DOI: 10.13718/j.cnki.xdzk.2017.01.001

一个生成两个相关随机比特串的方法①

包小敏[1]，　瞿云云[2]

1. 西南大学 数学与统计学院，重庆 400715；2. 贵州师范大学 数学科学学院，贵阳 550001

摘要：给出了一个方案，可以将量子密钥分配协议 BB84 的效率提高近 17%. 在这个方案中，通信的双方共享一个概率串，其元素都是从区间(0，1)中随机均匀地选取的. 在传送和接收一个量子比特的过程中，概率串的相应元素被双方作为各自从两个基中选择一个基的概率，之后利用一个细化的数据分析方案进行随后的数据分析.

关　键　词：量子密码；量子密钥分配；BB84；相关系数；概率串

中图分类号：TN918.1　　**文献标志码**：A　　**文章编号**：1673-9868(2017)01-0001-08

信息的保密一般是通过对信息进行加密来实现的，而按照 Kerckhoffs 原则，一个密码体制的安全应该基于密钥的保密而不是对算法本身的保密. 正因为如此，密钥分配成为密码学中非常重要的研究课题. 基于公钥算法的密钥分配方案的安全性大多是基于求解困难性并没有得到证明的数学难题上的，如 Diffie-Hellman 方案就是基于离散对数问题是难解的假设上的. 随着算法的改进和硬件的升级，特别是量子计算机的诞生，这些困难问题很可能变得容易[1-2]. 而量子密钥分配方案的安全性则由量子力学原理所保证，因此其安全性不受算法和硬件的影响. 目前最著名的量子密钥分配方案(也是史上第一个量子密钥分配方案)是 1984 年由 Bennett 和 Brassard 发明的 BB84[3]. BB84 在进行密钥分配过程中要用到两个信道：一个量子信道和一个可认证公开信道. 通信的双方 Alice 和 Bob 通过量子信道在他们之间传输量子信息，为建立一个共享的对称密钥做准备. Alice 和 Bob 利用公共信道对通过量子信道得到的信息进行处理，以便能得到共享的对称密钥，然后在公共信道上传输用共享对称密钥加密的数据. 可认证公开信道是假设任何人都能监听此信道，但不能改变其中的信息. BB84 密钥分配过程由下面的 3 个步骤构成：

1) 产生原始密钥(raw key generation)：通信的双方通过量子信道产生两个等长的比特串；

2) 信息调和(information reconciliation)：通信的双方通过可认证公开信道交换信息，对两个等长、但不一定相同的比特串进行纠(滤)错，使之变成相同的比特串；

3) 保密增强(privacy amplification)：通信的双方通过可认证公开信道交换信息，消除在进行前面两个步骤的过程中第三方可能获得的信息，以提高比特串的保密性.

量子密钥分配的一个重要特征是能探测到窃听者的存在，而要取得这一特征，在通过量子信道产生原始密钥时双方基串的选择要有一定的随机性，但这又会影响整个方案的效率. 本文给出一个基串的生成方法，在保证随机性的同时，又可以提高效率.

---

① 收稿日期：2016-04-05

基金项目：国家自然科学基金项目(61462016)；重庆市自然科学基金计划资助项目(CSTC2006BB2325)；贵州省科学技术基金项目(黔科合 J 字[2014]2125 号)；贵州省教育厅青年科技人才成长项目(黔教合 KY 字[2016]130)；贵州师范大学博士启动项目(0514021).

作者简介：包小敏(1959-)，新疆巴楚人，博士，教授，主要从事密码学，纠错码理论的研究.

通信作者：瞿云云，副教授.

1 相关概念和研究状况

先介绍一些概念和结论. 由于篇幅所限，只介绍与本文密切相关的. 若想了解更多，请参考文献[4－5]的相关章节. 在量子信息论里与经典信息论中比特(bit)对应的是量子比特(qubit). 比特只有两个状态：0和1，而量子比特的状态可以用2维Hilbert空间$\mathscr{H}$中的单位矢量表示. 设$|0\rangle$，$|1\rangle$是$\mathscr{H}$的一组标准正交基，记

$$|+\rangle=\frac{|0\rangle+|1\rangle}{\sqrt{2}}$$

$$|-\rangle=\frac{|0\rangle-|1\rangle}{\sqrt{2}}$$

则$|+\rangle$，$|-\rangle$也是$\mathscr{H}$的一组标准正交基，这两组基分别称为Z基和X基.

