Quantum Algorithm for Anomaly Detection of Sequences

Ming-Chao Guo, Hai-Ling Liu, Shi-Jie Pan, Wen-Min Li,* Fei Gao, Xin-Yi Huang, Su-Juan Qin, and Qiao-Yan Wen

Anomaly detection of sequences is a hot topic in data mining. Anomaly detection using piecewise aggregate approximation in the amplitude domain (called ADPAAD) is one of the widely used methods in anomaly detection of sequences. The core step in the classical algorithm for performing ADPAAD is to construct an approximate representation of the subsequence, where the elements of each subsequence are divided into several subsections according to the amplitude domain, and then the average of the subsections is computed. It is computationally expensive when processing large-scale sequences. In this paper, a quantum algorithm for ADPAAD is proposed, which can divide the subsequence elements and compute the average in parallel. The quantum algorithm can achieve polynomial speedups on the number of subsequences and the length of subsequences over its classical counterpart.

1. Introduction

Anomaly detection refers to the problem of finding patterns in data that do not conform to expected behavior. It can be divided into three modes according to the availability of data labels: supervised anomaly detection, semi-supervised anomaly detection, and unsupervised anomaly detection. Anomaly Detection using Piecewise Aggregate approximation in the Amplitude Domain (called ADPAAD) is an important method for unsupervised anomaly detection. As a widely used method in anomaly detection of sequences, it has been used in medical science, network security, finance, industrial engineering, and transportation.

The classical algorithm of ADPAAD can be divided into the following parts. (1) Division: the sequence is divided into subsequences through a sliding window. (2) Approximate representation: the elements of each subsequence are divided into several subsections according to the amplitude domain, and then the average of the subsections is computed. (3) Similarity: the similarity between subsequences is calculated according to their approximate representation. (4) Anomaly score: anomaly score for each subsequence is calculated. (5) Determination: anomalous subsequences are determined based on the comparison between their anomaly scores and predetermined threshold. Among them, obtaining the approximate representation of subsequences typically carries a linear time overhead with respect to the input size, and the running time of computing similarity depends quadratically on the number of subsequences. These are computationally expensive when processing large-scale sequences.

Quantum computing has been shown to be more computationally powerful than classical computing in solving certain problems, such as factoring integers, searching in unstructured databases, solving systems of linear and differential equations, cryptanalysis, and private queries. The combination of quantum computing and machine learning has made great progress in classification, clustering, neural networks, linear regression, association rule mining, dimensionality reduction, and quantum support vector machine. Therefore, it is worthwhile to explore quantum algorithms for anomaly detection to reduce its computational complexity.

Several works have been developed in the context of quantum computing to solve anomaly detection problems. In 2018, Liu et al. proposed a quantum kernel principal component analysis algorithm for anomaly detection. It calculates the inner product of two vectors based on the swap-test to obtain a value of proximity measure and achieves exponential speedup on the dimension of the training data set. Subsequently, Liang et al. presented a quantum anomaly detection algorithm based on density estimation. Its complexity is logarithmic in the dimension and the number of training data compared to the corresponding classical algorithm. In 2022, Guo et al. proposed a quantum algorithm for anomaly detection, which achieves exponential speedup on the number of training data points over its classical counterpart. It first regards the mean and variance of the parameters of the training data set as the inner product of two vectors, and then calculates the inner products in parallel by performing amplitude estimation. These quantum algorithms are aimed
at semi-supervised anomaly detection and determine the parameters of the training data set to achieve acceleration based on swap-test\cite{32} or amplitude estimation.\cite{35}

For the classical algorithm of ADPAAD, it is first required to divide the elements of each subsequence into several sub-segments based on the amplitude domain, and then to calculate the average value of the sub-segments. These average values cannot be directly calculated using amplitude estimate or swap-test before determining the number of elements belonging to a certain sub-segment. The published works\cite{32–34} calculate the mean value of a certain number of elements by amplitude estimation. Therefore the above quantum algorithms cannot be directly applied to ADPAAD.

In this paper, we focus on studying quantum algorithms for unsupervised anomaly detection. Specifically, we propose a quantum algorithm for ADPAAD. As shown above, the core step in the classical algorithm for performing ADPAAD is the approximate representation. To reduce computational complexity, quantum multiply-adder\cite{37,38} is used to divide the subsequence elements into several sub-sections, and amplitude amplification and estimation\cite{36} are used to calculate the average of each subsection without counting the number of elements belonging to the same sub-section. This enables both steps to be implemented in parallel. In practical application scenarios, due to the huge amount of data and the difficulty of collecting abnormal labels or normal label samples, the data is often unlabeled. However, we found that there is currently no quantum algorithm specifically proposed for unsupervised anomaly detection. It is shown that our quantum algorithm achieves polynomial speedups compared to its classical counterpart.

