CRL 分段2过量发布综合模型研究

谭 良[1,2], 刘 震[1], 佘 [1], 周明天[1]
(1. 电子科技大学计算机科学与工程学院 电子科技大学2卫士通信息安全实验室, 四川成都 610054;
2. 四川师范大学电子工程学院, 四川成都 610066)

摘 要: 提出了一种应用证书撤消列表 CRL(Certificate revocation List)发布公钥基础设施 PKI(Public key infrastruc2ture)证书状态信息新模型: 分段2过量发布综合模型, 该模型采用先将 CRL 分段, 然后各段独立过量发布的方式来实现. 通过分析表明, 该方式既可以减少 CRL 的长度, 使存储库以更快的速度提供请求服务, 又可以降低峰值请求率、峰值带宽和平均负荷, 减少时间碎片, 满足大规模 PKI 对证书撤消的要求.

关键词: 信息安全; 证书撤消列表 (CRL); 分段 CRL; 过量发布 CRL; 分段2过量发布; PKI
中图分类号: TP311 文献标识码: A 文章编号: 0372-2112 (2005) 02-0227-04

Research on the Segmented and Over2Issued CRL Synthesis Model

TAN Liang[1,2], LIU Zhen[1], SHE Kun[1], ZHOU Ming2tian[1]
(1. School of Comp. Sci. & Engn., Univ. of Electronic Sci. & Tech. of China Information Security United Lab of UESTC2Westone, Chengdu, Sichuan 610054, China; 2. College of Electronic Engineering, Sichuan Normal University, Chengdu, Sichuan 610066, China)

Abstract: An improved model: the segmented and over2issued CRL synthesis model, which issues certificate status information with certificate revocation list in public key infrastructure, is presented. It is realized by that CRL is segmented first, and then over2is2sued. Compared to other models, the improved model minimizes the size of CRL which can accelerate to request service, as well as the peak request rate, peak bandwidth, average loads and time piece on CRL repository. The improved model is fit for the large2scale PKIs according to the requirements.

Key words: information sacurity; certificate revocation list (CRL); segmented CRL; over2issued CRL; segmented and over2is2sued CRL; public key infrastructure(PKI)

1 引言

证书撤销是公开密钥基础设施中的一个关键性操作, 为此, 许多证书撤消机制[1~4]被提出. 目前, 关于证书撤消问题的解决方案主要是 X509 证书系统[5]的证书撤消列表, 商用 CA 或 PKI 系统主要使用定期颁布 CRL 来分发证书状态信息.

为了适应 PKI 规模的变化, 现有方法主要侧重通过更改 X.509 证书结构来对 CRL 的时间与空间特性进行改进, 而对存储性能的优化, 尤其是如何有效降低存储库峰值负荷方面仍有一些未解决的问题. 为此, 许多学者在这方面做了有益的探索, 建立了多个数学模型, 具体包括: 分段 CRLs[6], Delta2CRLs[7]和过量发布 CRLs[8]. 这些模型对如何优化 CRL 的存储性能、降低存储库峰值负荷、提高发布性能具有指导作用, 但均不完善. 特别是随着 Internet 的发展, CA 拥有的端实体数目迅速增加, 端实体对 CRL 的发布性能提出了更高的要求. 因此, 这些方法对大规模的 PKI 中是不实用的.

本文针对当前 PKI 应用规模的变化, 提出了一种新模型: 分段2过量 CRL 发布综合模型, 该模型采用先将 CRL 分段, 然后各段独立过量发布的方式来实现. 通过与其它模型比较表明, 通过分析表明, 该方式既可以减少 CRL 的长度, 使存储库以更快的速度提供请求服务, 又可以降低峰值请求率、峰值带宽和平均负荷, 减少时间碎片. 满足大规模 PKI 对证书撤消的要求.

2 PKI 应用规模的变化

X509 最初是在 80 年代中期开始设计的, 那时的 Internet 没有向现在这样爆炸性的发展, 它们被设计为离线环境中运行. 在这种情况下, 计算机只是偶尔地连接起来, 使用 CRL 的方式非常简单.