对于状态$\alpha=a|0\rangle+b|1\rangle$，当在$Z$基下测量(measure in basis Z)时，只能得到$|0\rangle$或$|1\rangle$，其中得到$|0\rangle$的概率为a^2，得到$|1\rangle$的概率为b^2. 特别地，在Z基下测量$|0\rangle(|1\rangle)$，则得到的一定是$|0\rangle(|1\rangle)$；在Z基下测量$|+\rangle$或$|-\rangle$，则得到$|0\rangle$或$|1\rangle$的概率均是$\frac{1}{2}$. 同样的，对于状态$\beta=c|+\rangle+d|-\rangle$，当在$X$基下测量时，则只能得到$|+\rangle$或$|-\rangle$，其中得到$|+\rangle$的概率为$c^2$，得到$|-\rangle$的概率为$d^2$. 特别地，在$X$基下测量$|+\rangle(|-\rangle)$，则得到的一定是$|+\rangle(|-\rangle)$；在$X$基下测量$|0\rangle$或$|1\rangle$，则得到$|+\rangle$或$|-\rangle$的概率均是$\frac{1}{2}$.

下面总假设n，m是正整数. 用$(a_1,\cdots,a_l)_l$表示长度为l的串(string)，在不引起混淆的情况下简记为$(a_1,\cdots,a_l)$.

定义1 设l是正整数，$\boldsymbol{a}_i=(a_{i1},\cdots,a_{im})$是长度为$m$的比特串，$1\leqslant i\leqslant l$. 数

$$\frac{|\{j\mid a_{1j}=a_{2j}=\cdots=a_{lj},\ 1\leqslant j\leqslant m\}|}{m}$$

称为$\boldsymbol{a}_1,\cdots,\boldsymbol{a}_l$的相关系数，记为Correlation$(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_l)$. 对于$\beta\in\{0,1\}$，数

$$\frac{|\{j\mid a_{1j}=a_{2j}=\cdots=a_{lj}=\beta,\ 1\leqslant j\leqslant m\}|}{m}$$

称为$\boldsymbol{a}_1,\cdots,\boldsymbol{a}_l$的$\beta$相关系数，记为Correlation$(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_l)_\beta$. 显然有

$$\text{Correlation}(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_l)=\text{Correlation}(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_l)_0+\text{Correlation}(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_l)_1$$

设$\boldsymbol{a}=(a_1,\cdots,a_m)$和$\boldsymbol{b}=(b_1,\cdots,b_m)$是两个比特串. 显然

$$1-\text{Correlation}(\boldsymbol{a},\boldsymbol{b})=\frac{|\{i\mid a_i\neq b_i,\ 1\leqslant i\leqslant m\}|}{m}$$

此数称为$\boldsymbol{a}$与$\boldsymbol{b}$的错码率.

下面介绍BB84协议[4].

1) Alice随机地选择$N=(4+\delta)n$个比特做成一个数据比特串$\boldsymbol{a}=(a_1,a_2,\cdots,a_N)$.

2) Alice再随机选择一个有N个比特的比特串$\boldsymbol{b}=(b_1,b_2,\cdots,b_N)$. 对于$1\leqslant i\leqslant N$，如果$b_i=0$，那么Alice把$a_i$编码为$\{|0\rangle,|1\rangle\}$；否则编码为$\{|+\rangle,|-\rangle\}$. 也就是把$\boldsymbol{a}$编码为具有$N$个量子比特的一个块：

$$|\psi\rangle=\bigotimes_{i=1}^{N}|\psi_{a_ib_i}\rangle$$

其中$|\psi_{00}\rangle=|0\rangle$，$|\psi_{10}\rangle=|1\rangle$，$|\psi_{01}\rangle=|+\rangle$，$|\psi_{11}\rangle=|-\rangle$.

3) Alice通过量子信道把$|\psi\rangle$发送给Bob.

4) Bob随机地选择一个比特串$\boldsymbol{b}'=(b'_1,b'_2,\cdots,b'_N)$. 对通过量子信道接收到的第$i$个量子比特，若$b'_i=0$，则在$Z$下测量，否则在$X$下测量. Bob将测量的结果按照

$$|\psi_{00}\rangle \longrightarrow 0,\ \&\ |\psi_{01}\rangle \longrightarrow 0$$
$$|\psi_{10}\rangle \longrightarrow 1,\ \&\ |\psi_{11}\rangle \longrightarrow 1$$

的编码规则逐个、顺序地转化成一个数据比特串 $\boldsymbol{a}'=(a'_1, a'_2, \cdots, a'_N)$.

