第41卷第7期 中国科学技术大学学报 Vol. 41, No. 7
2011年7月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA Jul. 2011

文章编号:0253-2778(2011)07-0594-05

基于二维区间 Hash 链的 RFID 安全协议

熊宛星,薛开平,洪佩琳,麻常莎

(中国科学技术大学电子工程与信息科学系信息网络实验室,安徽合肥 230027)

摘要:作为未来物联网(IOT)的核心技术之一,无线射频识别(RFID)系统由于设备的局限性而存在许多安全问题.在分析几种典型安全协议核心思想的基础上,提出了基于二维区间(two-dimensional region,TDR)Hash 链的安全协议.该协议以区间划分的方式标识各链,从而提高了数据库的检索效率;同时,由于在协议中引入了随机性,RFID 系统的安全性得到了进一步增强.

关键词:RFID;安全协议;二维区间 Hash 链

中图分类号:TP309.2 **文献标识码**:A doi:10.3969/j.issn.0253-2778.2011.07.005

RFID cryptographic protocol based on two-dimensional region Hash chain

XIONG Wanxing, XUE Kaiping, HONG Peilin, MA Changsha

(*Information Network Lab, Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230027, China*)

Abstract: Due to the limitation of relevant devices, a lot of security problems exist in a radio frequency identification (RFID) system, one of the core technologies of the future internet of things (IOT). A new protocol based on the two-dimensional region (TDR) Hash chains was proposed after the core ideas of several typical RFID cryptographic protocols were analyzed. TDR could significantly improve the efficiency of database retrieval by identifying each Hash chain with region division. Moreover, a random number was introduced to further enhance the security of RFID systems.

Key words: RFID; cryptographic protocol; two-dimensional region Hash chain

0 引言

无线射频识别(RFID)的基本原理是通过空间电磁感应或者电磁传播获取相关信息,以自动识别被标识对象.作为一种方便快捷的物品信息收集手段,RFID 技术是未来物联网不可或缺的组成部分[1].RFID 系统通常由标签、阅读器、后台数据库组成.标签从阅读器的电磁场中获取能量,向阅读器发送自身 ID,后台数据库中包含所有标签的相关信息,阅读器以 ID 为关键字从中提取需要的信息,从而实现了非接触的物品监控和信息管理[2].

相比阅读器与后台数据库的交互,阅读器与各标签的通信信道暴露在公共环境之下,通常被认为是不安全的.随着 RFID 的广泛应用,用户的安全和隐私问题日益凸显,携带标签的用户面临着被跟踪、假冒、重放、标签复制等多种攻击的危险.通过将安全协议集成在 RFID 系统中,能防止大部分对 RFID 系统的攻击,保证 RFID 系统的安全[3].但 RFID 标签的计算能力和存储能力有限,很多复杂的密码学算法无法使用.目前已经提出一些轻量级的方法[4],

收稿日期:2011-04-28;**修回日期**:2011-06-23

基金项目:国家自然基金(60903216),国家科技重大专项(2011ZX03005-006)资助.

作者简介:熊宛星,男,1987 年生,硕士生.研究方向:物联网网络体系结构和安全.E-mail:xwx1226@mail.ustc.edu.cn

通讯作者:洪佩琳,博士/教授.E-mail:plhong@ustc.edu.cn

包括物理方法、重加密方法、最小限度方法以及基于 Hash 的方法等，其中以 Hash 的方法最具代表性，受到人们的广泛关注.

目前已经提出的基于 Hash 方法的 RFID 安全协议包括 Hash 锁、随机 Hash 锁、Hash 链以及相应的改进协议[5-9]. 本文在 Hash 链协议的基础上，提出了二维区间 Hash 链机制，核心思想是：对各 ID 的 Hash 链进行修改，通过区间划分标识各链，保证后台数据库检索的效率；同时引入了随机性，实现标签和阅读器的双向认证，进一步增强 RFID 系统的安全性.