An outline of the paper follows. In Section 2, we review the related definitions of anomaly detection and briefly introduce the classical algorithm of ADPAAD. In Section 3, we propose a quantum algorithm for ADPAAD and analyze its complexity in detail. The conclusion is given in Section 4.

2. Review of ADPAAD

In this section, we introduce the relevant definitions for anomaly detection of sequences used in this paper and briefly review the classical algorithm of ADPAAD.\cite{2}

2.1. Definitions

For convenience, we begin with the following related definitions:\cite{2}

**Definition 1. Sequence**: A sequence $X(m) = \{x(1), x(2), \ldots, x(m)\}$ is a time series, where data elements are sorted by time, and $m$ represents the length of $X(m)$.

**Definition 2. Sliding window**: Sliding window is a user-defined window of length $n \leq m$, whose movement follows a sequence.

**Definition 3. Subsequence**: Given a sequence $X(m)$, the subsequence of length $n$ is extracted through a sliding window. If a sliding window moves $l$ steps each time, the $i$-th subsequence is expressed as:

$$X_i = \{x(1 + l(i - 1)), x(2 + l(i - 1)), \ldots, x(n + l(i - 1))\} \quad (1)$$

The $i$-th subsequence can be alternatively denoted as:

$$X_i = \{x(1), x(2), \ldots, x(n)\} \quad (2)$$

where $x(j)$ denotes the $j$-th element of the subsequence $X_i$ and $j = 1, 2, \ldots, n$.

**Definition 4. Amplitude domain**: Given a subsequence $X_i$, its amplitude domain is defined as $I_i = [L_i, H_i]$, where $L_i$ and $H_i$ represent the minimum and maximum values of the $X_i$, respectively.

**Definition 5. Subsection**: Given a subsequence $X_i$, its subsections are generated by dividing the amplitude domain $I_i$ of the subsequence $X_i$. The $t$-th subsection of the subsequence $X_i$ can be shown as follows:

$$l_t^i = [a_t^{i-1}, a_t^i], \quad t = 1, 2, \ldots, q \quad (3)$$

where $a_t^{i-1}$ and $a_t^i$ denote the lower and upper bounds of the $t$-th subsection, respectively. $q$ represents the number of subsections.

2.2. ADPAAD Algorithm

An important step of implementing ADPAAD is a piecewise aggregated approximate representation in the amplitude domain (called PAAD representation) for each subsequence, which reduces dimensionality of the subsequence and preserves its key information. In this step, $q$ subsections (as shown in Definition 5) are generated according to the amplitude domain of each subsequence $X_i$. The representation of subsequence $X_i$ is written as a vector of the mean elements in subsections. The whole procedure is depicted as follows.

(a) Divide the sequence $X(m)$ into subsequences $X_1, X_2, \ldots, X_K$ by sliding windows.

(b) Construct a PAAD representation $\bar{X}_i$ for each subsequence $X_i$ to get

$$\bar{X}_i = [\mu_1^i, \mu_2^i, \ldots, \mu_q^i]^T, \quad \mu_t^i = \frac{1}{n_t} \sum_{s_1 \in [a_t^{i-1}, a_t^i]} x(s)$$

where $i = 1, 2, \ldots, K$ and $n_t^i$ denotes the number of data points belonging to the subsection $[a_t^{i-1}, a_t^i]$ in subsequence $X_i$.

(c) Calculate the similarity between $X_i$ and $\bar{X}_i$, which is defined as follows:

$$S_\mu(X_i, \bar{X}_i) = \sqrt{\sum_{t=1}^{q} (\mu_t^i - \mu_t^i)^2}$$

(d) Compute the anomaly score of the $i$-th subsequence $X_i$ by

$$h_i = \frac{\sum_{s=1}^{K} S_\mu(X_i, X_s)}{K}$$

(e) Set a threshold $\delta$ in advance, if $h_i \geq \delta$, we mark the subsequence as an anomaly; otherwise, we judge it as normal.
The total runtime of this algorithm is $O(Kqn + K^2q)$. In the current era of big data, processing large-scale sequences will bring huge resource consumption. It is worthwhile exploring a quantum algorithm for ADPAAD to reduce its complexity.

