68	79	80	91
011	03	14	25	36	47	58	69	70	81	92
100	04	15	26	37	48	59	60	71	82	93
101	05	16	27	38	49	50	61	72	83	94
110	06	17	28	39	40	51	62	73	84	95
111	07	18	29	30	41	52	63	74	85	96

h) 继续选取满足条件的字段 A_i, 选取字段值 $r_i.A_i$, 计算编号 Index = Hash(key, Primary, $r_i.A_i$) , 重复 f) , 直到所有满足条件的字段均嵌入了相同的水印;

i) 向可信中心申请注册保护, 提供关系数据库的关键字段、概要说明和水印提取程序; 可信中心根据是否以往有类似的注册数据库和水印提取程序, 判定该数据库是否盗版, 若不是则在中心注册成功。

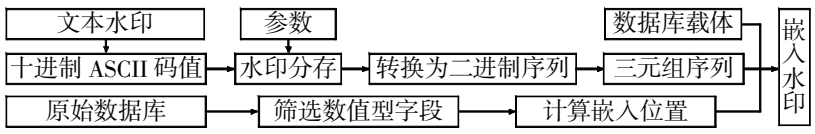

图 1　水印嵌入算法示意图

1.2　水印提取算法(图 2)

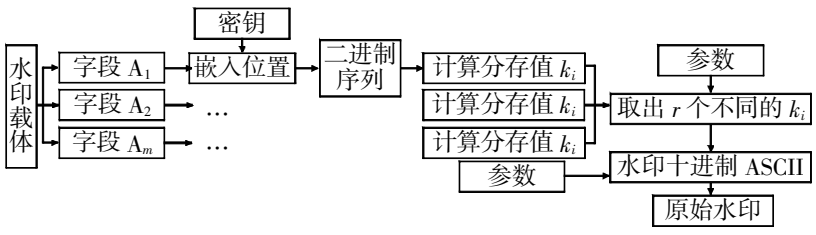

图 2　水印提取算法示意图

水印恢复时, 首先从水印载体的各个嵌入水印的字段中取出分存序列, 计算子水印的分存值; 然后根据参数选择 r 个不同的 k_i, 计算出原始水印信息的十进制 ASCII 码值, 恢复水印。

a) 从数据库中的 A_1 字段中, 计算 Index = Hash(Key, Primary, $r_i.A_1$) 。

b) 根据 Index 值, 顺序将各 $r_i.A_1$ 的末两位数取出, 对照表 1, 得到二进制三元组序列。

c) 根据二进制序列, 计算出 k_i 的值, 根据参数, 判定选取 r 个不同的 k_i 值。如果少于 r 个, 则继续从下一个字段中计算 k_i, 直到取出 r 个不同的 k_i 为止; 若取完所有的嵌入水印的字段, k_i 数小于 r, 则水印提取失败。

d) 根据参数和 r 个 k_i 值, 计算出原始水印的十进制 ASCII 码值, 恢复原始水印。

2　仿真试验

为了测试本算法的性能, 本文基于关系数据库进行了水印的嵌入和提取实验。实验中用到的关系数据库共有记录 6 500 条, 每条记录有数值型字段值 6 个。参数取值: $r=3$, $n=6$, 即 6 个子水印中的任意 3 个均可以恢复原水印。实验时, 测得数据库中共有 14 121 个字段值可以嵌入信息。嵌入的水印信息为 “central south”。

子集攻击时, 分别按 10%、20%、30%、40%、50%、60%、70%、80%、90%、100% 随机添加、修改、选择数据库中不同的记录。对水印进行鲁棒性仿真结果如图 3 ~5 所示。

