Whitenoise 密码 Wu 破译方法的分析与改进

金晨辉　　张　斌　　张远洋
(信息工程大学电子技术学院　郑州　450004)

摘　要　Whitenoise 是由 BSB Utilities 公司提出的一个序列密码算法。Wu 在 2003 年 8 月巧妙地给出了破译 Whitenoise 算法的一个解方程组方法。该文对 Wu 的破译算法进行了深入分析，证明了 Wu 方法的两个基本假设是错误的，因而 Wu 的方法不可能求出正确密钥。此外，该文还对 Wu 的破译方法进行了改进，给出了求解 Whitenoise 密码的秘密整数和秘密素数的方法，并给出了对 Whitenoise 密码的一个预测攻击方法，利用该方法可由其前 80445 个乱数求出其任一时刻的乱数。此外，该文还给出了求出其全部秘密要素的一个思路。

关键词　Whitenoise 序列密码, 密码分析, 预测攻击, 等效密钥
中图分类号：TP309.7　　**文献标识码**：A　　**文章编号**：1009-5896(2006)08-1530-03

Analysis and Improvment of the Wu Breaking Algorithm for the Cipher Whitenoise

Jin Chen-hui　　Zhang Bin　　Zhang Yuan-yang
(*Institute of Electronic Technology*, *the University of Information Engineering*, *Zhengzhou* 450004, *China*)

Abstract　Whitenoise is a stream cipher proposed by the BSB Utilities Inc. In August 2003, Wu Hongjun proposed a breaking algorithm for Whitenoise. In this paper, the authors make an in-depth analysis for Wu's breaking algorithm, and prove that the basic hypotheses of Wu are wrong, and that Wu's algorithm can not find the correct key of Whitenoise. Furthermore, Wu's algorithm is improved and a method obtaining the secret integer and the secret prime numbers is given. A forecast attack to Whitenoise is proposed, with which one can obtain signal for each clock. Furthermore, a think to obtain all the equivalent secret factors is given.

Key words　Stream cipher Whitenoise, Cryptanalysis, Forecast attack, Equivalent key

1　引言

Whitenoise 是由 BSB Utilities 公司提出的一个序列密码算法，文献[1]指出 Whitenoise 拥有序列密码的所有特性。文献[2]给出了 Whitenoise 的安全性评估报告，该报告没有发现 Whitenoise 的任何缺陷。但是不久，Wu 就在文献[3]给出了破译 Whitenoise 的一个巧妙算法。该方法将破译 Whitenoise 的问题归结为求解一个由 80189 个变量字节的线性方程组的问题，并认为只要有足够多的已知明文，就能求出正确密钥，从而达到破译 Whitenoise 的目的。然而，本文发现该破译方法依赖的两个基本假设是错误的，因而不可能求出正确密钥。此外，本文还对该破译方法进行了改进，提出了求出 Whitenoise 的前两个秘密要素的方法，并给出了对 Whitenoise 的预测攻击方法，给出了求解其全部秘密要素的一个思路。

2　Whitenoise 序列密码算法介绍

Whitenoise 密码算法的密钥由 Seed1 和 Seed2 两个种子密钥组成，其中 Seed1 为 1600–2400bit，Seed2 为 32bit。将 Seed1 开方取小数点后的序列，再通过特定的生成算法生成用于 Whitenoise 算法的 4 个秘密要素。Seed2 是上述算法中初始化函数的一个参数。

Whitenoise 序列密码的 4 个秘密要素分别为

(1) 整数 n，其取值范围为 $50 \le n \le 99$；

(2) n 个不同的素数 $p_{i_1}, p_{i_2}, \cdots, p_{i_n}$，其中 $p_1 < p_2 < \cdots < p_{172}$ 是 ≤ 1021 的全部素数；

(3) n 个代替表 $S^{(1)}, S^{(2)}, \cdots, S^{(n)}$，其中 $S^{(i)}: \{0,1,\cdots, p_i - 1\} \to \{0,1\}^8$。将 $S^{(i)}(j)$ 简记为 $S_j^{(i)}$；

(4) 双射 $T: \{0,1\}^8 \to \{0,1\}^8$.