### 3. Quantum Algorithm

In this section, we present a quantum algorithm for ADPAAD and analyze its complexity in detail.

Our quantum algorithm consists of four steps, corresponding to the four parts (b)–(e) of the classical ADPAAD algorithm. In step 1, we prepare the quantum state $\frac{1}{\sqrt{K}} \sum_{k=1}^{K} |i\rangle \sum_{t=1}^{q} \frac{1}{\sqrt{q}} \sum_{t=1}^{q} \langle t | \mu^t_i \rangle$ to obtain the PAAD representation of each subsequence; in step 2, we can get the quantum state $\frac{1}{K} \sum_{k=1}^{K} \sum_{k=1}^{K} S_p(X_k, X_k)/K^2$. In step 3, we perform amplitude estimation [36] to generate the state $\frac{1}{\sqrt{K}} \sum_{k=1}^{K} |i\rangle \sum_{t=1}^{q} \frac{1}{\sqrt{q}} \sum_{t=1}^{q} \langle t | h^p_i \rangle$. In step 4, we perform Grover’s algorithm [9] to search for anomalous subsequences satisfying $h_i \geq \delta$. The entire algorithm process is shown in Figure 1.

### 3.1. Preliminaries

In order to adapt to the quantum environment, we are going to use the following tools:

**Lemma 3.1.** (QRAM) Assume that the subsequences $\{X_k\}_{k=1}^{K}$, upper and lower bounds of subsections (that is, $a_i$ and $a_i^{-1}$, respectively) are stored in a Quantum Random Access Memory (QRAM) [39] which allows to efficiently perform the following two unitary operations in $O[\log(Kn)]$ and $O[\log(Kq)]$ time as given below:

$O_X : |i\rangle |j\rangle |0\rangle \rightarrow |i\rangle |j\rangle |x_i(j)\rangle$

$O_S : |i\rangle |t\rangle |0\rangle |0\rangle \rightarrow |i\rangle |t\rangle |a_i^t\rangle |a_i^{t-1}\rangle$ (7)

where $i = 1, 2, ..., K$, $j = 1, 2, ..., n$ and $t = 1, 2, ..., q$.

QRAM has been used for quantum state preparation in data classification [18, 19], clustering [20], neural networks [21], linear regression [22–24], association rule mining [25], dimensionality reduction [26, 27]. It’s worth noting that how to physically build QRAM is still an open question and may be a goal worthy of consideration in the future.

**Lemma 3.2.** (Inner Products Estimation [41]) Assume that there are unitaries satisfying $|i\rangle |0\rangle \rightarrow |i\rangle |v_i\rangle$, and $|j\rangle |0\rangle \rightarrow |j\rangle |w_j\rangle$, which can be performed in time $T$, the norms of $|v_i\rangle$ and $|w_j\rangle$ are known. There exists a quantum algorithm that can compute

$|i\rangle |j\rangle - \frac{1}{\sqrt{2}} ([0]|v_i\rangle + |1\rangle |w_j\rangle)|0\rangle$ (8)
with probability at least $1 - 2\delta$ for any $\delta \in (0, 1/2]$ with complexity $O\left[\frac{1}{\epsilon^2} \log \frac{1}{\epsilon} \log \log \frac{1}{\epsilon}\right]$, where $\epsilon$ is the error of $\langle \psi_l | \psi_r \rangle$.

### 3.2. Algorithm

The algorithm can be decomposed into the following four stages:

1. **Prepare the state**

   Step 1.1 Initialize the quantum state

   $$|0^\otimes K_1 \rangle \otimes |0^\otimes K_2 \rangle, |0^\otimes K_3 \rangle, |0^\otimes K_4 \rangle, |0\rangle, |0\rangle_2$$

   where the subscript numbers denote different registers.

