The Study on Quantum Algorithm using NMR Computers for Code Breaking of Secret Key Cryptosystems

Kazuo Ohta† Tetsuro Nishino† Seiya Okubo*

*The Graduate School of Electro-Communications
†The University of Electro-Communications
1-5-1 Chofugaoka, Chofu, Tokyo, 182-8585, Japan
e-mail: ota@ice.uec.ac.jp nishino@ice.uec.ac.jp s-okubo@ice.uec.ac.jp

Abstract: In this paper, we discuss quantum algorithms that find a secret key k_0 satisfying $c_0 = E(k_0, m_0)$ given m_0 and c_0, where an encryption algorithm E is publicly available, k_0 is a secret key, m_0 is plaintexts and c_0 is ciphertexts. We will propose a new algorithm suitable for implementation by NMR (Nuclear Magnetic Resonance) computer based on the technique to solve the counting problem. This algorithm can solve the problem with less complexity when the measurement of accuracy of NMR computer is improved more. We will consider the possibility that the proposed algorithm is superior to the Grover's algorithm based on the small experimental results. [1]

1 The Setting of Problem

A secret-key cryptosystem, a pair of $E(k, m)$ and $D(k, c)$, is given satisfying:

$$E : \mathcal{K} \times \mathcal{P} \longrightarrow \mathcal{C},$$
$$D : \mathcal{K} \times \mathcal{C} \longrightarrow \mathcal{P}, \text{ and}$$
$$D(k, E(k, m)) = m \text{ for } \forall k \in \mathcal{K} \text{ and } \forall m \in \mathcal{P}.$$

The secret-key search problem is defined as follows:

Definition 1 The inputs of the problem are a circuit for an encryption procedure of E and a pair of $(m_0, c_0) \in \mathcal{P} \times \mathcal{C}$ satisfying $c_0 = E(k_0, m_0)$ for some secret $k_0 \in \mathcal{K}$. The output of the problem is the value of k_0.

2 The Proposed Algorithm

2.1 Overview

Define a function f and g_r for a given (m_0, c_0) as follows:

$$f : \mathcal{K} \ni k \mapsto E(k, m_0) \in \mathcal{C},$$
$$r : \mathcal{C} \longrightarrow \mathcal{K} \text{ is any surjective map, and}$$
$$g_r = r \circ f : \mathcal{K} \ni k \mapsto r(E(k, m_0)) \in \mathcal{K}.$$

For any $d \in \mathbf{N}$, define a function $F_r^{(d)}$ ($F_r^{(d)} : \mathcal{K} \longrightarrow \{0, 1\}$) as follows:

$$\text{for } i \in \mathcal{K},$$
$$F_r^{(d)}(i) = 1 \quad \text{iff} \quad f(g_r^{(d-1)}(i)) = c_0,$$

where $g_r^{(0)}(i) = i$ and $g_r^{(j)}(i) = g_r(g_r^{(j-1)}(i))$.

The proposed algorithm consists of two phases:
Phase I: Select an appropriate subset $\mathcal{R}$ among all functions, $\{r : \mathcal{C} \longrightarrow \mathcal{K}\}$, and find d_0 and r_0 using NMR satisfying $1 \in F_{r_0}^{(d_0)}(\mathcal{K})$.
Phase II: Find i_0 and k_0 using NMR with an appropriate r_0 and d_0 satisfying the relations of $F_{r_0}^{(d_0)}(i_0) = 1$.

The fact of $F_{r_0}^{(d_0)}(i_0) = 1$ implies that $E(k_0, m_0) = f(k_0) = c_0$, where $k_0 = g_{r_0}^{(d_0-1)}(i_0)$, by the constructions of $F_r^{(d)}$ and g_r.

Define a target set $T_{d,r}$ for $F_r^{(d)}$ by $T_{d,r} = \{i \in \mathcal{K} \mid F_r^{(d)}(i) = 1\} = (F_r^{(d)})^{-1}(1)$. It has a tendency that when d increases, $T_{d,r}$ also increases.

2.2 Notations

Let T_0 be the threshold values with which NMR can distinguish the case where $t = 0$ and the case of $t > T_0$. The value of T_0 depends on the implementation technologies of NMR computers.

Let T_0 be represented as $T_0 = 2^{k-l}$, where l is a positive integer and is called measurement accuracy.

Let d_r be the minimal value of d with which $T_{d,r} \geq T_0$ holds for a given r. Generally d_r depends on r.

2.3 Procedures

Two phases are described as follows in detail:

Phase I: Input: E and (m_0, c_0).
Output: d_0, r_0 such that $1 \in F_{r_0}^{(d_0)}(\mathcal{K})$.

Step 1: Select an arbitrary subset $\mathcal{R}$ of maps, that is, $(\mathcal{R} \ni) r : \mathcal{C} \longrightarrow \mathcal{K}$ is surjective. Define

$$F^{(d)}(\cdot) : \mathcal{K} \times \mathcal{R} \ni (i, r) \mapsto F_r^{(d)}(i) \in \{0, 1\}$$

Step 2: Get the value of $T_{d,r}$ using NMR by increasing of d gradually, for a given $r \in \mathcal{R}$. If $T_{d,r} \geq T_0$ for some $r \in \mathcal{R}$, then go to **Phase II**. Denote the first pair of (d_r, r) satisfying $T_{d_r,r} \geq T_0$ by (d_0, r_0). Otherwise return to **Step 1**.