本文接下来的章节安排如下：节 1 对使用 Hash 函数的 RFID 安全机制进行综述；节 2 对二维区间 Hash 链机制进行详细说明；检索效率及安全性分析在节 3；最后是对全文的总结.

1　相关工作

Hash 函数是一种轻量级的密码学组件，便于在 RFID 标签内实现，由函数本身的不可逆性来保证 RFID 交互的安全性. Sarma 等[5]最早提出了 Hash 锁方案，其协议流程如图 1 所示. 标签在收到查询请求后会返回一个 Hash 值，称为 metaID，只有当阅读器返回正确的逆向 Hash 值 Key 时，标签才认证其身份的正确性，并向阅读器返回正确的 ID. Hash 锁方案强调标签对阅读器的认证，其不足之处在于：第 2 步中的 metaID 是固定不变的，且第 6 步中标签的 ID 是以明文发送的，容易被跟踪和假冒，从而无法实现对标签本身的身份认证.

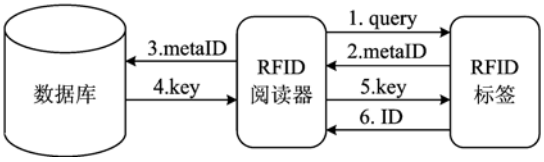

图 1　Hash 锁协议交互流程

Fig. 1　Process of Hash lock protocol

Weis 等[6]提出了改进的随机 Hash 锁方案，引入随机数来改变标签传输的 metaID 固定不变的问题，其协议流程如图 2 所示. 标签每次被请求时返回随机数 R 及 Hash 值 $H(\mathrm{ID}\,\|\,R)$，阅读器向数据库请求所有标签的 ID 值，之后进行大量计算比较以确定对应标签的 ID 值. 协议在阅读器上的检索效率较低，第 5 步的 ID 同样是明文传输，没有从根本上解决标签被跟踪和假冒的潜在问题.

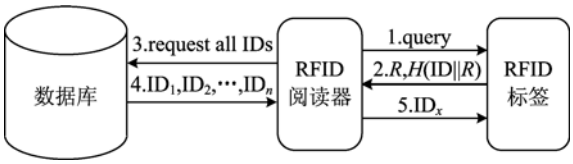

图 2　随机 Hash 锁协议交互流程

Fig. 2　Process of random Hash lock protocol

相比上述方案，Hash 链方案[7]引入了两个 Hash 函数 $H(*)$和 $G(*)$，其中 $H(*)$用于 Hash 链的更新，$G(*)$用于在 Hash 链值被保护的前提下计算实际的传输值. 后台数据库维护的针对每个标签的一串密值就对应于各标签的 Hash 链，具体地，第 i 个标签初始时的密值记为 S_i^0，之后第 j 次通信时的密值为 S_i^j，密值间满足关系式 $S_i^{j+1}=H(S_i^j)$. 其具体的协议流程如图 3 所示：第 j 次通信时，第 i 个标签将自身当前密值 S_i^j 的 Hash 结果 $a_i^j=G(S_i^j)$ 发送给阅读器，同时标签使用 $H(*)$函数将密值更新到 S_i^{j+1}，后台数据库对各个标签的 Hash 链进行检索，以得到对应标签的 ID. Hash 链方案的优势是标签不用传输 ID，从而避免了跟踪. 缺点是后台数据库缺少对各 Hash 链的状态维护，使得重放攻击成为可能；此外，阅读器可以认证标签的身份，但标签无法认证阅读器的身份.

