基于平滑熵的无条件安全秘密钥的提取

杨 波, 张 彤, 王育民
(西安电子科技大学 ISN 国家重点实验室 106 信箱, 陕西西安 710071)

摘 要: 本文讨论了通信双方在公共信道上基于平滑熵进行无条件安全秘密钥协商协议来提取秘密钥时, 窃听者在信息协调阶段获得的边信息对保密增强阶段所能提取出的秘密钥长度的影响, 得出了以一定概率所协商的秘密钥长度.

关键词: 平滑熵; 密钥协商; 信息协调; 保密增强
中图分类号: TN918.1 文献标识码: A 文章编号: 0372-2112 (2001) 07-0930-03

Distillation of Unconditionally-Secure Secret-Key Based on Smooth Entropy

YANG Bo, ZHANG Tong, WANG Yu-min
(*National Key Laboratory on ISN of Xidian University, Xi'an, Shanxi* 710071, *China*)

Abstract: This paper investigates the effect of side-information, obtained by the opponent through an initial reconciliation step, on the size of the secret-key that can be distilled safely by subsequent privacy amplification in uncoditionally-secure secret-key agreement protocol based on smooth entropy, and obtains the size of the secret-key with some probability.

Key words: smooth entroy; secret-key agreement; information reconciliation; privacy amplification;

1 引言

目前密码体制的安全性大都是基于计算安全模型, 即基于某一数学问题的难解性的假定之上的. 随着计算能力的日益提高, 数学困难性问题的假定则受到越来越大的挑战, 对具有无限计算能力的敌手从理论上来说现有的密码体制都可通过对密钥空间的穷搜索而破译. 无条件安全的密码体制则不对敌手的计算能力做任何限制, 而且也不用任何弱的安全性假定, 因而具有重要的理论意义和实用价值.

无条件安全的秘密钥协商协议[1]借助于有噪信道使通信双方 Alice, Bob 及敌手 Eve 分别得到概率分布服从 P_{XYZ} 的三个随机变量 X, Y, Z, 即使 Eve 的信道优于 Alice 和 Bob 的信道, Alice 和 Bob 也可在一个具有认证功能的公共信道(Eve 只能窃听, 不能篡改信道上的消息)上进行协商产生出高度保密的秘密密钥. 协议由以下三步组成[2]:

• 优先提取: 不要求 Alice 和 Bob 的信道优于 Eve 的信道, 即不要求 $I(X;Y) > I(X;Z)$ 和 $I(Y;X) > I(Y;Z)$; Alice 和 Bob 在公共信道上交换一些信息(由随机变量 C 来表示), Alice 可由 X 和 C 求出 W, 而 Bob 关于 W 的不确定性小于 Eve 关于 W 的不确定性, 即 $H(W|XC) = 0$, $H(W|YC) < H(W|ZC)$.

• 信息协调: Alice 和 Bob 交换冗余信息, 利用纠错技术使 Bob 以很高的概率确定 W, 而 Eve 却不能完全确定 W. 设 Alice 和 Bob 发送比特串 U, U 的长度 L 略大于 $H(W|YC)$ 使得 $H(W|YCU) \approx 0$, 而 $H(W|ZCU) \geq H(W|ZC) - L > 0$.

• 保密增强: Alice 和 Bob 从 W 中提取出一个更短的位串 K, 而 Eve 关于 K 的信息可忽略地小. 例如 Alice 和 Bob 可公开协商一个 universal hash 函数[3] g (Eve 知道 W 的部分信息和 g 的全部信息), 以 $K = g(W)$ 作为秘密钥.

文献[2]利用二阶 Rényi 熵研究保密增强, 然而未考虑信道协调阶段 Eve 获得的边信息 U 对 Alice 和 Bob 所协商的秘密钥长度的影响, 文献[4, 5]对此进行了研究, 文献[6]利用平滑熵研究保密增强, 然而也未考虑信息协调阶段 Eve 获得的边信息 U 对 Alice 和 Bob 所协商的秘密钥长度的影响. 本文研究这一问题.