5) 在公开信道上 Alice 公布 $\boldsymbol{b}$, Bob 公布 $\boldsymbol{b}'$.

6) Alice 和 Bob 分别保留 $\boldsymbol{a}$ 和 $\boldsymbol{a}'$ 中位置对应于 $\boldsymbol{b}-\boldsymbol{b}'$ 中分量为 0 的位置的比特, 即当 $b_i-b'_i=0$ 时, Alice 保留 a_i, Bob 保留 a'_i; 否则 Alice 删除 a_i, Bob 删除 a'_i. 这样 Alice 和 Bob 将分别得到一个长度约为 $2n$ 的比特串 $\bar{\boldsymbol{a}}$ 和 $\bar{\boldsymbol{a}}'$. $\bar{\boldsymbol{a}}$ 和 $\bar{\boldsymbol{a}}'$ 称为筛后比特串.

7) Alice 和 Bob 协商在 $2n$ 个位置中随机地选择 n 个位置, 然后通过公开信道公布各自这 n 个位置上的比特值.

8) 若双方这 n 个位置上对应比特串之间的错码率 e 超过可接受的限度 $e_{\max}$, 则终止协议.

9) Alice 和 Bob 在各自剩下的约 n 个比特(构成的两个比特串称为原始密钥) 上通过公开信道进行信息调和与保密增强以获得 m 个比特的共享密钥.

由于比特串 $\boldsymbol{b}$ 和 $\boldsymbol{b}'$ 分别确定 Alice 的编码基和 Bob 的测量基, 所以也将这两个比特串称为基串.

量子密钥分配与传统密钥分配的一个重要区别是前者能探测到窃听者的存在. BB84 能取得这一特性的关键是在步骤 2) 和步骤 4) 中 Alice 和 Bob 分别独立随机地选择基串 $\boldsymbol{b}$ 和 $\boldsymbol{b}'$. 在无窃听、无噪音的情况下, 在步骤 8) 的比较中, 对应比特值不同的个数为 0, 也就是说两个被比较的串的相关系数为 1. 如果有窃听, 因为 Alice 和 Bob 是独立地选择 $\boldsymbol{b}$ 和 $\boldsymbol{b}'$, Eve 选择的基只有 $\frac{1}{2}$ 的概率与 Alice 和 Bob 的选择相同. 在不相同的情况下, Bob 接收的量子比特将有 $\frac{1}{2}$ 的概率与 Alice 发送的量子比特不同, 因此当 Eve 以 λ 的概率窃听每一个量子比特时, 通过步骤 8) 的比较, Alice 与 Bob 对应比特值不同的出现概率将可能达到 $\frac{\lambda}{4}$. 如果 λ 较大时, 就有可能在步骤 8) 被检测出.

注意, 在步骤 8) 中双方各自参与对比的 n 个比特中, 两个对应比特之间, Alice 发送所选择的基与 Bob 测量所选择的基是一致的, 既可能是 Z 基, 也可能是 X 基. 在对比中, 双方没有把 n 个比特按基分成两个串分别比较, 而是组成一个串来比较. 称这种不区分基的对比检验为统一对比检验.

为了说明方便, 我们从操作的角度将上面 BB84 协议的步骤 2) 和步骤 4) 做一些形式上的变化, 为此要用到(伪) 随机数生成器, 本文考虑的(伪) 随机数生成器产生的都是区间 $[0, 1]$ 上的实数, 且具有均匀分布. 下文中产生一个数是指调用随机数生成器一次让其输出一个数.

定义 2 设 $\boldsymbol{p}=(p_1, \cdots, p_m)$ 是一实数串, 其中 $0<p_i<1 (i=1, \cdots, m)$, 称为概率串. 给定 $\boldsymbol{p}$, 可按下面方法产生一个比特串 $\boldsymbol{r}=(r_1, \cdots, r_m)$: 对整数 $i (1\leqslant i\leqslant m)$, 随机产生一个实数 x_i, $0\leqslant x_i\leqslant 1$, 令

$$r_i=\begin{cases}1 & \text{若 } x_i\leqslant p_i\\ 0 & \text{其它}\end{cases}$$

r_i 称为依照 p_i 产生, 而 $\boldsymbol{r}$ 称为依照 $\boldsymbol{p}$ 产生.