   Step 1.2 Prepare the Hadamard gates $H^\otimes K_1$, $H^\otimes K_2$, and $H^\otimes K_3$ on the first, second, and fifth registers to obtain

   $$\frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{q}} \sum_{i=1}^{q} |i\rangle |0\rangle_4 \otimes \frac{1}{\sqrt{n}} \sum_{j=1}^{n} |j\rangle |0\rangle_5 \rangle$$

   (9)

   Step 1.3 Apply the oracle $O_X$ and $O_t$ on the third, fourth, and sixth registers, which can be seen in Equation (7), to prepare

   $$\frac{1}{\sqrt{Kq}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{q}} \sum_{i=1}^{q} |i\rangle |0\rangle_4 \otimes \frac{1}{\sqrt{n}} \sum_{j=1}^{n} |j\rangle |x(j)\rangle_5 \rangle$$

   (10)

   Step 1.4 Prepare the quantum multiply-adder (QMA) gate on the third, fourth, sixth registers, and undo the redundant registers to create

   $$\frac{1}{\sqrt{Kd}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{q}} \sum_{i=1}^{q} |i\rangle |0\rangle_4 \otimes \frac{1}{\sqrt{n}} \sum_{j=1}^{n} |j\rangle |x(j)\rangle_5 \rangle$$

   (11)

   where $\rho(j) = (x(j) - a^0(i)) |x(j)\rangle - a^1(i) \rangle$. If $x(j) \in [a^0(i), a^1(i)]$, then $\rho(j) \leq 0$, otherwise, $\rho(j) > 0$. Equation (12) can be rewritten as:

   $$\frac{1}{\sqrt{Kd}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{q}} \sum_{i=1}^{q} |i\rangle |0\rangle_4 \otimes \frac{1}{\sqrt{n}} \sum_{j=1}^{n} |j\rangle |x(j)\rangle_5 \rangle$$

   (12)

   where $\rho(j) = (x(j) - a^0(i)) |x(j)\rangle - a^1(i) \rangle$. If $x(j) \in [a^0(i), a^1(i)]$, then $\rho(j) \leq 0$, otherwise, $\rho(j) > 0$. Equation (12) can be rewritten as:

   $$\frac{1}{\sqrt{Kd}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{q}} \sum_{i=1}^{q} |i\rangle |0\rangle_4 \otimes \frac{1}{\sqrt{n}} \sum_{j=1}^{n} |j\rangle |x(j)\rangle_5 \rangle$$

   (13)

   $$\times \left[ \sum_{x(j) \in [a^0(i), a^1(i)]} |j\rangle |x(j)\rangle \rangle \right]_{676}$$

   Step 1.5 Execute the amplitude amplification to magnify the part corresponding to $\rho(j) > 0$, we obtain

   $$\frac{1}{\sqrt{Kd}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{q}} \sum_{i=1}^{q} |i\rangle |0\rangle_4 \otimes \frac{1}{\sqrt{n}} \sum_{j=1}^{n} |j\rangle |x(j)\rangle_5 \rangle$$

   (14)

   where $|\Phi^+\rangle = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} |j\rangle |x(j)\rangle_5 \rangle$. $n^1_i$ denotes the number of data points belonging to the subsequence $F_i$ in subsequence $X_i$ and $n^1_i > 0$ is assumed. $\sqrt{p}$ represents the amplitude of $|\Phi^+\rangle$ and $|\Phi^+\rangle$ is the quantum state corresponding to $\rho(j) > 0$. For simplicity, we assume $p = 1$, then Equation (14) can be rewritten as

   $$\frac{1}{\sqrt{Kd}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{q}} \sum_{i=1}^{q} |i\rangle \frac{1}{\sqrt{n}} \sum_{j=1}^{n} |j\rangle |x(j)\rangle_5 \rangle$$

   (15)

   The general case $0 < p < 1$ is analyzed in detail in Appendix A.

   Step 1.6 Append one qubit, then rotate it from $|0\rangle$ to $\frac{\sqrt{\frac{2}{C}} |0\rangle + \sqrt{\frac{2}{C} - 1} |1\rangle}{\sqrt{2}}$ controlled on $|x(j)\rangle$. If $|x(j)\rangle \in \langle 10, 42, 43 \rangle$ discard the sixth and seventh registers, we can obtain

   $$\frac{1}{\sqrt{Kd}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{q}} \sum_{i=1}^{q} |i\rangle \frac{1}{\sqrt{n}} \sum_{j=1}^{n} |j\rangle \left( \frac{\sqrt{\frac{2}{C}} |0\rangle + \sqrt{\frac{2}{C} - 1} |1\rangle}{\sqrt{2}} \right) \rangle$$