2 Rényi 熵、平滑熵及一些结果

定义 1[6] 设 X 为取值于集合 $\mathcal{X}$ 的一个随机变量(文中随机变量用大写字母表示, 取值集合用相应的手写体字母表示), 概率分布为 P_X, $\alpha \geq 0$ 且 $\alpha \neq 1$, X 的 α 阶 *Rényi* 熵定义为 $H_\alpha(X) = \frac{1}{1-\alpha}\log\sum_{x\in\mathcal{X}} P_X(x)^\alpha$, 其中对数以 2 为底, 下同.

$\alpha = 2$ 时为 $H_2(X) = -\log\sum_{x\in\mathcal{X}} P_X(x)^2$, 而 $H(X) = -\sum_{x\in\mathcal{X}} P_X(x)\log P_X(x)$ 可看作是 $H_\alpha(X)$ 当 $\alpha \to 1$ 时的极限情况.

引理 1[6] $H_\alpha(X)$ 关于 α 是单调递减的, 即对 $\forall\, \alpha, \beta$, $0 < \alpha < \beta$, $H_\alpha(X) \geq H_\beta(X)$, 当且仅当 X 是均匀分布时, 等号成立.

收稿日期: 2000-04-03; 修回日期: 2000-11-06
基金项目: 国家自然科学基金(No. 69972034); 国防科技重点实验室基金(No. 99JS06.3.1.DZ0112)

引理 2[6]　设 X 为取值于集合 $\mathcal{X}$ 的一个随机变量, 概率分布为 P_X, $P_{min}=\min_{x\in\mathcal{X}}P_X(x)$, $P_{max}=\max_{x\in\mathcal{X}}P_X(x)$, 且 $P_{max}\le c\cdot P_{min}$, 其中 $c>1$, 那么

$$\frac{1}{|\mathcal{X}|-1+c}\le P_{min}\le\frac{1}{|\mathcal{X}|},\ \frac{1}{|\mathcal{X}|}\le P_{max}\le\frac{C}{|\mathcal{X}|-1+c}.$$

引理 3[6]　设 X、P_{min}、P_{max} 定义同上, 满足 $P_{max}\le c\cdot P_{min}$, 其中 $c>1$, 则 $H_2(X)>H(X)-2\log c$.

定义 2[6]　设 X、Y 为取值于同一集合 $\mathcal{X}$ 的两个随机变量, 概率分布分别为 P_X、P_Y, X 与 Y 的相对熵定义为 $D(P_X\|P_Y)=\sum_{x\in\mathcal{X}}P_X(x)\log\frac{P_X(x)}{P_Y(x)}$.

如果 P_Y 是 X 上的均匀分布, 则 $D(P_X\|P_Y)=\log|\mathcal{X}|-H(X)$.

下面引入平滑熵的概念, 对随机变量 X, 函数 $f: X\to Y$ 使得 $Y=f(X)$ 在其取值集 y 上是足够均匀分布的, 则称 f 为平滑函数, $|y|$ 称为相对于完全均匀分布在允许的偏差范围内 X 的平滑熵. 用 $M(X)$ 来度量偏差, 常取为相对熵 $D(P_X\|P_U)$ 其中 $P_U(x)=1/|\mathcal{X}|$ 是 X 上的均匀分布. 正式定义如下:

定义 3[6]　设 M 是非均匀性度量, $\Delta: R\to R$ 是一递减的非负函数, X 是取值于集合 $\mathcal{X}$ 的随机变量, 称 X 以概率 $1-\varepsilon$ 在偏差范围 $\Delta(s)$ (关于 M) 内有平滑熵 $\Psi(X)$, 如果 $\Psi(X)$ 是满足以下条件的最大的 Ψ: 对任意安全参数 $s\ge 0$, 存在一随机变量 T 和一个函数 $f: \mathcal{X}\times\mathcal{T}\to y$, $|y|=\lfloor 2^{\Psi-s}\rfloor$, 使得有一概率至多为 ε 的失败事件 ε, 在已知 T 和 $\bar{\varepsilon}$ 时, $Y=f(X,T)$ 的非均匀性度量 M 在 T 上的均值至多为 $\Delta(s)$, 即