由于 $x_1, \cdots, x_m$ 是随机产生的, 所以依照 $\boldsymbol{p}$ 产生的比特串不是唯一的. 不难证明下面这个定理.

定理 1 若 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 都是依据 $\boldsymbol{p}=(p, p, \cdots, p)_m$ 产生的, 其中 $0<p\leqslant\frac{1}{2}$, 则 Correlation$(\boldsymbol{a}, \boldsymbol{b})$ 的期望值为 $1-2p(1-p)$.

由定理 1 不难看出, BB84 协议中步骤 2) 和步骤 4) 中的 $\boldsymbol{b}$ 和 $\boldsymbol{b}'$ 都是依据 $\boldsymbol{p}=\left(\frac{1}{2}, \frac{1}{2}, \cdots, \frac{1}{2}\right)_N$ 产生的. 前面已经说明, 由于 Alice 和 Bob 是各自独立地依据 $\boldsymbol{p}=\left(\frac{1}{2}, \frac{1}{2}, \cdots, \frac{1}{2}\right)_N$ 分别产生基串 $\boldsymbol{b}$ 和基串 $\boldsymbol{b}'$,

因而通过步骤8)可以检测出窃听;但也正是因为双方独立地依据 $\boldsymbol{p}=\left(\frac{1}{2},\frac{1}{2},\cdots,\frac{1}{2}\right)_N$ 产生各自的基串,所以Bob接收到的数据比特串与Alice的数据比特串的相关系数在没有窃听的情况下只有$\frac{1}{2}$,因此在量子信道,BB84的效率$\left(\frac{\text{筛后比特串的长度}}{N}\approx 0.5\right)$只有约50%.

为了提高BB84的效率,文献[6]给出了一个方法,该方法产生 $\boldsymbol{b}$ 和 $\boldsymbol{b}'$ 时所依据的概率串为 $\boldsymbol{p}=(p,p,\cdots,p)_N$,其中 $0<p\leqslant\frac{1}{2}$. 此时,$\boldsymbol{b}$ 和 $\boldsymbol{b}'$ 中分量为0的比例约为 $1-p$,分量为1的比例约为 p. 这就是说Alice在步骤2)选择 Z 基编码,Bob在步骤4)选择 Z 基测量的比例约为 $1-p$,而双方选择 X 基的比例约为 p. 由定理1,当 $p\to 0$ 时,Correlation($\boldsymbol{b}$,$\boldsymbol{b}'$)的期望值趋向于1,从而Correlation($\bar{\boldsymbol{a}}$,$\bar{\boldsymbol{a}}'$)的期望值也趋向于1,因此量子信道的效率趋向于100%.

此时在步骤8)若采用统一对比检验,很可能无法检测出某些窃听. 文献[6]给出了一个窃听的方案,此方案是一个中间攻击方案(man in the middle attack). 对每个Alice发送给Bob的qubit,Eve按下面的策略操作:

1)以概率 p_Z 选择 Z 基测量,然后将测量结果重发给Bob;

2)以概率 p_X 选择 X 基测量,然后将测量结果重发给Bob;

3)以概率 $1-p_Z-p_X$ 什么都不做.

用 e_Z 表示Alice和Bob都选择 Z 基时的错码率,用 e_X 表示Alice和Bob都选择 X 基时的错码率. 由于在这两种情况下,Alice和Bob选择的基都相同,所以出错只能是由Eve选择了不同的基测量导致的,故

$$e_Z=\frac{p_X}{2}\qquad e_X=\frac{p_Z}{2}$$

若采用统一对比检验,Alice和bob只能发现平均错码率:

$$\bar{e}=\frac{p^2e_X+(1-p)^2e_Z}{p^2+(1-p)^2}=\frac{p^2p_Z+(1-p)^2p_X}{2[p^2+(1-p)^2]}$$

若Eve只选择 Z 基测量,即 $p_Z=1$,$p_X=0$,则有

$$\bar{e}=\frac{p^2}{2[p^2+(1-p)^2]}$$

当 $p\to 0$ 时,$\bar{e}\to 0$,因此采用统一对比检验,Alice和bob将无法检测出Eve的窃听.

这种攻击能够成功是利用了Alice和Bob偏向于选择 Z 基,在窃听时也偏向于选择 Z 基窃听,因此Alice和Bob在 Z 基中的对应比特的错码率小于其在 X 基中的对应比特的错码率. 由于统一对比检验得出的是两种错码率的平均值,尽管一个错码率可能很高,但由于另一个错码率较低,拉低了平均值,故而统一对比检测无法测出Eve的窃听.