   (16)

   where $C = \max |x(j)\rangle$. The state $|\psi_i\rangle$ can be rewritten as

   $$|\psi_i\rangle = |x(i)\rangle + |0\rangle \left( \cos \theta_i |\psi_i\rangle + \sin \theta_i |\psi_i\rangle \right),$$

   where $|\psi_i\rangle$ and $|\psi_i\rangle$ represent the normalized quantum states of $\frac{\sqrt{\frac{2}{C}} |0\rangle + \sqrt{\frac{2}{C} - 1} |1\rangle}{\sqrt{2}}$ and $\frac{\sqrt{\frac{2}{C}} |0\rangle + \sqrt{\frac{2}{C} - 1} |1\rangle}{\sqrt{2}}$, respectively. It can be easily calculated: $\sin^2 \theta_i = \frac{1}{n} \sum x(i) |x(i)\rangle |x(i)\rangle$. We define $Q_l = -A_l S_x A_l^T S_x$, where $A_l : |0\rangle_{58} \rightarrow |\psi_i\rangle$, $S_x = I - 2|0\rangle_{58} \langle 0|_{58}, S_y = I - 2|0\rangle_{58} \langle 0|_{58}$. The $Q_l$ is performed on the state $|\psi_i\rangle$ to get

   $$(Q_l |\psi_i\rangle = |x(2l + 1)\rangle |\psi_i\rangle + 2 \cos \theta_i |\psi_i\rangle)$$

   (17)

   for any $l \in N$. $Q_l$ acts as a rotation in 2-dimensional space $\text{Span} \{|\psi_i\rangle, |\psi_i\rangle\}$, and it has two eigenvalues $e^{\pm i \theta_i}$, with the eigenstates $|\psi_i\rangle$, where $|\psi_i\rangle = \frac{1}{\sqrt{2}} (|\psi_i\rangle \pm i |\psi_i\rangle)$, and $\theta_i = \sqrt{2}$. Step 1.7 Add an ancilla register, then perform amplitude estimation of $Q_l$ on $|\psi_i\rangle$ to generate

   $$\frac{1}{\sqrt{Kd}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{q}} \sum_{i=1}^{q} |i\rangle |0\rangle_4 \otimes \frac{\theta_i}{\pi} \rangle$$

   (18)

   Step 1.8 Compute $|\mu_i\rangle = \left| \cos \theta_i \left[ \langle x(i) | \right] \right| \langle \psi_i | \psi_i \rangle \rangle$ via the QMA and sine gate, and compute redundant registers to get

   $$\frac{1}{\sqrt{Kd}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{q}} \sum_{i=1}^{q} |i\rangle |0\rangle_4 \otimes \frac{\theta_i}{\pi} \rangle$$

   (19)

   2. **Prepare the state**

   $$\sum_{i=1}^{K} |i\rangle \langle k| S_x (X_i, X_k))$$

   We prepare the quantum state corresponding to the similarity between subsequences, and the specific steps are shown as follows.

   Step 2.1 Repeat the operators of step 1 to get

   $$\frac{1}{\sqrt{Kd}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{q}} \sum_{i=1}^{q} |i\rangle \langle j| \langle \mu_i | \mu_i \rangle_5$$

   (20)
Step 2.2 Add an ancilla register and perform the QMA gate, uncompute the fifth register to obtain

\[ \frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle |k\rangle \frac{1}{\sqrt{q}} \sum_{l=1}^{q} |l\rangle |\mu^l - \mu_i^l\rangle \]  

(21)

Step 2.3 Append an ancilla register, then apply controlled rotation operator, uncompute the fourth register we can get

\[ \frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle |k\rangle \frac{1}{\sqrt{q}} \sum_{l=1}^{q} |l\rangle \left[ \frac{\mu^l - \mu_i^l}{2C} |0\rangle + \sqrt{1 - \left( \frac{\mu^l - \mu_i^l}{2C} \right)^2} |1\rangle \right] \]  

(22)

where \( C = \max |x_j\rangle |j\rangle \).

Step 2.4 Perform the amplitude estimation and compute \( |\tilde{S}_k(X, X_i)\rangle = |\sin(\xi_i,\mu\rangle) \) via the QMA and sine gate (similar to steps (1.7)-(1.8)) to generate

\[ \frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle |k\rangle |\tilde{S}_k(X, X_i)\rangle \]  

(23)

where \( \tilde{S}_k(X, X_i) = \frac{1}{\sqrt{q}} \sum_{l=1}^{q} (e^{i(\xi_l/\mu\langle 0|)}/ C) \).