$$\Psi(X)=\max_{\Psi}\{\Psi\mid\forall s\ge 0:\exists\ T, f:\mathcal{X}\times\mathcal{T}\to y,\ |y|=\lfloor 2^{\Psi-s}\rfloor:$$
$$Y=f(X,T),\exists\ \varepsilon: P[\varepsilon]\le\varepsilon, M(Y|T\bar{\varepsilon})\le\Delta(s)\}$$

定理 1[2]　设 P_{VW} 是一个任意的概率分布, v 是 Eve 观测到的 V 的一个特定值, 若 Eve 关于 W 的二阶 Rényi 熵 $H_2(W|V=v)$ 至少为 c, 且 Alice 和 Bob 选择 $K=G(W)$ 作为其秘密钥, 其中 G 是从 $\mathcal{F}: \{0,1\}^l$ 的 Universal hash 函数族中随机选取的一个函数, 则

$$H(K|G,V=v)\ge H_2(K|G,V=v)\ge$$
$$l-\log(1+2^{l-c})\ge l-2^{l-c}/\ln 2.$$

令 $s=c-l$, 则当 $l<c$ 时, Eve 关于秘密钥 K 的熵值接近于最大, Alice 和 Bob 可获得最长的秘密钥, 其长度满足 $l=c-s\le H_2(W|V=v)-s$, 而 Eve 关于 K 的信息量 $\le 2^{l-s}/\ln 2$.

定理 2[6]　设 $1<\alpha<2$, r, $t>0$, m 是满足 $m-\log(m+1)>\log|\mathcal{X}|+t$ 的整数, s 是平滑熵的安全参数, 则随机变量 X 在误差范围 $2^{l-s}/\ln 2$ (以相对熵度量) 内的平滑熵 $\Psi(X)$ 以概率 $1-2^{-r}-2^{-t}$ 有以下关系: $\Psi(X)\ge H_\alpha(X)-\log(m+1)-\frac{r}{\alpha-1}-t-2$.

3　基于平滑熵所能提取出的无条件安全秘密钥的长度

定理 3　设 α, r, t, m, s 与定理 2 相同, v 是 Eve 观测到的 V 的一个特定值, Alice 和 Bob 选择 $K=G(W)$ 作为其秘密钥, 其中 G 是从 $\mathcal{F}: \{0,1\}^l$ 的 Universal hash 函数族中随机选取的一个函数, 则 K 的长度 l 以概率 $1-2^{-r}-2^{-t}$ 有以下关系:

$$l=H_\alpha(W|V=v)-\log(m+1)-\frac{r}{\alpha-1}-t-2-s.$$

且 Eve 关于 K 的信息量 $\le\frac{2^{-s}}{\ln 2}$.

证明　由平滑熵的定义, $|k|=\lfloor 2^{\Psi(W|V=v)-s}\rfloor\le 2^{\Psi(W|V=v)-s}$, $l=\log|\mathcal{K}|\le\Psi(W|V=v)-s$. 取 $l=H_\alpha(W|V=v)-\log(m+1)-\frac{r}{\alpha-1}-t-2-s$ 即满足平滑熵的定义 (以概率 $1-2^{-r}-2^{-t}$).

由定理 2 及相对熵的定义, $H(X)=\log|\mathcal{X}|-D(P_X\|P_U)\ge\log|\mathcal{X}|-2^{-s}/\ln 2$.

所以 $H(K|G,V=v)\ge\log|\mathcal{K}|-2^{-s}/\ln 2=l-2^{-s}/\ln 2$, 等价于 Eve 关于 K 的信息量 $\le 2^{-s}/\ln 2$.

与定理 1 比较, 定理 3 以一定的概率和定理 1 有相同的安全性, 但当 $m\to\infty$ 时, 可得 $\frac{1}{m}l=\frac{1}{m}H_\alpha(W|V=v)-\frac{1}{m}[\log(m+1)+\frac{r}{\alpha-1}+t+2+s]\to\frac{1}{m}H_\alpha(W|V=v)$.