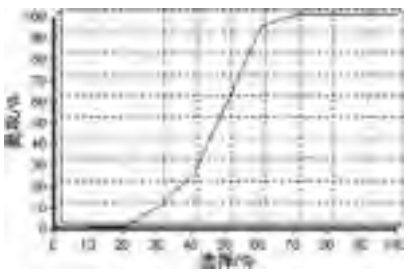

图 3　子集选取攻击

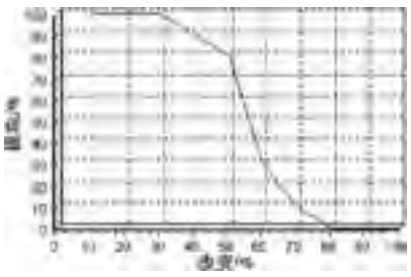

图 4　子集改变攻击

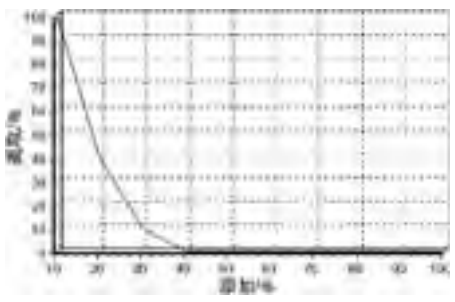

图 5　子集添加攻击

由图 3 可知, 对于子集选取攻击, 选取的数据比例越大, 选取的字段值就越多。相应地可用提取水印信息的字段越多, 因此选取的比例越大时, 准确恢复的比例越高。

由图 4、5 可知, 在受到子集增加或修改攻击时, 攻击的比例越大, 对水印数据破坏的程度越高, 因而准确恢复的比例越低。

3　算法性能分析

1) 鲁棒性　该算法采用了(r, n) 门限方案[8~10], 而且采用三元组, 使可嵌入的水印信息量加大, 在多个有效字段中嵌入相同的水印, 使恢复水印时仅仅需要部分分存水印, 大大提高了抗子集的攻击能力, 具有较好的鲁棒性, 可以防御各种各样的攻击, 包括正常的数据更新和恶意的攻击[3,4]。特别地, 该算法提出了基于可信中心的防添加攻击的思想, 即盗版者对盗版数据库嵌入自己的水印, 宣称自己对该数据库的所有权时和发生版权纠纷时, 数据库的原始所有者可以通过可信中心检测出该数据库含有自己的水印而取得所有权。

2) 不可见性高　选取有效的字段在误差允许范围之内对字段值进行嵌入, 水印不易被用户察觉, 也不会因为加了水印而影响关系数据的使用。

3) 嵌入的信息多　采用三元组的思想, 在一个字段值中就可嵌入 3 bit, 在各个有效的字段中均可嵌入相同的信息。

4) 误判率小　采用(r, n) 门限方案和各有效字段均嵌入水印, 恢复和检测时, 根据参数值, 只需要各有效字段中有 r 份不同的子水印, 即可提取出完整的水印。　(下转第 148 页)

5) (P |≡Q | ~(X, Y)) /(P |≡Q | ~X)

P 相信 Q 发送过消息(X, Y) , 则 P 相信 Q 发送过消息 X。

6) (P |≡#(X)) /(P |≡#(X, Y))

P 相信消息 X 的新鲜性, 则 P 相信消息(X, Y) 的新鲜性。

根据本协议中所用到的 WPKI 机制, 增加如下一条规则:

规则 1 当 P ◁K, P ◁M', P ◁{ M}K^{-1} 时, 若 M'= M, 则 P |≡K。

证明: 根据理想公钥体制, 对于 K^{-1}, 有且只有 K, 使得 K 和 K^{-1}可以互相解密, P 看到了 K, 并且用 K 可以成功解密 K^{-1}, 即 M'= M, 则 P |≡K。

3.2 协议目标

根据 BAN 逻辑中安全目标一般表示方法, 该协议的认证目标可以表示如下:

a) MS_2 相信 K_{MS_1}是 MS_1 的公钥, 即 $MS_2 \mid\equiv \mid \xrightarrow{K_{MS_1}} MS_1$。

b) MS_1 相信 K_{MS_2}是 MS_2 的公钥, 即 $MS_1 \mid\equiv \mid \xrightarrow{K_{MS_2}} MS_2$。

3.3 协议描述及分析

根据 WTLS 证书的格式, 证书的数据格式可以表示为 version, algorithm identifier, name of issuer, T, MS, K, { (version, algorithm identifier, name of issuer, T, MS, K) K_h} K_{CA}^{-1}, 省略与安全无关的项, 则协议可以详细描述为如下形式:

MS_1 - > MS_2(message1) : K_{MS_1}, T_{MS_1}, { (MS_1, K_{MS_1}, T_{MS_1}) K_h} $K_{CA_1}^{-1}$

MS_2 - > CA_1(message2) : Root Cert Req

CA_1 - > MS_2(message3) : K_{CA_1}, T_{CA_1}, { (CA_1, K_{CA_1}, T_{CA_1}) K_h} $K_{CA_1}^{-1}$

MS_2 - > MS_1(message4) : K_{MS_2}, T_{MS_2}, { (MS_2, K_{MS_2}, T_{MS_2}) K_h} $K_{CA_2}^{-1}$

MS_1 - > CA_2(message5) : Root Cert Req

CA_2 - > MS_1(message6) : K_{CA_2}, T_{CA_2}, { (CA_2, K_{CA_2}, T_{CA_2}) K_h} $K_{CA_2}^{-1}$

协议的详细分析如下:

对于 message3, 由规则 1 可知, 若 MS_2 验证通过, 可得

$$MS_2 \mid\equiv \mid \xrightarrow{K_{CA_1}} CA_1$$

对于 message1, 可得

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{MS_2 \mid\equiv \mid \xrightarrow{K_{CA_1}} CA_1 \quad MS_2 \quad \{(MS_1, K_{MS_1}, T_{MS_1})\, K_h\}\, K_{CA_1}^{-1}}{MS_2 \mid\equiv CA_1 \mid\sim (MS_1, K_{MS_1}, T_{MS_1})\, K_h}}{MS_2 \mid\equiv CA_1 \mid\sim MS_1, K_{MS_1}, T_{MS_1}\ MS_2 \mid\equiv \#(T_{MS_1})}}{MS_2 \mid\equiv CA_1 \mid\equiv K_{MS_1} \quad MS_2 \mid\equiv CA_1 \mid]\ K_{MS_1}}}{MS_2 \mid\equiv K_{MS_1}}}{MS_2 \mid\equiv \mid \xrightarrow{K_{MS_1}} MS_1}$$

同理, 可得 $MS_1 \mid\equiv \mid \xrightarrow{K_{MS_2}} MS_2$。

综上所述, 该协议基于 WPKI 机制, 完成移动终端间的身份认证, 达到了安全目标。

4 结束语

第三代移动通信系统的目标是全球范围的高度一致性、全球范围内使用的移动终端、全球的漫游能力, 对于不同的运营商来说, 保证相互之间的互连互通是非常必要的。但是不同的移动运营商有不同的安全设置和管理, 要想获得端到端的安全也是不容易的。本文设计了这样一种端到端的认证协议, 保证了不同运营商中移动终端的安全性, 形式化的分析也表明协议达到了认证目标。

参考文献:

[1] 3G TS 33.102, 3rd generation partnership project; technical specification group services and system aspects; 3g security; security architecture[S]. 2000.

[2] 刘东苏, 韦宝典, 王新梅. 改进的 3G 认证与密钥分配协议[J]. 通信学报, 2002, 23(5): 119-122.

[3] WAP-217-WPKI. Wireless application protocol public key infrastructure definition[S]. 2001.

[4] ITU-T Recommendation X. 509 |ISO/IEC 9594-8, Information technology-open systems interconnection the dircetory: public-key and attribute certificate frameworks[S]. 2001.

[5] BURROWS M, ABADI M, NEEDHAM R. A logic of authentication [J]. ACM Transactions on Computer Systems, 1990, 8(1): 18-36.

(上接第 145 页)

5) 盲测 该算法在检测关系数据库的水印时, 不需要原始数据库。

4 结束语

目前, 国内外对数据库的水印技术的研究还处于起步阶段, 本文提出的基于可信中心和密钥分存思想的数据库水印技术, 具有较好的应用前景。

参考文献:

[1] 易开祥, 石教英, 孙鑫. 数字水印技术研究进展[J]. 中国图象图形学报, 2001, 6(2): 111-117.

[2] 陈明奇, 钮心忻, 杨炎先. 数字水印的研究进展和应用[J]. 通信学报, 2001, 22(5): 71-79.

[3] AGRAWAL R, KIERNAN J. Watermarking relational databases [C]//Proc of the 28th VLDB Conference. Hong Kong: [s. n.], 2002: 155-166.

[4] SION R, ATALLAH. M, PRABHAKAR S. Rights protection for relational data[C]//Proc of the ACM SIGMOD International Conference on Management of Data San Diego. California: ACM SIGMOD, 2003: 98-109.

[5] 牛夏牧, 赵亮, 黄文军, 等. 利用数字水印技术实现数据库的版权保护[J]. 电子学报, 2003, 31(12A): 2050-2053.

[6] ZHANG Zhi-hao, JIN Xiao-ming, WANG Jian-min. Watermarking relational database using image [C]//Proc of the 3rd International Conference on Machine Learning and Cybernetics. Shanghai: [s. n.], 2004: 1739-1744.

[7] 黄敏, 张浩, 曹加恒. 一种基于关系数据库的水印技术[J]. 计算机工程与应用, 2005, 41(10): 153-155, 199.

[8] ASMUTH C, BLOOM J. A modular approach to key safeguarding [J]. IEEE Transaction on Information Theory, 1983, 29(2): 208-210.

[9] 施荣华. 一种多重密钥共享认证方案[J]. 计算机学报, 2003, 26(5): 552-555.

[10] 施荣华. 一种基于 Harn 数字签名的双向认证访问控制方案[J]. 计算机学报, 2001, 24(4): 400-404.

[11] SION R, ATALLAH M, PRABHAKARL S. Watermarking relational databases[R]. [S. l.]: CERIAS, 2002.