3. Prepare the quantum state \( \frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle |h_i\rangle \).

To demonstrate the process of obtaining anomaly scores for all subsequences, we describe it in the three parts: step 3.1 prepares the quantum state \( \frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle |\tilde{S}_k(X, X_i)\rangle \); step 3.2 generates the quantum state \( |\sum_{i=1}^{K} \frac{\tilde{S}_k(X, X_i)}{\sqrt{K}}\rangle \); step 3.3 obtains the quantum state \( \frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle |h_i\rangle \). The detailed process is as follows.

Step 3.1(a) Add two ancilla registers, then apply H gate on Equation (23) to create

\[ \frac{1}{\sqrt{2}} \left[ |0\rangle |1\rangle + |1\rangle |0\rangle \right] \sum_{i=1}^{K} |i\rangle |\tilde{S}_k(X, X_i)\rangle |0\rangle \]  

(24)

Step 3.1(b) Perform unitary operator \( I \otimes |0\rangle |0\rangle \otimes U + I \otimes |1\rangle |1\rangle \otimes I \), where \( U \) is a controlled rotation operator which rotates \( |0\rangle \rightarrow |\xi_i,\mu\rangle |0\rangle + \sqrt{1 - \xi_i^2} |1\rangle \) conditioned on \( |\tilde{S}_k(X, X_i)\rangle \), where \( \xi_i,\mu = \tilde{S}_k(X, X_i) \). Undo the fourth register to obtain

\[ \frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{2}} \left[ |0\rangle \sum_{k=1}^{K} |k\rangle (|\xi_i,\mu\rangle |0\rangle + \sqrt{1 - \xi_i^2} |1\rangle \rangle + |1\rangle \sum_{k=1}^{K} |k\rangle |0\rangle \right] \]  

(25)

The above equation can be rewritten as

\[ \frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{2}} (|0\rangle |\phi_i\rangle + |1\rangle |\rho\rangle) \]  

(26)

where \( |\phi_i\rangle = \frac{1}{\sqrt{K}} \sum_{k=1}^{K} |k\rangle |\xi_i,\mu\rangle |0\rangle + \sqrt{1 - \xi_i^2} |1\rangle \rangle \) and \( |\rho\rangle = \frac{1}{\sqrt{K}} \sum_{k=1}^{K} |k\rangle |0\rangle \).

Step 3.1(c) Add an ancilla register, then apply the Inner Products Estimation (Lemma 3.2) to generate

\[ \frac{1}{\sqrt{2}} \sum_{i=1}^{K} |i\rangle \frac{1}{\sqrt{K}} (|0\rangle |\phi_i\rangle + |1\rangle |\rho\rangle) \sum_{k=1}^{K} \frac{\tilde{S}_k(X, X_i)}{\sqrt{K}} \]  

(27)

where \( \sum_{i=1}^{K} \frac{\tilde{S}_k(X, X_i)}{\sqrt{K}} = |\phi_i\rangle \).

Step 3.2(a) The Equation (25) is equivalent to

\[ \frac{1}{\sqrt{2}} \left[ |0\rangle \sum_{k=1}^{K} |i\rangle (|\xi_i,\mu\rangle |0\rangle + \sqrt{1 - \xi_i^2} |1\rangle \rangle + |1\rangle \sum_{k=1}^{K} |i\rangle |k\rangle |0\rangle \right] \]  

(28)

where \( |\Psi\rangle = \frac{1}{\sqrt{K}} \sum_{i=1}^{K} (|i\rangle |\xi_i,\mu\rangle |0\rangle + \sqrt{1 - \xi_i^2} |1\rangle \rangle |\Phi\rangle = \frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle |k\rangle |0\rangle \) and \( \xi_i,\mu = \tilde{S}_k(X, X_i) \).

Step 3.2(b) Append an ancilla register, then perform the Inner Products Estimation and discard redundant registers to get

\[ |\langle \Psi|\Phi\rangle\rangle = \left| \sum_{i=1}^{K} \sum_{k=1}^{K} \frac{\tilde{S}_k(X, X_i)}{K^2} \right| \]  

(29)

According to Equation (27), we can obtain the quantum state

\[ \frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle \sum_{k=1}^{K} \frac{\tilde{S}_k(X, X_i)}{K} \left| \frac{1}{\sqrt{K}} \sum_{k=1}^{K} \tilde{S}_k(X, X_i) \right|^2 \]  

(30)

Step 3.3 Apply the QMA gate on Equation (30) to obtain

\[ \frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle \sum_{k=1}^{K} \tilde{S}_k(X, X_i) \left| \frac{1}{\sqrt{K}} \sum_{k=1}^{K} \tilde{S}_k(X, X_i) \right|^2 = \frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle |h_i\rangle \]  

(31)

where \( \frac{1}{\sqrt{K}} \sum_{k=1}^{K} \tilde{S}_k(X, X_i) = \frac{1}{\sqrt{K}} \sum_{k=1}^{K} \tilde{S}_k(X, X_i) = h_i \).