$l\approx H_\alpha(W|V=v)\ge H_2(W|V=v)\ge l'+s$ (其中 l' 为定理 1 产生的秘密钥长度), 可见定理 3 的界好于定理 1 的界. 与定理 1 一样, 定理 3 也未考虑 Eve 在 Alice 和 Bob 进行信息协调阶段所获得的边信息 U 对秘密钥长度的影响. 下面首先考虑 $H_\alpha(W|V=v)-H_\alpha(W|V=v,U)$ 的上界, 然后考虑 Alice 和 Bob 最终所能得到的秘密钥长度.

定理 4　设 $1<\alpha<2$, X 为取值于集合 $\mathcal{X}$ 的一个随机变量, f 是 $\mathcal{X}\to\mathcal{U}$ 的任一函数, 令 $U=f(X)$, 取值于集合 $\mathcal{U}$, 则

$$H_\alpha(X)-H_\alpha(X|U=u)\le-\log P_U(u)+r/(\alpha-1)$$

至少以概率 $1-2^{-r}$ 成立, 其中 r 为大于 0 的任一常数.

证明　因为 $U=f(X)$, 所以 $H_\alpha(X)=H_\alpha(XU)=\frac{1}{1-\alpha}\log\sum_{x\in\mathcal{X},u\in U}P_{XU}(x,u)^\alpha=\frac{1}{1-\alpha}\log\sum_{u\in U}P_U(u)P_U(u)^{\alpha-1}\sum_{x\in\mathcal{X}}P_{X|U=u}(x)^\alpha=\frac{1}{1-\alpha}\log\sum_{u\in U}P_U(u)2^{(\alpha-1)\log P_U(u)+(1-\alpha)H_\alpha(X|U=u)}$

将 $H_\alpha(X|U=u)$ 看作 u 的函数, 上式变为

$$\sum_{u\in U}P_U(u)2^{(\alpha-1)\log P_U(u)+(1-\alpha)H_\alpha(X|U=u)}=2^{(1-\alpha)H_\alpha(X)}$$

即 $E_U[2^{(\alpha-1)\log P_U(u)+(1-\alpha)H_\alpha(X|U=u)}]=2^{(1-\alpha)H_\alpha(X)}$,

$$E_U[2^{(\alpha-1)\log P_U(u)+(1-\alpha)H_\alpha(X|U=u)-(1-\alpha)H_\alpha(X)}]=1,$$

取 r 为大于 0 的任一常数, 则

$$E_U[2^{(\alpha-1)\log P_U(u)+(1-\alpha)H_\alpha(X|U=u)-(1-\alpha)H_\alpha(X)-r}]=2^{-r}$$

由不等式 $P[X\ge r]\le E[e^{(X-r)t}]$, 其中 $t\in R^+$, $r\in R$, 取 $t=\ln 2$, 则 $P[X\ge r]\le E[2^{(X-r)}]$

得 $P_U[(\alpha-1)\log P_U(u)+(1-\alpha)H_\alpha(X|U=u)-(1-\alpha)\cdot H_\alpha(X)\ge r]\le 2^{-r}$

即 $(\alpha-1)\log P_U(u)+(1-\alpha)H_\alpha(X|U=u)-(1-\alpha)H_\alpha(X)\le r$ 至少以概率 $1-2^{-r}$ 成立, 两边同除以 $(\alpha-1)$ 得 $H_\alpha(X)-H_\alpha(X|U=u)\le-\log P_U(u)+r/(\alpha-1)$.

若已知边信息 U 的 α 阶 Rényi 熵 $H_\alpha(U)$, 则可得以下定理,

定理 5　设 $1<\alpha<2$, X, U, x, u 的定义同定理 4, $P_{min}=$

$\min_{u\in\mathcal{U}} P_U(u)$, $P_{\max}=\max_{u\in\mathcal{U}} P_U(u)$.