随着 Internet 的发展, PKI 的应用规模逐渐增加, CA 拥有的端实体数目迅速增加, CRL 可能变得很大, 分发 CRL 将占用太多的网络资源, 信任方得到它可能会用困难, 因为它们访问 CA 的带宽是有限的. 并且由于 CRL 是由 CA 签过名的, 在使用之前必须检验它的签名, 检验一个庞大的 CRL 签名, 所

收稿日期: 2004-03-25; 修回日期: 2004-08-01
基金项目: 现代通信国家重点实验室基金(No. 51436050203DZ0210)

花的时间将会很长. 再则, 随着 CRL 的增大, 对信任方和应用程序提出了更大的存储要求. 另一方面, 随着 CA 拥有的端实体数目迅速增加, CRL 库不能很好处理进来的所有访问, 特别是请求率达到高峰时, 部分请求失去了合理的响应时间.

因此, 大规模的 PKI 对 CRL 分发机制提出了更高的要求, 就是尽可能减小信任方所需下载的 CRL 并应降低 CRL 峰值请求率、峰值带宽和平均负荷.

3 分段2过量发布综合模型

为了改善大规模 PKI 的 CRL 发布性能, 可以采用先将 CRL 分段, 然后各段独立过量发布来实现. 下面建立分段2过量发布 CRL 模型, 这里假设 CRL 被随几分段, 并且每个确认企图尽可能地访问每个分段.

如果新的 CRL 的第 1 段在时刻 0 发布, 那么, 当且仅当信任实体在[t, t+ dt] 内确认证书, 并且确认需要用到第 1 段 CRL 而信任实体在[0, t] 内又没有第 1 段 CRL 的证书时, 信任实体才会在[t, t+ dt] 内向存储库请求第 1 段 CRL.

首先, 要确定信任实体在[0, t] 没有请求第 1 段 CRL 的概率. 假定确认企图是以指数函数相互到达的, 从泊松定律可知在时间长度 t 内发生 n 个确认企图的概率是:

$$\left[\frac{(vt)^n}{n!}\right] e^{-vt} \tag{1}$$

如果有 s 个 CRL 分段, 执行任何确认企图时需要用到 CRL 第 1 段的概率为 1/ s, 则对于任意 n 个确认企图不需要第 1 段的概率为:

$$\left(1-\frac{1}{s}\right)^n \tag{2}$$

结合式(1)、(2)可得出任何信任实体在[0, t] 内不请求第 1 段的概率:

$$\sum_{n=0}^{\infty}(1-1/s)^n\left[\frac{(vt)^n}{n!}\right] e^{-vt} \tag{3}$$

其次, 要确定信任实体在[t, t+ dt] 内请求第 1 段 CRL 的概率. 根据泊松定律, 在时间段[t, t+ dt] 内发生确认企图的概率为 $ve^{-vdt}dt = vdt(\lim_{dt\to 0} e^{-vdt}dt = 1)$. 任何确认企图需要用到第 1 段的概率为

$$vdt/s \tag{4}$$

结合式(3)、(4), 可得一个信任实体在[t, t+ dt] 内请求第 1 段的概率为

$$\sum_{n=0}^{\infty}(1-1/s)^n\left[\frac{(vt)^n}{n!}\right] e^{-vt}\cdot\frac{vdt}{s} \tag{5}$$

化简得:

$$ve^{-vt/s}dt/s \tag{6}$$

而第 1 段又采用过量发布方式, 即在同一时刻, 同组的信任实体可能存在几个有效的第 1 段 CRL. 这样可以把向存储库的请求在一定程度上散开, 从而降低存储库的峰值请求率.

把第 1 段 CRL 发布所隔时间定义为一个时间间隔, 一个信任实体仅仅在给定的时间间隔内需要执行一次确认而且它的缓存中不存在未过期的第 1 段 CRL 时才会向存储库发送一个请求. 如果 O 代表给定间隔内有效的第 1 段 CRL 的数量, P_{val} 代表信任实体在给定的时间间隔内对第 1 段 CRL 执行确认的概率, 则信任实体在时间间隔 n 请求第 1 段 CRL 的概率为 P_{val} 乘以其在前 n- 1 个时间间隔内没有执行确认的概率, 即