设第 i 个明文字节为 m_i，则 Whitenoise 序列密码对 m_i 的加密过程为

(1) 计算 $z_i = S^{(1)}_{i \bmod p_{i_1}} \oplus S^{(2)}_{i \bmod p_{i_2}} \oplus S^{(3)}_{i \bmod p_{i_3}} \oplus \cdots \oplus S^{(n)}_{i \bmod p_{i_n}}$，其中 $\oplus$ 为逐位模 2 加；

(2) 计算 $y_i = T(z_i)$，并算出密文 $c_i = m_i \oplus y_i$。

3　Whitenoise 密码的 Wu 破译算法

直接攻击 Whitenoise 的种子密钥是不可行的。比较现实的是求出由种子密钥产生的 4 个密钥要素。但是，由于 n 共有 50 种选择，且 $p_{i_1}, p_{i_2}, \cdots, p_{i_n}$ 共有 $C_{172}^n \ge C_{172}^{50} \approx 2^{150}$ 种可能，因而对密钥要素(1)和(2)进行穷举也是不可行的。文献[3]巧妙地回避了对秘密要素(1)和(2)的穷举问题，通过假设没有选用的代替表的输出值全为 0，将 Whitenoise 的加密过程转

2004-12-29 收到，2005-05-08 改回
河南省杰出青年科学基金(0312001800)资助课题

化为所有代替表全部选用的模型:

$$z_i = S^{(1)}_{i \bmod p_{i_1}} \oplus S^{(2)}_{i \bmod p_{i_2}} \oplus S^{(3)}_{i \bmod p_{i_3}} \oplus \cdots \oplus S^{(172)}_{i \bmod p_{172}}$$

再根据 $z_i = z_j$ 等价于 $y_i = y_j$，就可利用两个相等的乱数 y_i 和 y_j 建立一个以 $S^{(t)}_j, 1 \le t \le 172,\ 0 \le j \le p_i - 1$ 为未知变量的线性方程：

$$\bigoplus_{t=1}^{172} S^{(t)}_{i \bmod p_t} \oplus \bigoplus_{t=1}^{172} S^{(t)}_{j \bmod p_t} = 0 \tag{1}$$

文献[3]认为，由于该方程只有 $\sum_{i=1}^{172} p_i = 80189$ 个未知变量，故当已知的乱数个数远大于 256+80189 时，就可得到一个满秩的方程组，从而求出诸 $S^{(t)}_j$ 的唯一解；再将 $S^{(t)}_0 = S^{(t)}_1 = \cdots = S^{(t)}_{p_t - 1} = 0$ 的代替表 $S^{(t)}$ 确定为没有选用的代替表，其它的确定为已选用的代替表，就可求出秘密要素(1)至(3)；最后利用双射 T 的输入和输出都已知这个条件，就可确定 T，从而实现对 4 个秘密要素的已知乱数攻击。整个攻击的计算复杂性为 $O(2^{48.4})$。

但不幸的是，本文将证明 Wu 建立的方程组的秩一定 $\le 80189 - 172$，因而求出正确的 $S^{(t)}_m$ 的计算复杂性至少为 256^{172}。我们还证明了 $S^{(t)}$ 未被选用也不等价于 $S^{(t)}$ 的输出值全为 0，故 Wu 的方法无法求出正确密钥，因而无法实现对 Whitenoise 的破译。

4　Wu 的破译算法的分析与改进

4.1　Wu 破译算法的分析

首先指出，由于 Wu 建立的方程组的系数矩阵是 0,1 矩阵，故可仅利用模 2 加运算完成方程组的求解，因而可将之看作二元域上的方程组进行求解，这等价于并行地求解由字节的 8 个分量构成的 8 个方程组。