为抵抗这种攻击,文献[6]还给出了一个较统一对比检验细化的方法:在步骤7),Alice和Bob协商在 $2n$ 个位置中分别随机地选择 m_1 个双方选择 Z 基的位置和 m_2 个双方选择 X 基的位置,这里 $m_1\approx\frac{n}{2}\approx m_2$,然后通过公开信道公布各自这 m_1+m_2 个位置上的比特值. 在步骤8)将各自 m_1 个采用 Z 基的比特组成一个子串,采用 X 基的组成另一个子串,然后分别计算这两个子串与对方对应子串的错码率 e_Z 和 e_X. 在步骤8),只有当 $e_Z,e_X<e_{\max}$ 时,才进行步骤9). 称这种对比检验为分类对比检验. 文献[6]中,步骤9)选择的 n 个比特值都来自发送和接收都选择 Z 基的比特值(因为选择 Z 基的比特数在 $p<\frac{1}{2}$ 时要多于选择 X 基的比特数).

文献[6]证明了在BB84中依据 $\boldsymbol{p}=(p,p,\cdots,p)_N$ 来产生 $\boldsymbol{b}$ 和 $\boldsymbol{b}'$,然后用分类对比检验来替换统一对比检验,这样改进后的方案是安全的. 为方便起见,称改进后的方案为LCA-BB84方案.

LCA-BB84 方案虽然能提高 BB84 产生原始比特的效率，但这个效率更多体现在量子信道上. BB84 之所以要在步骤 7) 从 $2n$ 个比特中选取 n 个来进行对比检验，是为了使得到的检验结果具有较高的精确度，即能够使检验得到的错码率尽可能接近双方剩余的 n 个比特串的实际错码率. 从统计学的角度来说，参与检验的比特数越大，得到的结果的精确度也就越高. 文献[6]的作者也认为如果要求 Eve 攻破系统的概率是剩余比特长度 n 的指数函数的倒数(因为量子密钥分配要达到的安全等级是无条件安全，所以这个攻破系统的概率要求是合理的)，那选择参与检测的比特数与 n 同级是必须的.

为了在步骤 7) 有 $2n$ 个比特，BB84 通过量子信道要发送约 $4n = N$ 个 qubit. 若 LCA-BB84 方案在取 $p = \frac{1}{3}$ 时针对两个基的错码率也要达到和 BB84 相同的精确度，也就是对应于每个基的比特串的长度都是 n，加上剩余的 n 个比特，则至少要在步骤 6) 得到 $3n$ 个比特，因此通过量子信道发送的 qubit 至少要 $\frac{3n}{1-2p(1-p)} = \frac{27n}{5}$ 个，大于 BB84 的 $4n$. 如果只要求总的参与比较的比特数是 n，即 $m_1 = m_2 = \frac{n}{2}$. 由于 $m_2 \leqslant Np^2$，因此通过量子信道发送的 qubit 至少要 $\frac{m_2}{p^2} = \frac{9n}{2} = 4.5n$，也大于 BB84 的 $4n$. 从另一个角度来看，若要求发送 qubit 的总数不超过 $4n$，即要保证 $N \leqslant 4n$，则 $m_2 \leqslant Np^2 \leqslant 4np^2$，故 $p \geqslant \sqrt{\frac{1}{8}} > \frac{1}{3}$.

从上面的分析看出，文献[6]声称的 LCA-BB84 的效率高于 BB84，主要体现在量子信道上. 从总体来看，LCA-BB84 效率的提高是有限的. 在不增加传送的 qubit 的总数的情况下，产生基串时依据的概率串 $\boldsymbol{p} = (p, p, \cdots, p)_N$ 中的 $p > \frac{1}{3}$，故由定理 1 知其效率小于 $\frac{5}{9}$，并不能无限接近 100%. 注意，这里还没有考虑 LCA-BB84 方案在步骤 9) 进行信息调和时的比特串全部来自 Z 基的比特. 导致 LCA-BB84 方案在量子信道效率高，但经过检错后效率回落的原因是：要提高量子信道的效率，就要减少概率串 $\boldsymbol{p} = (p, p, \cdots, p)_N$ 中 p 的值. 而 p 值的减少，就导致步骤 6) 中剩余的比特里 X 基的比特较少. 要保证有足够数量的 X 基比特来进行错码率对比，要么加大 qubit 的发送量，要么 p 的值不能太小.