4. Grover’s algorithm\(^{(9)}\) is applied to search all indices \( i \) of the abnormal subsequences that satisfy \( h_i \geq \delta \).

3.3. Complexity and Error Analysis

(i) The complexity of step 1 is \( O(\sqrt{n\log(K)n}) \), the specific analysis is as follows:

In steps 1.2–1.4, \( H \) and QMA gates, the oracle \( O_X \) and \( O_i \) are executed with complexity \( O(\log(K)n) \). In step 1.5, the complexity of amplitude amplification is \( O(\sqrt{n\log(K)n}) \). In step 1.6, it contains a controlled rotation operator with complexity \( O(1) \). In step 1.7, the amplitude estimation block needs \( O(1/\varepsilon_1) \) applications of \( Q_1 \) to achieve error \( \varepsilon_1 \), and the complexity of performing unitary operator \( Q_1 \) is \( O(\sqrt{n\log(K)n}) \). In step 1.8, it takes QMA and sine gates with complexity \( O(\log(1/\varepsilon_1)) \), which is smaller than \( O(1/\varepsilon_1) \), the complexity of these gates can be omitted.\(^{(38)}\)
Now, we analyze the error of \( \mu' \), which mainly comes from the amplitude estimation in step 1.7,
\[
|\vec{\mu'} - \mu'| = C|\sin\theta_i - \sin\theta_j| \leq |\sin(\theta_i - \theta_j)| \leq C \cdot |\theta_i - \theta_j|
\]
\[
\leq CE_1
\]
(32)
where \( \vec{\mu'} \) represents the estimate of \( \mu' \) and \( |\theta_i - \theta_j| \leq \epsilon \), comes from step 1.8. In this paper we use \( \epsilon \) to denote the estimated value.

(ii) The time complexity of step 2 is \( O(\sqrt{\log(Kqn)}) \), which mainly stems from the amplitude estimation of step 2.4.
In step 2.1, the operations of step 1 are performed with complexity \( O(\sqrt{\log(Kqn)}) \). In steps 2.2–2.3, QMA gate and controlled rotation are performed with complexity \( O(\log(1/\epsilon_2)) \), which is smaller than \( O(1/\epsilon_1) \) and can be ignored. In step 2.4, the amplitude estimation costs \( O(\sqrt{\log(Kqn)}) \) time to ensure the error \( \epsilon_2 \). The error of step 2.4 is \( |\hat{S}_2(X_i, X_k) - \hat{S}_2(X_i, X_k)| \leq |\hat{\theta}_{i,k} - \alpha_{i,k}| \leq \epsilon_2 \). We analyze the error of \( \hat{S}_2(X_i, X_k) \) as follow:
\[
|\hat{S}_2(X_i, X_k) - \hat{S}_2(X_i, X_k)| = |\hat{S}_2(X_i, X_k) - \hat{S}_2(X_i, X_k) + \hat{S}_2(X_i, X_k)|
\]
\[
\leq \epsilon_2 \leq \sqrt{2}\epsilon_2
\]
\[
\leq 2\epsilon_2
\]
(33)
where \( \hat{S}_2(X_i, X_k) = \frac{1}{\sqrt{2}} \sum_{l=1}^{q} \left( \frac{|\mu'_l - \mu_k'|}{2C} \right)^2 \). We assume that at least half of the values of \( \left( \frac{|\mu'_l - \mu_k'|}{2C} \right)^2 \) are greater than a constant \( E \), that is
\[
\hat{S}_2(X_i, X_k) \geq \sqrt{2}\frac{E}{2}
\]
\[
\geq \frac{1}{\sqrt{2}} \sum_{l=1}^{q} \left( \frac{|\mu'_l - \mu_k'|}{2C} \right)^2 \geq \frac{\sqrt{2}E}{2}
\]
(34)
The second term of Equation (33) is as follows:
\[
|\hat{S}_2(X_i, X_k) - \hat{S}_2(X_i, X_k)| = |\hat{S}_2(X_i, X_k) - \hat{S}_2(X_i, X_k) + \hat{S}_2(X_i, X_k)|
\]
\[
\leq \epsilon_2 \leq \sqrt{2}\epsilon_2
\]
\[
\leq 2\epsilon_2
\]
(35)
Therefore, we get \( \hat{S}_2(X_i, X_k) \) with error \( \epsilon_2 + \frac{2\epsilon_2}{E} \).