设 $\boldsymbol{S} = \left(S^{(1)}_0, S^{(1)}_1, S^{(2)}_0, S^{(2)}_1, S^{(2)}_2, \cdots, S^{(172)}_0, \cdots, S^{(172)}_{1020}\right)^{\mathrm{T}}$ 是 GF(256) 上的 80189 维向量，且 $\alpha_i \boldsymbol{S} = 0$ 是形如式(1)的方程，这里 $1 \le i \le m$，m 是方程的个数。再设 $\boldsymbol{A}$ 是该方程组 $\boldsymbol{AS} = 0$ 的系数矩阵，则有

定理 1　(1) 矩阵 $\boldsymbol{A}$ 的秩 $\le \sum_{i=1}^{172} p_i - 172$，因而方程组 $\boldsymbol{AS} = 0$ 的解 $\boldsymbol{S}$ 至少有 172 个自由未知量；

(2)　$\forall i$，$S^{(i)}_0, S^{(i)}_1, \cdots, S^{(i)}_{p_i - 1}$ 中自由未知量的个数 $\ge p_i - $ 秩$(A_i) \ge 1$；

(3)　秩 $(A) = \sum_{i=1}^{172} p_i - 172 \Leftrightarrow \forall i$，均有 $S^{(i)}_0 = S^{(i)}_1 = \cdots = S^{(i)}_{p_i - 1}$。

证明　(1) $\forall i: 1 \le i \le 172$，记 $\boldsymbol{A}_i$ 是 $\boldsymbol{A}$ 的第 $1 + \sum_{t<i} p_t$ 列至第 $\sum_{t \le i} p_t$ 列构成的矩阵，则 $\boldsymbol{A}_i$ 的每行要么全为 0，要么恰有两个 1，故 $\boldsymbol{A}_i$ 的各列的模 2 和为 0，因而 $\boldsymbol{A}_i$ 的列向量线性相关，从而秩 $(A_i) \le p_i - 1$，这说明秩 $(A) \le \sum_{i=1}^{172}$秩$(A_i) = \sum_{i=1}^{172} p_i - 172$，即(1)成立。

(2) 记 $A_i = \left(a_{i,0}, \cdots, a_{i,p_i - 1}\right)$，则方程组 $\boldsymbol{AS} = 0$ 等价于 $\bigoplus_{i=1}^{172} \bigoplus_{j=0}^{p_i - 1} a_{i,j} S^{(i)}_j = 0$。$\forall i$，记秩 $(A_i) = r_i$，则存在 $t_1, \cdots, t_{r_i}$，使得 $\forall j: 0 \le j < p_i$，均存在 $c_{i,j,1}, c_{i,j,2}, \cdots, c_{i,j,r_i} \in \{0,1\}$，使得 $a_{i,j} = \bigoplus_{m=1}^{r_i} c_{i,j,m} a_{i,t_m}$，从而

$$\begin{aligned}\boldsymbol{AS} = \bigoplus_{i=1}^{172} \bigoplus_{j=0}^{p_i - 1} a_{i,j} S^{(i)}_j &= \bigoplus_{i=1}^{172} \bigoplus_{j=0}^{p_i - 1} \left[\bigoplus_{m=1}^{r_i} c_{i,j,m} a_{i,t_m} \right] S^{(i)}_j \\ &= \bigoplus_{i=1}^{172} \bigoplus_{m=1}^{r_i} a_{i,t_m} \left[\bigoplus_{j=0}^{p_i - 1} c_{i,j,m} S^{(i)}_j \right]\end{aligned}$$

令 $T^{(i)}_m = \bigoplus_{j=0}^{p_i - 1} c_{i,j,m} S^{(i)}_j$，则方程组 $\boldsymbol{AS} = 0$ 等价于 $\bigoplus_{i=1}^{172} \bigoplus_{m=1}^{r_i} a_{i,t_m} T^{(i)}_m = 0$。由秩 $(A_i) = r_i$ 知 $T^{(1)}_m, T^{(2)}_m, \cdots, T^{(r_i)}_m$ 线性无关，再由 $r_i \le p_i - 1$ 说明(2)成立。