$$p_{i,n} = p_{val}\left[1-\sum_{j=n-o+1}^{n-1} p_{i,j}\right] \tag{7}$$

当系统处于稳定状态时, 信任实体在连续的时间间隔内对第 1 段 CRL 执行确认的概率将相同, 即

$$p_{i,n} = p_{i,n-1} = \cdots = P_{i,1} \tag{8}$$

所以在稳定状态下:

$$p_i = p_{val}[1-(O-1)p_i] \tag{9}$$

解得:

$$p_i = \frac{p_{val}}{(O-1)p_{val}+1} \tag{10}$$

如果时间间隔从 0 时刻开始, 那么信任实体在时间 t 到 t + dt 内向 CRL 存储库发送请求第 1 段 CRL 的概率等于信任实体在时间段 t 到 t+ dt 内对第 1 段 CRL 执行它的第一次确认请求的概率乘以信任实体在缓存中没有有效的第 1 段 CRL 的概率, 信任实体在[t, t+ dt] 内向第 1 段 CRL 发送请求的概率由式(6)确定; 信任实体在缓存中没有有效第 1 段 CRL 的概率可以通过信任实体在时间间隔内请求第 1 段 CRL 的概率除以信任实体在时间间隔内对第 1 段 CRL 执行确认的概率计算出来(即式(10) 除以 P_{val}). 这样, 信任实体在[t, t+ dt] 内请求第 1 段 CRL 的概率为:

$$\frac{ve^{-vs/s}dt}{s(O-1)P_{val}+s} \tag{11}$$

由于确认呈指数概率分布, 信任实体在给定时间间隔对第 1 段 CRL 不执行确认的概率为 $e^{-vl/o}$, 这里 l 为 CRL 的有效期, l/ O 即为一个时间间隔, 所以:

$$P_{val} = 1-e^{-vl/o} \tag{12}$$

带入式(11), 得:

$$\frac{ve^{-vt/s}dt}{s(O-1)(1-e^{-vl/o})+s} \tag{13}$$

式(13)乘以信任实体总数 N 和除以总的分段数 s, 再除以 dt, 得在时刻 t 对存储库的总请求率:

$$\frac{Nve^{-vt/s}dt}{(O-1)(1-e^{-vl/o})+1} \tag{14}$$

峰值请求率为:

$$R(0) = \frac{Nv}{(O-1)(1-e^{-vl/o})+1} \tag{15}$$

当 O= 1, s= 1 时, 式(15) 即为传统方式发布 CRL 的峰值请求率. 当 O> 1, s= 1 时, 式(15) 即为过量发布 CRL 的峰值请求率; 当 O= 1, s> 1, 式(15) 即为分段发布 CRL 的峰值请求率; 当 CRL 不断发布时即 $\lim O\to\infty$.

$$\lim_{o\to\infty}\left[\frac{Nv}{(O-1)(1-e^{-vl/o})+1}\right] = \frac{Nv}{vl+1} \tag{16}$$

式(16)代表了分段) 过量 CRL 发布综合模型在理论上可达到的最小峰值请求率, 但实际中达到最小峰值请求率是不可能的.

假设 CRL 的有效期为 T, 则分段) 过量 CRL 发布综合模

型的平均请求率为

$$\overline{R(t)}=\frac{1}{T}\int_0^T\frac{Nve^{-vt/s}}{(O-1)(1-e^{-vl/o})+1}dt$$
$$=\frac{sN(1-e^{-vT/s})}{T[(O-1)(1-e^{-vl/o})+1]} \quad (17)$$

假设每个 CRL 首部的大小为 H_h, CRL 中每项的大小为 H_e. 假设系统中平均每天有 r 个证书被吊销, 证书有效期平均为 L_C 天, 每个证书从被吊销到该证书过期的平均时间为 $L_C/3$ 天. 基于上述假设, 一个完全 CRL 的平均大小为:

$$S_C=H_h+rH_eL_C/3$$

分段) 过量 CRL 发布综合模型的峰值带宽为:

$$B=S_C@R(0)=\frac{Nv}{(O-1)(1-e^{-vl/o})}[H_h+(rH_eL_C)/(3s)] \quad (18)$$

平均负荷为:

$$\overline{B}=S_C@\overline{R(t)}$$
$$=\frac{sN(1-e^{-vT/s})}{T[(O-1)(1-e^{-vl/o})+1]}[H_h+(rH_eL_C)/(3s)] \quad (19)$$

4　分段2过量综合模型与其它模型的比较

对于当前的 CRL 发布模型, 除传统模型外, 还有分段 CRLs, Delta2CRLs 和过量发布 CRLs, 下面将这几种模型的性能参数与传统模型进行定性比较, 见表 1.