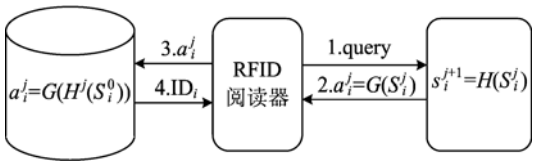

图 3　Hash 链协议交互流程

Fig. 3　Process of Hash chain protocol

文献[8]提出对 Hash 链方案的随机化改进，在每次通信时引入随机数 R，第 2 步的传输值为 $G(S_i^j\,\|\,R)$，从而使得每次标签的响应都不同，能成功对抗重放攻击. 但后台数据库对某个 Hash 值进行检索以认证标签时，需要该随机数参与运算，从而对后台的计算和检索能力要求很高，在管理较多标签时协议执行的效率会受到影响. 文献[9]修改了第 2 步中传输的值，通过两次 G 函数的运算得到 $G(S_i^j)\,\|\,G(a_i^j\,\|\,\mathrm{DB})$，其中 DB 是数据库的标识. 可以证明，此方式可以实现双向认证，但检索效率依然较低.

2　基于二维区间 Hash 链的安全协议

经典的 Hash 链方案没有实现双向认证，而其

改进方案的检索效率较低，针对以上问题，我们提出了二维区间 Hash 链的方案，简称为 TDR 方案. TDR 方案将原本的 Hash 链结构进行了扩展：标签 i 除了持有当前密值 S_i^j，还要保存 A_i+B_i 个最近使用过的传输值 $a_i^{j-A_i-B_i}$，$a_i^{j-A_i-B_i+1}$，…，a_i^{j-1}；后台数据库则持有每个标签的当前密值 S_i^j，及 A_i+B_i 个历史密值 $S_i^{j-A_i-B_i}$，$S_i^{j-A_i-B_i+1}$，…，S_i^{j-1}. 区间的分布如图 4 所示，箭头代表执行一次 Hash 运算，标签存储的内容即图中的阴影值，分为 3 个部分：当前密值、A_i 区间和 B_i 区间. 其中，A_i 区间用于标识检索，B_i 区间用于标签对阅读器认证，两者合称为二维区间.

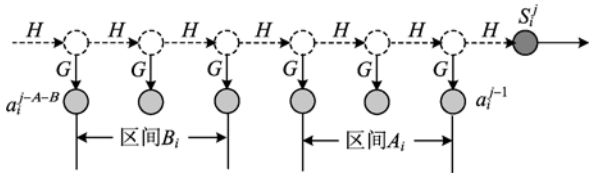

图 4 二维区间 Hash 链方案标签的存储内容

Fig. 4 Storage content of tags in TDR Hash chain scheme

初始时，后台数据库对每个被管理的标签 i 指定一个索引标识 (A_i, B_i)，索引标识决定了数据库中保留的相应标签的 Hash 链的长度，及各标签的 A_i，B_i 区间大小. 各标签的 A_i 值是平均分布的，即允许多个标签对应相同的值；而 A_i+B_i 值会受到 RFID 标签 i 的存储能力限制.

基于二维区间 Hash 链的安全协议完整流程如图 5 所示：

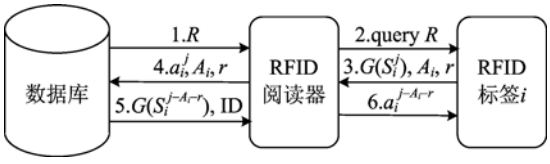

图 5 基于二维区间 Hash 链协议的交互流程

Fig. 5 Process of TDR Hash chain protocol

①当阅读器发起通信会话，由后台数据库生成随机数 R，并将 R 发送给阅读器；

②阅读器向标签发送 Query 请求，并将 R 发送给标签；

③标签 i 按已分配的标识 (A_i, B_i) 计算 $r=(R \bmod B_i)+1$，从而 r 是属于区间 $[1, B_i]$ 的正整数，标签根据当前密值 S_i^j，发送 3 个值 $a_i^j=G(S_i^j)$，A_i，r，并暂存 r；

④阅读器收到 a_i^j，A_i，r 并转发给后台数据库，数据库对所有标签的 Hash 链 (X, Y) 进行标识索引，在满足 $X==A_i$ 且 $r==(R \bmod Y)+1$ 的链中进一步寻找 S_i^j，使得 a_i^j 和 $G(S_i^j)$ 相等. 如找到，则对标签的认证成功；