(3) 充分性　$S^{(i)}_0 = S^{(i)}_1 = \cdots = S^{(i)}_{p_i - 1}$ 对诸 i 成立说明 $S^{(i)}_0, S^{(i)}_1, \cdots, S^{(i)}_{p_i - 1}$ 中只有一个自由变量，因而方程 $\boldsymbol{AS} = 0$ 的自由变量的个数为 172，故秩 $(A) = \sum_{i=1}^{172} p_i - 172$。

必要性　设 $q_0 = 0$，且 $q_i = \sum_{j=1}^{i} p_j$，$\boldsymbol{e}_i$ 为仅第 i 分量为 1 的 q_{172} 维二元行向量。$\forall i \ne q_i$，令 $\boldsymbol{d}_i = \boldsymbol{e}_i \oplus \boldsymbol{e}_{i+1}$，则易证如此定义的 $q_{172} - 172$ 个向量线性无关。现证 $\boldsymbol{A}$ 的行向量都可由这些 $\boldsymbol{d}_i$ 在二元域上线性表出。

事实上，设 $b = \bigoplus_{i=1}^{172} b_i$ 是 $\boldsymbol{A}$ 的一个行向量，其中 b_i 除第 $q_{i-1} + 1$ 至第 q_i 分量外全部为 0，则由 b_i 的重量为偶数或 0 知 b_i 可由 $d_{q_{i-1}+1}, d_{q_{i-1}+2}, \cdots, d_{q_i}$ 在二元域上线性表出，即存在 $c_{i,1}, c_{i,2}, \cdots, c_{i,p_i - 1} \in \{0,1\}$，使得

$$b_i = \bigoplus_{j=1}^{p_i - 1} c_{ij} d_{q_{i-1}+j}$$

从而 $b = \bigoplus_{i=1}^{172} b_i = \bigoplus_{i=1}^{172} \bigoplus_{j=1}^{p_i - 1} c_{ij} d_{q_{i-1}+j}$ 可由诸 $\boldsymbol{d}_i$ 在二元域上线性表出。

又因当秩$(A) = q_{172} - 172$ 时，$\boldsymbol{A}$ 的行向量的秩与 $\boldsymbol{d}_i$ 的个数相同，因而 $\boldsymbol{A}$ 的行向量生成的子空间与所有 $\boldsymbol{d}_i$ 生成的子空间相同，故诸 $\boldsymbol{d}_i$ 可由 $\boldsymbol{A}$ 的行向量表出，即存在二元域上的 $m \times m$ 可逆矩阵 $\boldsymbol{B}$，使得 $\boldsymbol{D} = \boldsymbol{BA}$ 的第 i 行在 $i \ne q_i$ 时为 $\boldsymbol{d}_i$，否则为 $\boldsymbol{0}$ 向量，这说明方程组 $\boldsymbol{AS} = 0$ 与 $\boldsymbol{DS} = 0$ 等价，即 $\forall i$，均有 $S^{(i)}_0 = S^{(i)}_1 = \cdots = S^{(i)}_{p_i - 1}$。　证毕

定理 1 说明，Wu 建立的方程组至少有 172 个自由未知量，因而至少有 256^{172} 个解，要想从中找出正确的解是不可能的。定理 1 之(3)还说明，Wu 建立的方程组的自由未知量的个数一般都大于 172。

4.2　Whitenoise 的破译方法

下面我们给出破译 Whitenoise 的一种有效方法。我们的方法是找出秘密要素(1)和(2)，并给出所有等效的秘密要素(3)和(4)。

显然，并非 Wu 建立的方程组的每个解都是 Whitenoise

的等效解。我们的目的是找出其全部的等效解。

首先给出秘密要素(1)和(2)的求解方法。

设 $s_1,\cdots,s_k$ 是 Wu 方程组解的自由未知量，$\boldsymbol{s}=(s_1,\cdots,s_k)^{\mathrm{T}}$，则通过对 Wu 方程组的求解，对 $\forall i:1\le i\le 172$，$\forall t:0\le t<p_i$，都可求出 k 维二元向量 $\boldsymbol{\beta}_t^{(i)}$ 使得 $\boldsymbol{S}_t^{(i)}=\boldsymbol{\beta}_t^{(i)}\boldsymbol{s}$。记 $\boldsymbol{B}_i$ 是以 $\boldsymbol{\beta}_t^{(i)}$ 为第 $t+1$ 行构成的 $p_i\times k$ 矩阵，则有