表 1　模型定性比较

性能参数 / 发布模型	峰值请求率	平均请求率	峰值带宽	平均负荷
分段 CRLs	等于传统模型	大于传统模型	小于过量模型	小于传统模型
Delta2CRLs	等于传统模型	小于传统模型	小于传统模型且与时间跨度有关	小于传统模型且与时间跨度有关
过量模型	小于传统模型	小于传统模型	小于传统模型	小于分段模型
分段2过量模型	等于过量模型	大于传统模型, 但小于分段	小于分段模型	小于过量模型

从表 1 可以看出, 分段2过量发布模型较其他模型优, 既具有分段模型的优点, , 使存储库以更快的速度提供请求服务, 又可以降低峰值请求率、峰值带宽和平均负荷, 减少时间碎片, 但平均请求率增加了.

如果做如下假设: (1) CRL 的有效期是 24 小时; (2) CRL 在时刻 0 发布; (3) 共存在 300000 个信任实体, 即 N= 300000; (4) 每个信任实体平均每天确认 10 个证书, 即 v= 10; (5) 系统中平均每天吊销 3000 个证书, CRL 的首部大小为 51 个字节, CRL 中每项大小为 9 个字节; (6) 证书的有效期为 365 天, l= 24 小时, O= 4, 定量比较传统模型、过量模型、分段模型和分段2过量综合模型, 见表 2.

从表 2 可以看出, 对于分段模型和分段2过量模型, 随着分段数的增加, 峰值带宽减小, 平均负荷减小, 但平均请求率增加. 因此, 只要合理确定和调整参数 O 和 s, 就可以把平均负荷、平均请求率和峰值带宽控制在要求的范围内, 满足大规模 PKI 的要求. 下面给出相应的确定和调整算法.

表 2　模型定量比较

模型 \ 参数 / 值		峰值请求率 (个/s)	平均请求率 (个/s)	峰值带宽 (MB/s)	平均负荷 (MB/s)
传统模型		34.72	3.43	36.26	3.63
过量模型		9.25	0.92	9.66	0.96
分段模型	S= 10	34.72	21.95	3.63	2.29
	S= 100	34.72	33.04	0.36	3.45
	S= 1000	34.72	34.55	0.04	3.61
分段2过量综合模型	S= 10	9.25	5.85	0.97	0.61
	S= 100	9.25	8.80	0.097	0.092
	S= 1000	9.25	9.20	0.01	0.0097

算法:

(1) 根据 PKI 的信任实体数目 N 和信任实体平均每天确认证书的次数 v, 确定该 PKI 的最大峰值请求率 Nv

(2) 根据式(22) 确定该 PKI 系统能够达到的最小峰值请求率 Nv/(vl+1)

(3) 设 CRL 存储库所在服务器的处理能力 U, U 表示存储库所在服务器处理信任实体请求时, 能够及时响应的最大连接数(可用相关工具进行测试获得, 如对 Web 服务器, 可以用 SPEC Web99[9]), 此时存储库所在服务器得到最充分利用.

(4) 如果 U< Nv/(vl+1), 则退出(服务器的处理能力不够), 如果 U> Nv, 则不需要过量发布, 即 O= 1, 如果 Nv> U> Nv/(vl+1), 则 U 应大于或等于服务器的峰值请 R(0). 相等时, 处理器的资源得到充分利用, 可由下式:

$U=R(0)=Nv/[(O-1)(1-e^{-vl/o})+1]$ (注: $1-e^{-vl/o}\approx 1$)

解得:

$O\approx Nv/U$

即可确定 O 的值

(5) 设 CRL 存储库服务器的网络带宽为 W, 显然峰值带宽不能大于网络带宽, 所以有分段数:

$$s=\frac{S_C@R(0)}{W}=\frac{NvS_C}{W[(O-1)(1-e^{-vl/o})+1]}$$

代入 O 的值, 就可以确定 s 的值.