且 $P_{\max}\le c\cdot P_{\min}$, 其中 $1<c\le\sqrt[3]{2}$, 则当 $1\le\beta\le 1/\sqrt{c^3-1}$ 时,

$$H_\alpha(X)-H_\alpha(X\mid U=u)\le H_\alpha(U)-\log(1-\beta\sqrt{c^3-1})+r/(\alpha-1)$$

至少以概率 $1-2^{-r}-\beta^{-2}$ 成立.

证明 由 $H_2(U)=-\log\sum_{u\in\mathcal{U}}P_U(u)^2=-\log E[P_U(u)]$ 及 $H_2(U)\le H_\alpha(U)$ 得

$$E[P_U(u)]=2^{-H_2(U)}\ge 2^{-H_\alpha(U)}$$

$$H_3(U)=-\frac{1}{2}\log\sum_{u\in\mathcal{U}}P_U(u)^3$$

$$H(U)-H_3(U)\le\log|\mathcal{U}|+\frac{1}{2}\log\sum_{u\in\mathcal{U}}P_U(\omega)^3\le\log|\mathcal{U}|+\frac{1}{2}\log(|\mathcal{U}|P_{\max}^3)\le\frac{3}{2}\log|\mathcal{U}|+\frac{3}{2}\log P_{\max}=\frac{3}{2}\log(|\mathcal{U}|P_{\max})\le\frac{3}{2}\log\left[|\mathcal{U}|\frac{c}{|\mathcal{U}|-1+c}\right]\le\frac{3}{2}\log c$$

(第 4 个不等式由引理 2 得出).

$$H_3(U)\ge H(U)-\frac{3}{2}\log c\ge H_\alpha(U)-\frac{3}{2}\log c.$$

所以 $P_U(u)$ 的方差

$$\sigma^2_{P_U}=E[P_U(u)^2]-E^2[P_U(u)]=\sum_{u\in\mathcal{U}}P^3_U(u)-\left(\sum_{u\in\mathcal{U}}P^2_U(u)\right)^2=2^{-2H_3(U)}-2^{-2H_2(U)}$$

$$\le 2^{-2H_\alpha(U)+3\log c}-2^{-2H_\alpha(U)}\le 2^{-2H_\alpha(U)}(c^3-1).$$

因为 $1<c\le\sqrt[3]{2}$, 所以 $1\le 1/\sqrt{c^3-1}$, 当 $1\le\beta\le 1/\sqrt{c^3-1}$ 时, 由契比雪夫不等式可得

$$P\{|P_U(u)-\mu_{P_U}|<\beta\sigma_{P_U}\}\ge 1-\beta^{-2}$$

由 $|P_U(u)-\mu_{P_U}|<\beta\sigma_{P_U}$ 得

$$P_U(u)\ge\mu_{P_U}-\beta\sigma_{P_U}\ge 2^{-H_\alpha(U)}-\beta 2^{-H_\alpha(U)}\sqrt{c^3-1}=2^{-H_\alpha(U)}(1-\beta\sqrt{c^3-1})$$

所以 $-\log P_U(u)\le H_\alpha(U)-\log(1-\beta\sqrt{c^3-1})$ 至少以概率 $1-\beta^{-2}$ 成立.

结合定理 4 得 $H_\alpha(X)-H_\alpha(X\mid U=u)\le H_\alpha(U)-\log(1-\beta\sqrt{c^3-1})+r/(\alpha-1)$

至少以概率 $(1-2^{-r})(1-\beta^{-2})>1-2^{-r}-\beta^{-2}$ 成立.

在定理 5 中, 以 $P(W\mid V=v)$ 代替 $P(X)$, 并令 $\tau=H_\alpha(U)-\log(1-\beta\sqrt{c^3-1})+r/(\alpha-1)$, 则得, $H_\alpha(W\mid V=v)-H_\alpha(W\mid V=v,U=u)\le\tau$, $H_\alpha(W\mid V=v,U=u)\ge H_\alpha(W\mid V=v)-\tau$, 重新利用定理 3 则得 $l=H_\alpha(W\mid V=v,U=u)-\log(m+1)-r/(\alpha-1)-t-2-s\ge H_\alpha(W\mid V=v)-\log(m+1)-r/(\alpha-1)-t-2-s-\tau$.