⑤为了使标签能认证阅读器的身份，数据库发送 $a_i^{j-A_i-r}=G(S_i^{j-A_i-r})$ 及该标签 ID 给阅读器，并将对应数据项的密值更新为（或增加）$S_i^{j+1}=H(S_i^j)$；

⑥阅读器将 $a_i^{j-A_i-r}$ 发送给标签，标签查看存储区间及 r，验证 $G(S_i^{j-A_i-r})$ 是否和 $a_i^{j-A_i-r}$ 一致，如一致，则完成对阅读器的认证，标签删除暂存的 r，并开始与阅读器的正式通信.

⑦该次通信结束时，节点需要将密值更新为 $S_i^{j+1}=H(S_i^j)$，并保留最新的 A_i+B_i 个 a_i^x 值，以实现和后台 Hash 链同步更新.

需要说明的是，最初各标签 Hash 链长度有限，RFID 系统只能使用经典的 Hash 链方案；当且仅当某标签 i 使用过的 a_i^x 个数超过 A_i+B_i 时，开始使用二维 Hash 链方案来生成和管理此标签的 Hash 链.

方案的优势体现于：保留了原 Hash 链方案标签防跟踪的特性，能实现阅读器对标签的认证；通过对二维区间 Hash 链标识 (A, B) 的索引，可以加快后台对标签 Hash 值的检索速度，减少不必要的计算比较；新方案中引入的随机性，满足了标签对阅读器和后台的认证需求，在不增加检索量的同时，提升了对阅读器及数据库认证的可信度. 具体的分析见下一节.

3 检索效率及安全性分析

3.1 检索效率分析

假定后台数据库对应唯一的阅读器，阅读器可管理 N 个标签，一次比较的开销是单位时间，计算与所有标签通信一轮的平均检索开销. 本节首先对两种 Hash 链方案的协议性能进行评估和比较，主要是指后台数据库对标签的 Hash 值的检索效率. 参与比较的方案是经典 Hash 链方案和二维区间 Hash 链方案，随机化 Hash 链方案还要考虑使随机数参与进来的计算开销，在提升安全性的同时会比经典 Hash 链方案的效率更低，故不在此处列入比较范围.

检索效率的提升体现在两个方面：首先，数据库中的每个标签对应的 Hash 链以 (X, Y) 标识，只需在满足 $X==A_i$ 且 $r==(R \bmod Y)+1$ 的链中进

行检索即可，故整体上需要检索的 Hash 链数目大大减少；其次，在标签 i 的链中，只需要保留最多 A_i+B_i+1个链值即可，包括一个当前密值 S_i^j，A_i+B_i 个历史值 a_i^x，相比原方案，每条链中的检索数与计算量都减少了.

TDR 方案中，假定所有标签的 A_i 均匀分布于区间$[1, A_{max}]$，且每条链中只要比较一次最新密值，我们通过仿真比较新方案在不同 A_{max}值及不同标签规模情况下的平均检索时间，结果如表 1 所示. 由表 1 可以看出，随着标签规模的增加，各 A_{max}值方案的检索时间也增加，并且两者成正比关系. 另外 A_{max}越大，则实际参与检索的 Hash 链的数目越少，检索的效率越高. 经典 Hash 链方案的检索时间在上述假定下可以简单地通过 N/2 计算得到，由表中数据易看出 TDR 方案的检索时间非常接近经典方案 $1/A_{max}$，即平均检索时间和 A_{max}成反比关系.

表 1　A_{max} 值对检索效率的影响

Tab. 1　Influence of A_{max} to retrieval efficiency

N	A_{max}					
	3	4	5	6	7	8
100	17.45	13.33	10.81	9.21	8.06	7.17
200	34.17	25.83	20.83	17.53	15.21	13.39
300	50.72	38.3	30.86	25.9	22.31	19.65
400	67.41	50.81	40.89	34.22	29.41	25.89
500	84.1	63.44	50.78	42.59	36.67	32.12
600	100.8	75.8	60.94	50.87	43.74	38.44
700	117.46	88.34	70.86	59.18	50.87	44.68
800	134.08	100.89	80.84	67.53	58.08	50.97
900	150.67	113.39	90.92	75.9	65.22	57.11
1000	167.47	125.14	100.8	83.17	72.39	63.44

但实际使用时，由于标签本身的存储能力限制，A_{max}不可能很大，而且标签在更新 Hash 区间时，还有数据的复制和移动等过程，导致 A_{max}也不宜过大.

