Some Properties of Quantum Data Search Algorithms

Keisuke Arima[1], Hiromi Miyajima[2], Noritaka Shigei[3] and Michiharu Maeda[4]
[1] Graduate School of Science and Engineering, Kagoshima University, Japan
1-21-40 Korimoto, Kagoshima 890-0065, Japan
[2,3] Faculty of Engineering, Kagoshima University, Japan
1-21-40 Korimoto, Kagoshima 890-0065, Japan
[4] Fukuoka Institute of Technology, Japan
3-30-1 Wajiro-higashi Higashiku, Fukuoka 811-0295, Japan
E-mail: [1]k0483174@kadai.jp, [2,3]{miya,shigei}@eee.kagoshima-u.ac.jp

Abstract: This paper deals with some properties of quantum data search algorithms. First, Grover's and Ventura's algorithms for quantum data search are introduced and compared with each other. As a result, it is shown that both algorithms are not always universal with the number of stored data. Further, some properties on the data search algorithms are shown.

1. Introduction

Many studies on the ability of quantum computer have been done. By using quantum calculation, some algorithms on practical applications such as the factorization and the file searching (Grover's algorithm) are proposed[1]. Grover's algorithm is fast one to find one item in unsorted database. It is effective in the case where all the data are stored in the database[2], [4]. However, it is known that it is not always effective in the case where all the data are not stored. On the other hand, Ventura is proposed an algorithm as generalized Grover's one and insists that it is effective in all cases[3], [5]. In this paper, we will point out that it is not always true and give some properties on the data search algorithms.

2. Grover's and Ventura's quantum algorithms for data searching

The basic unit in quantum computation is a qubit, a superposition of two independent states $|0\rangle$ and $|1\rangle$ corresponding to the states 0 and 1, denoted $c_0|0\rangle + c_1|1\rangle$, where c_0 and c_1 are complex numbers such that $|c_0|^2 + |c_1|^2 = 1$. We use the Dirac bracket notation, where the ket $|i\rangle$ is analogous to a column vector. Let n be a positive integer and $N = 2^n$. A system with n qubits is described using N independent state $|i\rangle$, $0 \leq i \leq N-1$, each associated with probability amplitude c_i, a complex number, as follows: $\sum_{i=0}^{N-1} c_i|i\rangle$ where $\sum_{i=0}^{N-1} |c_i|^2 = 1$. The direction of c_i on the complex plane is called the phase of state $|i\rangle$ and the absolute value $|c_i|$ is called the amplitude of state $|i\rangle$. In quantum system, starting from any quantum state, the desired state is formed by multiplying column vector of the quantum state by unitary matrix. Finally, we can obtain the desired state with high probability through observation[2]. The problem is how we find unitary matrixes. Grover has proposed the fast data searching algorithm.

Let explain the Grover's algorithm shown in Fig.1. Grover has proposed an algorithm for finding one data in an unsorted database. In the conventional computation, if there are N data in the database, it would require $O(N)$ queries to the database. However, Grover has shown how to perform this using the quantum computation with only $O(\sqrt{N})$ queries[2]. Let $Z_N = \{0, 1, \cdots, N-1\}$. Define the following operators:

$$\begin{aligned} I_a &= \text{identity matrix except for} \\ &\quad I(a+1, a+1) = -1,\ a \in Z_N, \end{aligned} \tag{1}$$

which inverts any state $|\psi\rangle$ and

$$\begin{aligned} W(x,y) &= \frac{1}{\sqrt{N}}(-1)^{x_0y_0+\cdots+x_{N-1}y_{N-1}} \\ &\quad \text{for } x = \sum_{i=0}^{N-1} x_i 2^i,\ y = \sum_{i=0}^{N-1} y_i 2^i, \end{aligned} \tag{2}$$

which is called the Walsh or Hadamard transform and performs a special case of discrete Fourier transform. We begin with the $|\bar{0}\rangle$ state and apply W operator, where $|\bar{0}\rangle$ means that the number of 0's is N. As a result, all the states have the sane amplitude $-1/\sqrt{N}$. Next, we apply the I_τ operator, where $|\tau\rangle$ is the state of searching data. Further, we apply the operator

$$G = -WI_0W. \tag{3}$$

Followed by the I_τ operator $T = (\pi/4)\sqrt{N}$ times and observe the system[2]. G operator has been described as inverting all the state's amplitudes around the average one of all states.