(6) 如果平均负荷和峰值带宽满足要求, 则此时的 O, s 的值即为最终的 O, s 值. 如果不满足要求, 需对 O, s 的值进行合理微调, 可固定 O 的值, 增加 s, 系统的峰值带宽和平均负荷将减小(注意: 不能无限增大 s, 因为 s 增大, 每段的长度减小, 每段的有效数据比例减小, 浪费网络带宽), 当平均负荷和峰值带宽满足大规模 PKI 的要求时, 可选择一个最佳的 s.

以第 1 部分所做的假设为例, 并假设 CRL 存储库所在的服务器的处理能力 U= 10 个/s, 网络有效带宽为 0.5M/s, 由算法第一步算得 Nv= 37.42, 第二步得 Nv/(vl+1)= 3.16,

则: Nv> U> Nv/(vl+ 1)

所以由算法第四步得:

O U 4

第五步得

s U 4

当 O= 4, s= 4 时, 平均负荷和峰值带宽都远小于网络有效带宽, 因此, O、s 不在需要调整.

5 结论

本文分析了发布证书的传统模型, 在此基础上, 提出了一种新的模型: 分段2过量发布综合模型, 通过比较可以看出, 该模型既具有分段模型的优点, 即 CRL 的长度小, 存储库以更快的速度提供请求服务; 又具有过量模型的优点, 即峰值带宽降低, 时间随片减小. 虽然分段2过量模型的平均请求率比过量模型大, 但只有合理调整参数 O 和 s, 就可以把就可以把平均负荷、平均请求率和峰值带宽控制在要求的范围内, 满足大规模 PKI 的要求.

参考文献:

[1] S Micali. Efficient Certificate Revocation[M]. Cambridge, MA, USA: Massachusetts Institute of Technology, 1996. 542- 563.

[2] Paul C Kochar. On certificate revocation and validation[A]. Proceedings of the Second International Conference on Financial Cryptography[C]. Berlin: Springer2Verlag, 1998. 171- 177.

[3] Moni, Naor, Kobbi, Nissim. Certificate revocation and certificate update[J]. IEEE Journal on Selected Areasin Communications, 2000, 18(1): 561- 170.

[4] 王尚平, 张亚玲, 王育民. 证书吊销的线索二叉排序 Hash 树解决方案[J]. 软件学报, 2001, 12(9): 1343- 1350.

[5] Hously R, Ford W, Polk W, et al. Internet X. 509 publickey infrastructure certificate and CRL profile[S]. IETF RFC2459, 1999, http://www.ietf.org/rfc/rfc2459.html.

[6] Andr ! mes, Mike Just, Svein J, et al. Selecting revocation solutions for PKI[A]. Proceedings of The Fifth Nordic Workshop on Secure IT Systems (NORDSEC 2000)[C]. Reykjavik, Iceland, 2000. 360- 376.

[7] Cooper A Cooper. A more efficient use of Delta2CRLs[A]. The Proceedings of the 2000 IEEE Symposium on Security and Privacy[C]. Berkeley, 2000. 190- 202.

[8] David A Cooper. A model of certificate revoction[A]. The Proceedings of Fifeenth Annual Comperter Security Application Conference[C]. Phoenix, 1999. 256- 264.

[9] 朱晶, 沈美明, 汪东升. Web 服务系统的性能分析与测试[J]. 计算机工程与应, 2001, 037(015): 9- 11.

作者简介:

谭 良 男, 1973 年出生于四川泸县, 博士研究生, 主要研究方向为网络计算, 信息安全, 软件工程. E2mail: tanliang2008cn@yahoo.com.cn.

刘 震 男, 1976 年出生于吉林市, 博士研究生, 主要研究方向为网络计算, 分布式计算, 信息安全.

佘 男, 1967 年出生于四川成都市, 副教授, 在职博士研究生, 主要研究方向为网络计算, 信息安全

周明天 男, 1939 年出生于广西容县, 教授, 博士生导师, 主要研究方向为网络计算, 信息安全, 并行分布计算, 移动计算.