定理 2 $\boldsymbol{S}^{(i)}$ 未被选用的充要条件秩 $(B_i)=1$。

证明 由于秩 $(B_i)=1$ 等价于 $S_0^{(i)}=S_1^{(i)}=\cdots=S_{p_i-1}^{(i)}$，即 $S^{(i)}$ 是常值函数，因而它等价于 $S^{(i)}$ 未被选用。

由定理 2，根据诸秩 $(B_i)=1$ 是否成立，就可确定 Whitenoise 的秘密要素(2)和(1)。这样，就解决了秘密要素(1)和(2)的求解问题。显然，$S^{(i)}$ 未被选用等价于 $S_0^{(i)}=S_1^{(i)}=\cdots=S_{p_i-1}^{(i)}$，但文献[3]认为 $S^{(i)}$ 未被选用等价于 $S_0^{(i)}=S_1^{(i)}=\cdots=S_{p_i-1}^{(i)}=0$ 是不正确的。

接着给出对 Whitenoise 的预测攻击方法。

为简单起见，我们将未被选用的代替表所含的自由变量都设置为 0，从而消去未被选用的代替表所含的自由变量。记经过这样处理后保留下来的自由变量为 $s'_1,s'_2,\cdots,s'_{k'}$，并令 $s'=(s'_1,s'_2,\cdots,s'_{k'})^{\mathrm{T}}$。同时我们再做一个合理的假设，即 $z_0,z_1,\cdots,z_t,\cdots$ 能够取遍 GF(256) 中的所有值。

定理 3 (1) $\forall t\ge 0$，均可求出 k' 维二元向量 $\boldsymbol{\alpha}_t$，使得 $\boldsymbol{z}_t=\boldsymbol{\alpha}_t s'$；

(2) $\forall t_1,t_2\ge 0$，$y_{t_1}=y_{t_2}$，$z_{t_1}=z_{t_2}$ 和 $\alpha_{t_1}=\alpha_{t_2}$ 三者等价。

证明 (1) 由 $\boldsymbol{z}_t=S_{t\bmod p_1}^{(1)}\oplus S_{t\bmod p_2}^{(2)}\oplus S_{t\bmod p_3}^{(3)}\oplus\cdots\oplus S_{t\bmod p_{172}}^{(172)}=\left[\bigoplus_{j=1}^{172}\beta_{t\bmod p_j}^{(j)}\right]s$ 即证。(2)显然。

根据定理 3，我们可对 t 按照乱数 y_t 的值分为 256 个等价类，每个等价类中的 y_t 对应相同的 $\boldsymbol{z}_t$，也对应相同的 $\boldsymbol{\alpha}_t$；不同类中的 y_t 对应不同的 $\boldsymbol{z}_t$，也对应不同的 $\boldsymbol{\alpha}_t$，这说明 $\boldsymbol{\alpha}_t$ 共有 256 个不同取值。据此我们就可提出对 Whitenoise 的预测攻击方法。

对 Whitenoise 的预测攻击算法:

步骤 1 解方程组(1)得到诸 $\boldsymbol{\beta}_t^{(i)}$；

步骤 2 找出 256 个互不相同的乱数 $y_{i_0},y_{i_1},\cdots,y_{i_{255}}$，并利用诸 $\boldsymbol{\beta}_t^{(i)}$ 求出上述乱数对应的 $\alpha_{i_0},\alpha_{i_1},\cdots,\alpha_{i_{255}}$；

步骤 3 $\forall t\ge 1$，利用 $\boldsymbol{\alpha}_t=\bigoplus_{j=1}^{172}\beta_{t\bmod p_j}^{(j)}$ 求出 $\boldsymbol{\alpha}_t$，进而求出使 $\boldsymbol{\alpha}_t=\alpha_{i_k}$ 的 α_{i_k}，则 $y_t=y_{i_k}$。