Example 1:

Let $n = 4$ and $N = 16$. Let searching data $|\tau\rangle = |6\rangle = |0110\rangle$ and the number stored data $m = 16$.

At step 2,

$$|\psi\rangle = \frac{1}{4}(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1)^{\mathrm{t}}.$$

1. Initial state $|\bar{0}\rangle$
2. $|\psi\rangle = W|\bar{0}\rangle = |\bar{1}\rangle$
3. Repeat T times
4. $|\psi\rangle = I_\tau|\psi\rangle$
5. $|\psi\rangle = G|\psi\rangle$
6. Observe the system

Figure 1. Grover's algorithm.

Then, $|\psi\rangle$ is updated as follows.

$$
\begin{aligned}
I_\tau|\psi\rangle &= \frac{1}{4}(1,1,1,1,1,1,-1,1,1,1,1,1,1,1,1,1)^{\mathrm{t}},\\
G|\psi\rangle &= \frac{1}{16}(3,3,3,3,3,3,11,3,3,3,3,3,3,3,3,3)^{\mathrm{t}},\\
I_\tau|\psi\rangle &= \frac{1}{16}(3,3,3,3,3,3,-11,3,3,3,3,3,3,3,3,3)^{\mathrm{t}},\\
G|\psi\rangle &= \frac{1}{64}(5,5,5,5,5,5,61,5,5,5,5,5,5,5,5,5)^{\mathrm{t}},\\
I_\tau|\psi\rangle &= \frac{1}{64}(5,5,5,5,5,5,-61,5,5,5,5,5,5,5,5,5)^{\mathrm{t}},\\
G|\psi\rangle &= \frac{1}{256}(-13,-13,-13,-13,-13,-13,251,-13,\\
&\quad -13,-13,-13,-13,-13,-13,-13,-13)^{\mathrm{t}}.
\end{aligned}
$$

Finally, the desired pattern 0110 is obtained with the probability 0.96. We can get the searching data with high probability.

Next, supposing that stored data are $|0\rangle$, $|3\rangle$, $|6\rangle$, $|9\rangle$, $|12\rangle$ and $|15\rangle$, and searching data is $|6\rangle$. At step 2 in Fig.1, supposing that

$$|\psi\rangle = \frac{1}{\sqrt{6}}(1,0,0,1,0,0,1,0,0,1,0,0,1,0,0,1).$$

Continuing the same process as above, we obtain

$$
\begin{aligned}
|\psi\rangle = \frac{1}{8\sqrt{6}}(&5,-3,-3,5,-3,-3,13,\\
&-3,-3,5,-3,-3,5,-3,-3,5).
\end{aligned}
$$

In this case, the desired pattern 0110 is obtained with the probability 0.44. It is shown that, in the case where the subset of all the patterns is memorized, Grover's algorithm does not always get a good result. □

On the other hand, Ventura has proposed the generalized algorithm. Let explain Ventura's algorithm as shown in Fig.2. The differences between Grover's and Ventura's algorithms are steps before the repeating part and initial state $|\psi\rangle$. For Ventura's algorithm, the initial state $|\psi\rangle$ is as follows:

$$
|\psi\rangle_i = \begin{cases} 1 & \text{for any } i \text{ that is} \\ & \text{the number for stored data} \\ 0 & \text{otherwise.} \end{cases} \tag{4}
$$

Let define the following operator.

$$
\begin{aligned}
I_\rho &= \text{identity matrix except for}\\
&\quad I(\rho+1,\rho+1) = -1 \text{ for any } \rho\\
&\quad \text{that is the number for stored data.}
\end{aligned} \tag{5}
$$

1. Initial state $|\psi\rangle$
2. $|\psi\rangle = I_\tau|\psi\rangle$
3. $|\psi\rangle = G|\psi\rangle$
4. $|\psi\rangle = I_\rho|\psi\rangle$
5. $|\psi\rangle = G|\psi\rangle$
6. Repeat $\frac{\pi}{4}\sqrt{N} - 2$ times
7. $|\psi\rangle = I_\tau|\psi\rangle$
8. $|\psi\rangle = G|\psi\rangle$
9. Observe the system

Figure 2. Ventura's algorithm.