3.2　安全性分析

（Ⅰ）阅读器对标签的认证：G 是单向的 Hash 函数，攻击者只能监听到标签认证时传输的 a_i^j，A_i，并不能得到 Hash 链中的密值 S_i^j，且每次通信标签都要更新密值，因此，只要后台数据库能检索到相同的密值 S_i^j，即可确认标签身份的正确性. 另外，新方案中对标识为（A_i，B_i）的标签，有关系 $r=(R \bmod B_i)+1$，也从另一个角度验证了标签身份.

（Ⅱ）重放攻击：随机数 R 是每次通信时由后台数据库生成的，不能提前得知，标签在流程第③步传输的 a_i^j，r 每次都是不同的，因而攻击者无法使用上次传输值来进行重放；对于第⑥步，阅读器每次传输的 $a_i^{j-A_i-r}$ 也是不同的，攻击者无法重放上次的值来冒充阅读器.

（Ⅲ）不可追踪性：在协议执行过程中，没有出现明文的标签 ID，并且传输的内容是变化的，因此，攻击者无法确认通信与某个标签有关联，从而无法追踪该标签.

（Ⅳ）标签对阅读器的认证：阅读器需要向标签出示之前使用过的 $a_i^{j-A_i-r}$ 以证明身份，虽然攻击者可能截获过 $a_i^{j-A_i-r}$，但无法得知标签 i、密值 j 及随机数 r 的具体情况，因而无法推断并使用 $a_i^{j-A_i-r}$，标签可以认为通过认证的阅读器身份是真实的.

（Ⅴ）接入控制：标签只在收到第⑥步的 $a_i^{j-A_i-r}$，并实现对阅读器的认证后，才与阅读器正式通信，因此协议满足了接入控制的需求.

（Ⅵ）哄骗或修改攻击：和重放攻击一样，哄骗或修改数据的攻击方式无法通过第④步的验证；如果是修改第⑥步的数据，则无法通过标签的验证.

3.3　优化分析

TDR 方案中，以二维区间的方式对各 Hash 链作了标识. 对标识为（A_i，B_i）的 Hash 链，其对应的标签需要存储 A_i+B_i 个历史密值. 其中，A_i 取值区间上界 A_{max}的大小对检索的效率有影响，A_{max}越大则各链的分类越细致，检索时效率越高；而各标签每次的 r 是属于区间$[1, B_i]$的，B_i 的大小对随机数 r 的变化性有影响，基于安全性考虑，标签的 B_i 越大则安全性越高. 但标签的存储能力有限，设为 T，导致 $B_i=T-A_i$，实际中需要在各标签的 A_i 和 B_i 的分配上作均衡的考虑.

对每个存储能力有限的标签 i，希望 A_i 和 B_i 尽量大，则设定评价指标：$Q_i=A_i * B_i$. 对于所有 N 个标签，评价指标和为：

$$Q=\sum_{i=1}^{N}Q_i=\sum_{i=1}^{N}(A_i * B_i) \tag{1}$$

仍假定所有标签的 A_i 均匀分布于区间$[1, A_{max}]$，则所有标签按 A_i 值可分为 A_{max}类，每类 k 中的 N/A_{max}个标签评价指标为 $k*(T-k)$，则式(1)变化形式为

$$Q=\sum_{k=1}^{A_{max}}\sum_{i=1}^{N/A_{max}}(k*(T-k)) \tag{2}$$

进一步计算得到结果如式(3)所示：

$$Q=(-2A_{max}^2+3(T-1)A_{max}+(3T-1))*N/6 \tag{3}$$

Q是一个开口向下的二次函数,根据二次函数的性质,其最大值在导函数为 0 时取得,此时有

$$A_{max}^{best} = 3(T-1)/4 \quad (4)$$

A_{max}^{best}可能不是个整数,Q的最优值在此值附近的整数处取得.图 6 给出了各种 T 限制下 Q 随 A_{max} 变化的具体情况,从而可以验证此结论.