且 $H(K\mid G,V=v,U=u)\ge l-2^{-s}/\ln 2$ 等价于 Eve 关于 K 的信息量 $\le 2^{-s}/\ln 2$.

上两式成立的概率至少为 $(1-2^{-r}-2^{-t})(1-2^{-r}-\beta^{-2})$.

综合以上讨论, 得本文的主要结果如下:

定理 6 Alice 和 Bob 进行无条件安全的秘密钥协商协议, 在优先提取后, 设 Eve 所知道的关于 W 的信息由 $V=v$ 表示; 在信息协调阶段, Eve 获得的边信息由 $U=u$ 表示; 若已知 $H_\alpha(U)$ $(1<\alpha<2)$ 及 U 的概率分布满足的限制条件 $\max_{u\in U}P_U(u)\le c\cdot\max_{u\in U}P_U(u)$ $(1<c<\sqrt[3]{2})$, 则 Alice 和 Bob 在保密增强阶段能够以至少 $(1-2^{-r}-2^{-t})(1-2^{-r}-\beta^{-2})$ 的概率提取长为

$$H_\alpha(W\mid V=v)-\log(m+1)-r/(\alpha-1)-t-2-s-\tau$$

比特的无条件安全秘密钥 K, 且 Eve 所知道的关于 K 的信息量不大于 $2^{-s}/\ln 2$.

其中 $\tau=H_\alpha(U)-\log(1-\beta\sqrt{c^3-1})+r/(\alpha-1)$, $r>0$, $t>0$, $1\le\beta\le 1/\sqrt{c^3-1}$), m 满足 $m-\log(m+1)>\log|\mathcal{X}|+t$, $s\ge 0$ 是安全参数.

4 结论

在基于平滑熵进行无条件安全的秘密钥协商协议中, 窃听者在通信双方进行信息协调阶段所获得的边信息必将对通信双方最终所提取的秘密钥长度产生影响. 若已知边信息的 α 阶 Rényi 熵及边信息概率分布的一定限制条件, 就可以以一定的概率得出通信双方所能提取的秘密钥长度.

参考文献:

[1] U M Maurer. Secret key agreement by public discussion from common information [J]. IEEE Trans, Inform. Theory, 1993, 39(3): 733- 742.

[2] Charles H Bennett, et al. Generalized privacy amplification [J]. IEEE Trans. Theory, 1995, 41(6): 1951- 1923.

[3] J L Carter, M N Wegman. Universal classes of hash functions [J]. J. Comput. Syst. Sci., 1979, 18(2): 143- 154.

[4] Christian Cachin, Ueli Maurer. Linking information reconciliation and privacy amplification [A]. EUROCRYPT' 94 [C], Lecture Notes in Computer Science, Springer-Verlag, 1995, 266- 274.

[5] 刘胜利, 王育民. 无条件安全密钥的提取[J]. 电子学报, 1999, 27(10): 128- 130.

[6] Christian Cachin. Smooth entropy and Rényi entropy [A]. EUROCRYPT' 97[C], Lecture Notes in Computer Science, Springer-Verlag, 1997, 193- 208.

作者简介:

杨 波 男. 1963 年 5 月生于陕西. 西安电子科技大学副教授、密码学博士, 中国电子学会高级会员, 信息产业部信息安全技术专家组成员, 陕西省信息及网络安全保密专家顾问小组成员. 研究方向为信息论、电子商务中的安全理论与技术.

张 彤 男. 1967 年 6 月生于陕西. 副研究员, 西安电子科技大学博士生, 研究方向为信息论、信息隐匿理论与技术.

王育民 男. 生于 1936 年, 教授、博士生导师、中国电子学会会士, 长期从事信息论、信道编码、密码学以及通信网的安全等方面的研究.