本文给出一种方案，在方案中 Alice 和 Bob 都采用一种特殊的方法来产生基串，然后利用分类对比检验，可以使效率达到 $\frac{2}{3}$. 这个方案的效率不仅优于 BB84 方案，也优于 LCA-BB84 方案.

3 新方案 —PARK

在 BB84 方案中的概率串为 $(\frac{1}{2}, \cdots, \frac{1}{2})_N$；而在 LCA-BB84 方案中，概率串为 $(p, \cdots, p)_N$，其中 $0 < p \leqslant \frac{1}{2}$. 一个自然的考虑是将概率串换成 $(p_1, \cdots, p_N)$ 后会有什么结果. 为此先证明下面的定理.

定理 2 设概率串 $\boldsymbol{p} = (p_1, \cdots, p_m)$ 的各项都是相互独立随机产生的. 若 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 都是依据 $\boldsymbol{p}$ 产生的，则当 $m \to \infty$ 时，Correlation($\boldsymbol{a}$, $\boldsymbol{b}$) 的期望值为 $\frac{2}{3}$.

证 考察 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 在位置 i 的对应元素 a_i 和 b_i. 容易看出 $a_i = b_i$ 当且仅当 $a_i = 1 = b_i$ 或 $a_i = 0 = b_i$. 但

$$\Pr(a_i = 1 = b_i) = p_i^2,\ \Pr(a_i = 0 = b_i) = (1 - p_i)^2$$

所以

$$\Pr(a_i = b_i) = p_i^2 + (1 - p_i)^2$$

因此 Correlation($\boldsymbol{a}$, $\boldsymbol{b}$) 的期望值为

$$E(\text{Correlation}(\boldsymbol{a}, \boldsymbol{b})) = \int_0^1 [x^2 + (1-x)^2]\mathrm{d}x = \frac{2}{3}$$

从这个定理可以看出，若概率串的各项都是独立随机产生的，则双方依据此概率串产生的基串的相关系数就为$\frac{2}{3}$，因此相应的量子信道的效率就是$\frac{2}{3}$，高于BB84.

下面给出我们的方案. 首先假设Alice和Bob共享一个概率串$\boldsymbol{p}=(p_1, \cdots, p_N)$，其各项是互相独立随机产生的.

1) Alice随机地选择$N=(3+\delta)n$个比特做成一个数据比特串$\boldsymbol{a}=(a_1, a_2, \cdots, a_N)$.

2) Alice依据$\boldsymbol{p}$产生基串$\boldsymbol{b}=(b_1, b_2, \cdots, b_N)$. 对于$1 \leqslant i \leqslant N$，如果$b_i=0$，那么Alice把$a_i$编码为$\{|0\rangle, |1\rangle\}$；否则编码为$\{|+\rangle, |-\rangle\}$. 也就是把$\boldsymbol{a}$编码为具有$N$个量子比特的一个块：

$$|\psi\rangle=\bigotimes_{i=1}^{N}|\psi_{a_i b_i}\rangle$$

其中$|\psi_{00}\rangle=|0\rangle$，$|\psi_{10}\rangle=|1\rangle$，$|\psi_{01}\rangle=|+\rangle$，$|\psi_{11}\rangle=|-\rangle$.

3) Alice通过量子信道把$|\psi\rangle$发送给Bob.

4) Bob依据$\boldsymbol{p}$产生基串$\boldsymbol{b}'=(b'_1, b'_2, \cdots, b'_N)$. 对通过量子信道接收到的第$i$个量子比特，若$b'_i=0$，则在$Z$下测量，否则在$X$下测量. Bob将测量的结果按照

$$|\psi_{00}\rangle \longrightarrow 0 \qquad |\psi_{01}\rangle \longrightarrow 0$$
$$|\psi_{10}\rangle \longrightarrow 1 \qquad |\psi_{11}\rangle \longrightarrow 1$$

的编码规则逐个、顺序地转化成一个数据比特串$\boldsymbol{a}'=(a'_1, a'_2, \cdots, a'_N)$.

5) 在公开信道上Alice公布$\boldsymbol{b}$，Bob公布$\boldsymbol{b}'$.