Ventura's idea is to try to make the amplitudes of all the states except for searching data uniform using steps 3 and 4.

Example 2:

Let apply Ventura's algorithm to the latter problem in Example 1.

$$
\begin{aligned}
|\psi\rangle &= \frac{1}{\sqrt{6}}(1,0,0,1,0,0,1,0,0,1,0,0,1,0,0,1)^{\mathrm{t}},\\
I_\tau|\psi\rangle &= \frac{1}{\sqrt{6}}(1,0,0,1,0,0,-1,0,0,1,0,0,1,0,0,1)^{\mathrm{t}},\\
G|\psi\rangle &= \frac{1}{2\sqrt{6}}(-1,1,1,-1,1,1,3,1,\\
&\qquad 1,-1,1,1,-1,1,1,-1)^{\mathrm{t}},\\
I_\rho|\psi\rangle &= \frac{1}{2\sqrt{6}}(1,1,1,1,1,1,-3,1,1,1,1,1,1,1,1,1)^{\mathrm{t}},\\
G|\psi\rangle &= \frac{1}{4\sqrt{6}}(1,1,1,1,1,1,9,1,1,1,1,1,1,1,1,1)^{\mathrm{t}}.
\end{aligned}
$$

Finally, we obtain

$$
\begin{aligned}
|\psi\rangle &= \frac{1}{16\sqrt{6}}(-1,-1,-1,-1,-1,-1,39,-1,\\
&\qquad -1,-1,-1,-1,-1,-1,-1,-1)^{\mathrm{t}},
\end{aligned}
$$

where $T = 1$. In this case, the desired pattern 0110 is obtained with the probability 0.99. We can get the searching data with high probability.

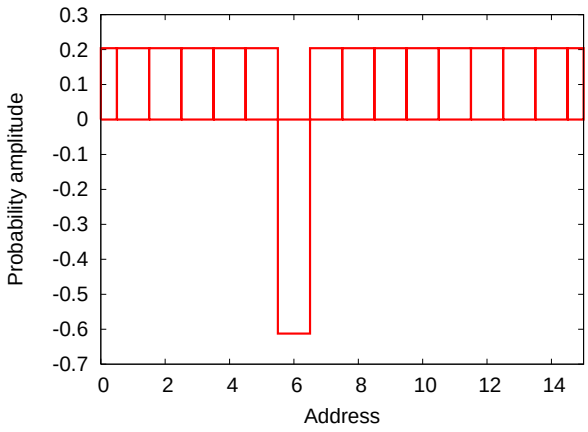
Next, supposing that searching data is $|6\rangle$ and stored data set are as follows:

$$|\psi\rangle = \frac{1}{16}(1,0,1,1,1,1,1,1,1,1,1,0,1,0,1,1)^{\mathrm{t}}.$$

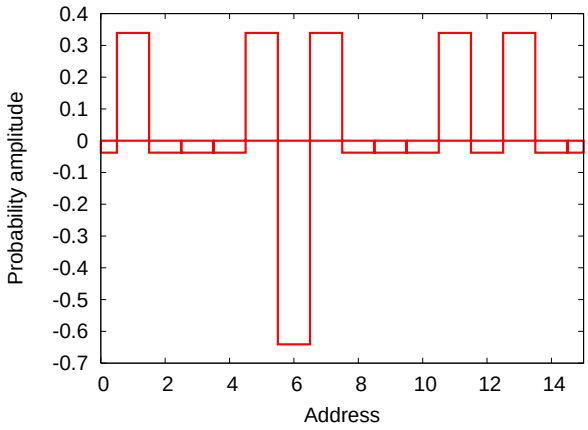
After applying I_τ, G, I_ρ and G, the following state is obtained:

$$|\psi\rangle = \frac{1}{8\sqrt{3}}(1,1,1,1,-5,1,9,1,-5,1,1,1,-5,1,1,-5)^{\mathrm{t}}.$$

As the maximum number T of repetition is 6. The desired pattern $|6\rangle$ is obtained with the probability 0.441. It is also shown that Ventura's algorithm does not always get a good result. □



(a) When the searching data is found with high probability.