上述预测攻击说明，由 Whitenoise 的前 $80189+256=80445$ 个乱数就可求出该算法任一时刻的乱数，其乱数序列完全由其前 80445 个乱数决定。

最后给出求解 Whisenoise 的秘密要素(3)和(4)的等效解的一个思路。

现设 $y_{j_0},y_{j_1},\cdots,y_{j_{255}}$ 是互不相同的乱数，则自由变量 $s'_1,s'_2,\cdots,s'_{k'}$ 的真实取值一定使 $z_{j_0},z_{j_1},\cdots,z_{j_{255}}$ 互不相同。由于等价的 t 对应相同的 α_t，故自由变量的每种取值都使等价的 t 产生相同的 $\boldsymbol{z}_t$，因而产生相同的 y_t。因此，只要自由变量的一种取值能使产生的 $z_{j_0},z_{j_1},\cdots,z_{j_{255}}$ 互不相同，就可通过令诸 $T(z_{j_t})=y_{j_t}$ 构造出 T，再由 $s'_1,s'_2,\cdots,s'_{k'}$ 的取值求出诸 $S_j^{(i)}$，从而求出一个 T 和一组 $S_j^{(i)}$。

按上述方法确定的 T 和诸 $S_j^{(i)}$ 未必是真实的 T 和诸 $S_j^{(i)}$，但由于 t 的每个等价类均对应相同的 $\boldsymbol{\alpha}_t$，因而按这种方法确定的 T 和诸 $S_j^{(i)}$ 产生的乱数一定和真实的乱数相同，即按这种方法求出的 T 和诸 $S_j^{(i)}$ 是等效的秘密要素(3)和(4)。因此，通过求出等效的秘密要素(3)和(4)来完成 Whitenoise 的破译问题，就转化为如何给出自由变量的一种取值，使得 $z_{j_0},z_{j_1},\cdots,z_{j_{255}}$ 互不相同的问题。该问题可具体描述为:

在已知 256 个互不相同的 k 维二元向量 $\alpha_0,\alpha_1,\cdots,\alpha_{255}$ 的条件下，如何求出一个 $s\in[\mathrm{GF}(256)]^k$ 使得 $\alpha_0 s,\alpha_1 s,\cdots,\alpha_{255}s$ 互不相同。其中 $\alpha_i s$ 是 $[\mathrm{GF}(256)]^k$ 中向量的点积。

这个问题是本文留下的一个公开问题。如果能求出其一个解，就可得到 Whitenoise 的一个等效密钥；如果能求出其全部解，就可得到 Whiteniose 密码全部等效密钥。

我们对将 172 修改为 4 后得到的缩减 Whitenoise 算法进行了攻击实验。实验结果表明，我们的确能够求出 Whitenoise 的等效密钥。

5 结束语

本文对 Wu 提出的破译 Whitenoise 密码算法的方法进行了深入分析，发现该方法不能按作者预期的那样求出正确密钥。同时，我们提出了求出 Whitenoise 的前两个秘密要素的方法，并给出了对 Whitenoise 的预测攻击方法，给出了求解其全部秘密要素的一个思路。如何最终实现对 Whitenoise 的全部秘密要素的求解，还有待进一步研究。

参 考 文 献

[1] Dr Traore Issa, Michael Liu Yanguo. Evaluation of Whitenoise cryptosystem. http://eprint.iacr.org/2003/B539_0003, 2003(2).

[2] Wagner David. A security evaluation of Whitenoise. http://eprint.iacr.org/2003/218.

[3] Wu Hongjun. Breaking the stream cihper Whitenoise. http://eprint.iacr.org/2003/250.

[4] Boren Stephen, Brisson Andre. Software specificationsfor tinnitus utilizing Whitenoise substitution stream cihper. http://eprint.iacr.org/2003/249.

金晨辉： 男，1965 年生，博士，教授，博士生导师，主要研究方向为密码理论和信息安全.

张 斌： 男，1982 年生，硕士，研究方向为密码学.

张远洋： 男，1982 年生，硕士，研究方向为密码学.