6) Alice和Bob分别保留$\boldsymbol{a}$和$\boldsymbol{a}'$中位置对应于$\boldsymbol{b}-\boldsymbol{b}'$中分量为0的位置的比特，即当$b_i-b'_i=0$时，Alice保留$a_i$，Bob保留$a'_i$；否则Alice删除$a_i$，Bob删除$a'_i$. 这样Alice和Bob将分别得到一个长度约为$2n$的比特串$\overline{\boldsymbol{a}}$和$\overline{\boldsymbol{a}}'$. $\overline{\boldsymbol{a}}$和$\overline{\boldsymbol{a}}'$称为筛后比特串.

7) Alice和Bob协商在$2n$个位置中随机地选择n个位置，然后通过公开信道公布各自这n个位置上的比特值.

8) 双方分别计算n个位置中使用Z基和X基的比特构成的两个串与对方对应串的错码率e_Z和e_X. 若$e_Z > e_{\max}$或$e_X > e_{\max}$，则终止协议.

9) Alice和Bob在各自剩下的约n个比特上通过公开信道进行信息调和和保密增强以获得m个比特的共享密钥.

在上述方案中，基串的每一项取1或取0的概率都要根据概率串中相应的项来调整，且这种调整具有一定的随机性，因此将这个方案命名为PARK(probabilistic adaptive random key generation).

在PARK方案中，由于概率串的第i项是p_i，所以双方基串的第i项为1和0的概率分别为p_i和$1-p_i$. 由大数定律可知基串中1和0出现频率的期望值分别为

$$\int_0^1 x\,\mathrm{d}x=\frac{1}{2}$$

和

$$\int_0^1 (1-x)\,\mathrm{d}x=\frac{1}{2}$$

也就说当基串长度足够长时，基串中1和0出现的频率基本相同.

由定理2知，在PARK的步骤6)，$\boldsymbol{a}$和$\boldsymbol{a}'$的长度约为

$$N\times\frac{2}{3}=(3n+\delta)\,\frac{2}{3}\approx 2n$$

因此整体来看，要产生长度为n的原始密钥，BB84需发送约$4n$个qubit，而PARK只需发送约$3n$个. 若只考虑量子信道，PARK的效率是$\frac{2}{3}$，比BB84高约$17\% \approx \frac{1}{6}=\frac{2}{3}-\frac{1}{2}$.

对 LCA-BB84，如果要求发送的 qubit 数不超过$(4+\delta)n$，那么产生基串时所依据的概率串$\boldsymbol{p}=(p, p, \cdots, p)_N$ 中的 p 值必须大于$\frac{1}{3}$. 由于 $1-2p(1-p)$ 作为 p 的函数在区间$(\frac{1}{3}, \frac{1}{2}]$上是严格单调递减的，故 LCA-BB84 此时在量子信道上的效率不会超过$1-2p(1-p)<\frac{5}{9}<\frac{2}{3}$. 在发送 qubit 的总数为$(3+\delta)n$ 的情况下，即便按$\frac{5}{9}$ 来算，筛后比特串的长度的期望值也只能是 $3n\times\frac{5}{9}<2n$. 去掉 n 个参与对比检验的比特，原始密钥的长度就不足 n 了. 因此不论是从量子信道看还是从整体来看，PARK 也都优于 LCA-BB84.

4 安全性讨论

在对 BB84 的安全性证明中，文献[7]给出的证明方法可能是最简单的. 文献[7]实际上给出了一个基于量子纠缠纯化协议(EPP)的量子密钥分配协议，并证明其是安全的，然后证明这个协议的安全蕴含 BB84 的安全. 文献[6]给出了分类对比检验中的两个错码率 e_X 和 e_Z 与 EPP 中的两个错误率 — 比特翻转差错(bit-flip error)$e^{\text{bit-flip}}$ 和相位差错(phase error)e^{phase} 之间的一个转换关系：

$$e^{\text{bit-flip}}=qe_Z+(1-q)e_X$$

$$e^{\text{phase}}=qe_X+(1-q)e_Z$$

其中 q 和 $1-q$ 分别为最后的密钥中对应于 Z 基和 X 基的比特所占的比例. 可以看出，适当选择 $e_{\max}$，只要

$$0\leqslant e_Z, e_X<e_{\max}-\delta_e$$

其中 δ_e 是一个小的正整数，就能保证

$$0\leqslant e^{\text{bit-flip}}, e^{\text{phase}}<11\%$$

成立. 因此文献[7]的论证可移植来证明 LCA-BB84 的安全性[6]. PARK 与 LCA-BB84 最主要的区别是在基串的产生上，故 PARK 的安全性可类似地证明. 在 LCA-BB84 中，$q=1$，所以 $e^{\text{bit-flip}}=e_Z$，$e^{\text{phase}}=e_X$；而在 PARK 中，可以取 $q=\frac{1}{2}$，这时 $e^{\text{bit-flip}}=e^{\text{phase}}=\frac{e_Z+e_X}{2}$.