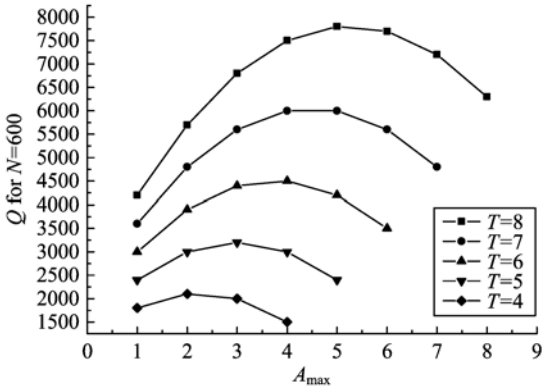

图 6 不同存储限制下的评价指标

Fig. 6 Metric for different storage limitation

4 结论

RFID 是未来物联网的核心技术,但由于 RFID 标签的存储和计算能力相对较低,设计高效可靠的 RFID 安全协议是相当困难的.本文针对经典的 Hash 链方案提出改进,以二维区间标识的方式对各链增加标识,在加快检索效率的同时,以随机性增强了系统对抗重放攻击的能力,并实现了阅读器和标签的双向认证,在协议设计上给出了令人满意的回答,具有较高的实用价值.TDR 方案在实际的应用中受到标签的存储空间限制,通过合理地调整整体的 A_i 取值区间上界 A_{max},能够均衡各标签 A_i 和 B_i 的取值,实现在检索效率和安全性上的进一步优化.

参考文献(References)

[1] Violino B. Leveraging Internet of things [J]. RFID Journal,2005: 1-2.

[2] Han S, Dhillon T S, Chang E. Anonymous mutual authentication protocol for RFID tag without back-end database [C]//Proceedings of the 3rd International Conference on Mobile Ad-hoc and Sensor Networks. Berlin: Springer-Verlag, 2007: 623-632.

[3] Knospe H, Pohl H. RFID security [J]. Information Security Technical Report, 2004, 9(4): 39-50.

[4] Juels A. RFID security and privacy: A research survey [J]. IEEE Journal on Selected Areas in Communication, 2006, 24(2): 381-394.

[5] Sarma S E, Weis S A, Engels D W. Radio frequency identification: Secure risks and challenges [J]. RSA Laboratories Cryptobytes, 2003, 6(1): 2-9.

[6] Weis S A, Sarma S E, Rivest R L, et al. Security and privacy aspects of low-cost Radio Frequency Identification systems [J]. Lecture Notes in Computer Science, 2004, 2802: 50-59.

[7] Ohkubo M, Suzuki K, Kinoshita S. Hash-chain based forward secure privacy protection scheme for low-cost RFID [C]// Proceedings of the 2004 Symposium on Cryptography and Information Security (SCIS 2004). Sendai, 2004: 719-724.

[8] Li Zhanglin, Lu Guizhang, Xin Yunwei. Scalable hash-chain-based RFID authentication protocol [J]. Computer Engineering, 2008, 34(4): 173-175.
李章林,卢桂章,辛运帏.基于 Hash 链的可扩展 RFID 验证协议 [J].计算机工程,2008,34(4):173-175.

[9] Zhang Yaling, Guo Hu. An improved RFID privacy protection scheme based on Hash-chain [C]// Proceedings of the 2010 International Conference on Logistics Engineering and Intelligent Transportation Systems (LEITS 2010). New York: IEEE Press, 2010: 1-4.