(b) When the searching data is found with low probability.

Figure 3. Probability amplitude examples after applying I_ρ.

The maximum number T of repetition of Ventura's algorithm is shown as follows[3]:

$$T = \frac{\frac{\pi}{2} - \arctan\left[\frac{\bar{k}(0)}{\bar{l}(0)}\sqrt{\frac{1}{N-1}}\right]}{\arccos\left[\frac{N-2}{N}\right]},$$

$$\frac{\bar{k}(0)}{\bar{l}(0)} = \frac{\left[8(m-2)(N-m)+N^2\right](N-1)}{4(m-2)(N-m)(N-2)-N^2(m-1)}.$$

3. Some results for data searching algorithms

Grover's algorithm is effective in the case where all the data are stored in the database. However, it is not always effective in the case where all the data are not stored as shown in Example 1. Hence, Ventura has proposed an algorithm as generalized Grover's one and insisted that it is effective in all cases[3], [5]. But, in Example 2, we have shown that it is not always true. Then, let consider the reason why it happens. The cases of good results are ones that all the probability amplitudes except for searching data are identical with each other before the repeating step as shown in Fig.3. In Grover's algorithm, it is clear. Then, how is Ventura's algorithm. The following result is obtained.

Table 2. The observed probabilities of searching data for $m = \frac{16}{4} + 2$.

n	N	m	Probability
3	8	4	1.000000
4	16	6	0.990234
5	32	10	0.999894
6	64	18	0.998078
7	128	34	0.994887
8	256	66	0.999915
9	512	130	0.999415
10	1024	258	0.999472
11	2048	514	0.999997
12	4096	1026	0.999946
13	8192	2050	0.999916

1. Initial state $|\psi\rangle$
2. Repeat T times
3. $|\psi\rangle = I_\tau|\psi\rangle$
4. $|\psi\rangle = G|\psi\rangle$
5. $|\psi\rangle = I_\rho|\psi\rangle$
6. $|\psi\rangle = G|\psi\rangle$
7. Observe the system

Figure 4. Proposed algorithm.

Proposition. The maximum probability amplitude in Ventura's algorithm is obtained in the case where $m = \frac{N}{4} + 2$, where m is the number of stored data.

Proof. Let show the number of stored data with the maximum probability amplitude in Ventura's algorithm as the one of Grover's algorithm. Let $m \neq N$. In Grover's algorithm, the probability amplitude of all data becomes the same values before entering the repeating steps. In Ventura's algorithm, starting the initial data $|\psi\rangle$, the probability amplitude of all data after applying I_τ, G and I_ρ is shown in Table 1, where a is the initial probability amplitude for a stored data. The number of stored data that provides the maximum probability amplitude is obtained when the probability amplitudes of stored data after applying I_ρ are identical with ones of unstored data. From Table 1, the case is calculated as follows:

$$a - 2 \times \frac{a(m-1)-a}{N} = 2 \times \frac{2(m-1)-a}{N}$$

From the equation, we can get

$$m = \frac{N}{4} + 2.$$

□

From the proposition, it is shown that Ventura's algorithm achieves its best performance at $m = \frac{N}{4} + 2$. In Example 2, the former case is $m = \frac{16}{4} + 2 = 6$ and the latter case is $m = 13$. Table 2 shows the observed probabilities of searching data for $m = \frac{N}{4} + 2$. Next, let consider the relation between Grover's and Ventura's algorithms.

Fig.5 shows the relation between Grover's and Ventura's algorithms for the number m of stored data at $N = 2^{20}$. It is

Table 1. The probability amplitude after applying each step of I_τ, G and I_ρ.

	I_τ	G	I_ρ
When the corresponding data is stored.	a	$-a + 2 \times \frac{a(m-1)-a}{N}$	$a - 2 \times \frac{a(m-1)-a}{N}$
When the corresponding data is not stored.	0	$2 \times \frac{a(m-1)-a}{N}$	$2 \times \frac{a(m-1)-a}{N}$

shown that, Grover's and Ventura's algorithms are effective in the case where the number of stored data are large and small, respectively.

If we know the value m, we can select a better algorithm. If not, we may perform two algorithms in parallel.