在 PARK 中，Alice 和 Bob 要事先共享一个概率串 $\boldsymbol{p}$. $\boldsymbol{p}$ 可以由双方共同协商产生，也可用

$$(\text{authenstr}, \text{nonce}_A, \text{nonce}_B)$$

作为种子，通过一个安全的伪随机数生成器产生. 这里 authenstr 是 Alice 和 Bob 用作身份认证的信息，而 nonce_A 和 nonce_B 分别是由 Alice 和 Bob 即时随机产生的数. 在这种情况下，因为 authenstr 只有 Alice 和 Bob 才有，所以 $\boldsymbol{p}$ 也只有 Alice 和 Bob 才能产生. 这时 PARK 还具有身份认证的作用：因为 Alice 和 Bob 能产生 $\boldsymbol{p}$，所以在步骤 6) 基串的比较中 $\text{Correlation}(\boldsymbol{b}, \boldsymbol{b}')\approx\frac{2}{3}$；如果有一方是假冒的，那么 $\text{Correlation}(\boldsymbol{b}, \boldsymbol{b}')\approx\frac{1}{2}$，因此可被另一方发现.

5 结束语

提高量子密钥分配的效率对量子密码的实际应用和普及有着积极的作用. 在 LCA-BB84 方案中，双方通过加大选择一个基的概率，减少选择另一个基的概率，使得量子信道的效率得到了提高. 但这也导致得到的数据中对应两个基的两类数据中有一类要少于另一类. 为了保证检测数据有足够的精确度，两类数据的数目都要足够多，才能保证有足够多的数据被抽样来进行检验. 因此要么加大 qubit 的传送数目，要么减少选择两种基的概率之差. 这样一来，LCA-BB84 方案的总体效率就受到制约. 在 PARK 方案中，由于两种基的选择总体上是平均的，因此得到的两类数据比例也是平均的，总体效率不受检测的影响. 另外，适当调整 PARK 中概率串的生成方式，就可使 PARK 具有身份认证的功能.

参考文献:

[1] SHOR P. Polynomial－Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer [J]. SIAM J Comput, 1997(26): 1484－1509.

[2] SHOR P W. Algorithms for Quantum Computation: Discrete Logarithms and Factoring [C]// 1994 Proceedings Symposium on Foundations of Computer Science. New York: IEEE Computer Society Press, 1994: 124－134.

[3] BENNETT C H. Quantum Cryptography : Public Key Distribution and Coin Tossing [C]// IEEE International Conference on Computers Systems and Signal Processing. New York: IEEE Computer Society Press, 1984: 175－179.

[4] NIELSEN M A, CHUANG I L. Quantum Computation and Quantum Information [M]. Cambridge: CambridgeUniversity Press, 2000.

[5] ASSCHE GILLES VAN. Quantum Cryptography and Secret－key Distillation [M]. Cambridge: CambridgeUniversity Press, 2006.

[6] LO H K, CHAU H F, ARDEHALI M. Efficient Quantum Key Distribution Scheme and a Proof of Its Unconditional Security [J]. Journal of Cryptology, 2005, 18(2): 133－165.

[7] SHOR P W, PRESKILL J. Simple Proof of Security Of The BB84 Quantum Key Distribution Protocol [J]. Physical Review Letters, 2000, 85(2): 441－444.

A Scheme for Generating Two Correlated Random Bit Strings

BAO Xiao-min[1], QU Yun-yun[2]

1. *School of Mathematics and Statistics, Southwest University, Chongqing 400715, China;*
2. *School of Mathematics Science, Guizhou Normal University, Guiyang 550001, China*

Abstract: This paper presents a method that can increase the efficiency of the quantum key distribution scheme BB84 by 17%. In this method, a probabilistic string, whose elements are randomly and uniformly selected from the interval (0, 1), is shared by the two communication parties. During a quantum bit transmission and reception process, a corresponding element from the probabilistic string is used as a probability for the two communication parties to choose one of the two polarization bases. A refined data analysis process is then used to do the subsequent data analysis.

Key words: quantum cryptography; quantum key distribution; BB84; correlation; probabilistic string

责任编辑 张 枸