[1] Detail information on this paper is available at http://www.tnlab.ice.uec.ac.jp/papers/ota/eqis02.ps

In **Phase I**, the following property of the measurement of $T_{d,r}$ is essential to ensure the complexity discussed in Section 3.3: the superposition of NMR will not be destroyed for a while, that is, we can measure the tape cells in the superposition some constant times before the superposition is destroyed. Thus the value of d_0 can be obtained by **Step 2**, where the value of d increases gradually. The optimal selection of d will be discussed in Section 3.2.

Phase II: Input: (d_0, r_0) obtained in **Phase I**.
Output: k_0 satisfying $c_0 = E(k_0, m_0)$.

Define a function $f_{j=0}$ and $\widetilde{F}^{(d)}_{r,j=0}$ using a given (m_0, c_0) as follows and do the decision procedure of the value of k_0 described below:

$$f_{j=0} : \mathcal{K} \ni (x_1, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_k) \mapsto E(x_1, \ldots, x_{j-1}, 0, x_{j+1}, \ldots, x_k, m_0) \in \mathcal{C},$$

for $i \in \mathcal{K}$,

$$\widetilde{F}^{(d_0)}_{r,j=0}(i) = 1 \quad \text{iff} \quad f_{j=0}(g_r^{(d_0-1)}(i)) = c_0.$$

begin
  for $j = 1$ **to** k **do**
    apply NMR to obtain $\#(\widetilde{F}^{(d_0)}_{j=0})^{-1}(1)$.
    if $\#(\widetilde{F}^{(d_0)}_{j=0})^{-1}(1) \geq T_0$, **then** $[k_0]_j = 0$.
    otherwise $[k_0]_j = 1$.
end.

Figure 1: Decision Procedure of the value of k_0.

3 Experimental Results

We will evaluate the repetition number in Phase I at first using a small examples, where $\#\mathcal{K} = 2^{16}$ and $\#\mathcal{M} = \#\mathcal{C} = 2^{64}$.

3.1 The Tendency of $\frac{K}{\#U_r^{(d-1)}}$

Let $U_r^{(d-1)} = g_r^{(d-1)}(\mathcal{K})$ and $V_r^{(d-1)}(T_0) = \{k \in U_r^{(d-1)} \mid (g_r^{(d-1)})^{-1}(k) \geq T_0\}$.

Figure 2 implies the following conjecture for reasonable $d < (\sqrt{K})$:

$$\textbf{[Conjecture1]} \quad \frac{d \times \#U_r^{(d-1)}}{K} = constant(\approx 2) \qquad (1)$$

3.2 The Success Probability in Phase I

The size of $\mathcal{R}$ used in **Phase I** will be evaluated here.

Table 1 summarizes the optimal points (the value of d-the size of V) and their success probability of each type of r in the cases of $T_0 = 4, 16, 64, 256$, respectively.

T_0	004	016	064	256
Type 1	10-6330	30-1592	180-449	410-49
Type 2	10-6336	30-1508	120-378	330-127
Type 3	10-6419	30-1600	110-333	160-62
Average	10-6362	30-1567	137-387	300-79
Success Prob.(%)	9.70	2.39	0.508∼ 0.685	0.0748∼ 0.193

Table 1: The Optimal Points and Their Success Probability

Table 1 implies the following conjecture for reasonable $d(< 2^{k/4})$:

$$\textbf{[Conjecture2]} \qquad d_0 \approx 2 \times T_0. \qquad (2)$$

3.3 Complexity

We will estimate the complexity of the proposed algorithm by the number of calls of E and r.
Phase I: $\frac{K}{\#V_{r_0}^{(d_0-1)}(T_0)} \times d_0$
Phase II: $k \times d_0$
Total:

$$C(\#\text{ of calls of } E, r) = \frac{K}{\#V_{r_0}^{(d_0-1)}(T_0)} d_0 + k d_0 \qquad (3)$$

Since $U_r^{(d-1)} > V_r^{(d-1)}(T_0)$ and Conjecture 1 (Eq. (1)), $\frac{K}{\#V_{r_0}^{(d_0-1)}(T_0)} d_0 > \frac{d_0^2}{2}$ holds.

Let us consider the possibility that the proposed algorithm is superior to the Grover's algorithm based on the small experimental results. $d_0^2 \leq 2^{\frac{k}{2}+1}$ implies $d_0 \leq 2^{\frac{k}{4}+\frac{1}{2}}$.

Since Conjecture 2 (Eq. (2)) holds and $T_0 = 2^{(k-l)}$, $k - l + 1 \leq \frac{k}{4} + \frac{1}{2}$. So $k \leq \frac{4}{3}(l - \frac{1}{2})$.

This means that if the accuracy of measurement of NMR computers is improved, that is, the value of l becomes larger, the range of k where the proposed algorithm is superior to the Grover's algorithm is $k \leq \frac{4}{3}(l - \frac{1}{2})$.

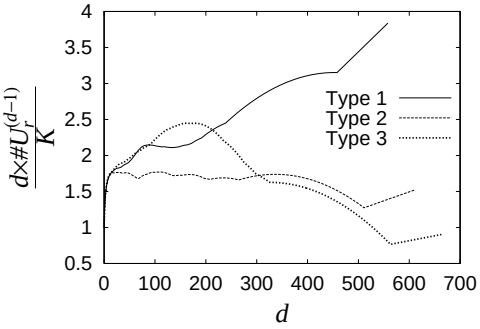

Figure 2: Three types of $\frac{d \times \#U_r^{(d-1)}}{K}$