Lastly, we propose a new algorithm that repeats I_τ, G, I_ρ and G as shown in Fig.4. Let us show an example for the proposed algorithm.

Example 3:

Let $n = 4$, $N = 16$. Let searching data $|\tau\rangle = |0110\rangle$ and the number of stored data $m = 8$.

$$
\begin{aligned}
|\psi\rangle &= \frac{1}{2\sqrt{2}}(1,0,1,0,1,0,1,0,1,0,1,0,1,0,1,0)^{\mathrm{t}} \\
I_\tau|\psi\rangle &= \frac{1}{2\sqrt{2}}(1,0,1,0,1,0,-1,0,1,0,1,0,1,0,1,0)^{\mathrm{t}} \\
G|\psi\rangle &= \frac{1}{8\sqrt{2}}(-1,3,-1,3,-1,3,7,3,-1,3,-1,3, \\
&\qquad -1,3,-1,3)^{\mathrm{t}} \\
I_\rho|\psi\rangle &= \frac{1}{8\sqrt{2}}(1,3,1,3,1,3,7,3,1,3,1,3,1,3,1,3)^{\mathrm{t}} \\
G|\psi\rangle &= \frac{1}{4\sqrt{2}}(1,0,1,0,1,0,5,0,1,0,1,0,1,0,1,0)^{\mathrm{t}}
\end{aligned}
$$

The following result is obtained after applying I_τ, G, I_ρ, G:

$$
\begin{aligned}
|\psi\rangle = \frac{1}{8\sqrt{2}}(&-1,0,-1,0,-1,0,11,0,-1,0,-1,0, \\
&-1,0,-1,0)^{\mathrm{t}}
\end{aligned}
$$

The desired pattern $|6\rangle$ is obtained with the probability $(\frac{11}{8\sqrt{2}})^2 \approx 0.95$. □

It has a maximum point at the half of N as shown in Fig.5 from numerical simulation. But we have not yet got its detailed analysis as the number of repeating time.

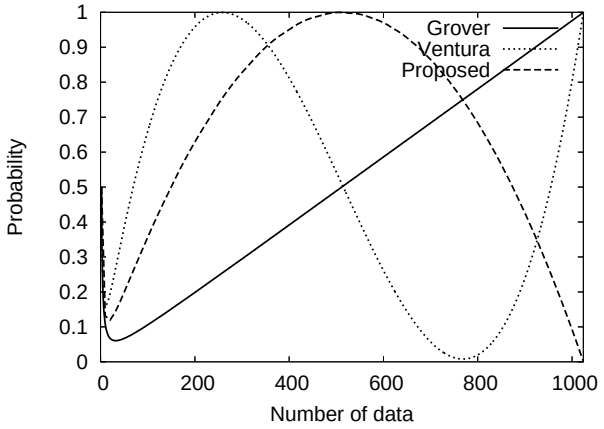


Figure 5. The relation between the observed probability of searching data and the number of stored data for $N = 1024$.

4. Conclusion

In this paper, we show some properties of quantum data search algorithms. First, we have shown that Grover's and Ventura's algorithms do not always provide the good performance for arbitrary initial distribution of stored data and they are effective in the case where the number of stored data are large and small, respectively. Further, we propose a new algorithm and a method combining these algorithms. In the future works, we will explain the effectiveness of the proposed algorithm theoretically.

References

[1] P. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer", SIAM Journal of Computing, vol.26, no.5, pp.1484-1509, 1997.

[2] L. Grover, "A Fast Quantum Mechanical Algorithm for Database Search", Proc. of 28th ACM Symp. on Theory of Computing, pp.212-219, 1996.

[3] D. Ventura and T. Martinez, "Quantum Associative Memory", Information Science, vol.124, pp.273-296, 2000.

[4] D. Biron, O. Biham, E. Biham, M. Grassl and D.A. Lidar, "Generalized Grover Search Algorithm for Arbitrary Initial Amplitude Distribution", Lecture Notes In Computer Science, Vol.1509, pp.140–147, 1998.

[5] D. Ventura and T. Martinez, "Initializing The Amplitude Distribution Of A Quantum State", Foundations of Physics Letters, Vol.12, No.6, pp.547–559, 1999.