(iii) The complexity of step 3 is \( O(\sqrt{\log(Kqn)}) \), which mainly coming from the Inner Product Estimation (Lemma 3.2) and amplitude estimation of steps 3.1 and 3.2.
In steps 3.1(a)–3.1(b), \( H \) gate and unitary operation \( I \otimes |0\rangle\langle 0| \otimes I \otimes |1\rangle\langle 1| \otimes I \) are performed with complexity \( O(\log(1/\epsilon_3)) \). According to Lemma 3.2, the state \( |\phi_1|\rho\rangle = \frac{1}{K} \sum_{k=1}^{K} |\mu'_k - \mu_k'| \geq \frac{1}{K} \sum_{k=1}^{K} |\mu'_k| \geq \frac{1}{K} \sum_{k=1}^{K} \frac{1}{2C} \]
\[
\leq \epsilon_3
\]
(36)
Now, we analyze the errors of \( \hat{S}_p(X_i, X_k)/K \) and \( \sum_{k=1}^{K} \hat{S}_p(X_i, X_k)/K \) as follow:
\[
|\phi_1|\rho\rangle - \frac{1}{K} \sum_{k=1}^{K} \frac{1}{\sqrt{q}} \sum_{l=1}^{q} \left( \frac{|\mu'_l - \mu_k'|}{2C} \right)^2
\]
\[
= |\phi_1|\rho\rangle - |\phi_1|\rho\rangle + |\phi_1|\rho\rangle - \frac{1}{K} \sum_{k=1}^{K} \sum_{l=1}^{q} \frac{1}{2C} \]
\[
\leq \epsilon_3 + |\phi_1|\rho\rangle - \frac{1}{K} \sum_{k=1}^{K} \sum_{l=1}^{q} \frac{1}{2C} \]
\[
\leq \epsilon_3 + \epsilon_2 + \frac{2\epsilon_2}{E}
\]
(36)
where \( \langle \phi | \rho \rangle = \sum_{i=1}^{K} \tilde{S}_p(X_i, X_k) / K \).

\[
\left| \langle \tilde{\Psi} | \Phi \rangle - \frac{1}{K^2} \sum_{i=1}^{K} \sqrt{q_i} \sum_{j=1}^{q_i} \left( \frac{\mu_i - \mu_j}{2C} \right)^2 \right| \leq \varepsilon_1 + \varepsilon_2 + \frac{2\varepsilon_1}{E}
\]

Finally, we analyze the error of \( h_i \), as follows:

\[
\left| \hat{h}_i - h_i \right| = \left| \frac{\langle \phi | \tilde{\rho} \rangle}{\langle \tilde{\Psi} | \Phi \rangle} - \frac{1}{K^2} \sum_{i=1}^{K} \sqrt{q_i} \sum_{j=1}^{q_i} \left( \frac{\mu_i - \mu_j}{2C} \right)^2 \right| \leq \frac{(\varepsilon_1 + \varepsilon_2)}{E} + \frac{2\varepsilon_1}{E^2}
\]  

If \( \varepsilon_1 = \frac{\varepsilon}{6} \), \( \varepsilon_2 = \frac{\varepsilon}{3} \), \( \varepsilon_3 = \frac{\varepsilon}{6} \), and \( \varepsilon_4 = \varepsilon \), we can get \( \left| \hat{h}_i - h_i \right| \leq \varepsilon \). That is, we can obtain an \( \varepsilon \)-approximate of the state \( \sum_{i=1}^{K} \langle i | h_i \rangle \) with complexity \( \tilde{O}(\sqrt{K^2 \log(K^p)}) \) in step 3.

The complexity of executing Grover’s algorithm to obtain all abnormal subsequences is \( \tilde{O}(\sqrt{KT \frac{p \log(Kp)}{\varepsilon^3}}) \), where \( T \) is the number of abnormal subsequences.

We know that \( E = O(1) \), and in general the number of abnormal subsequences is much smaller than the number \( K \) of subsequences, the overall runtime will be \( O(\sqrt{\frac{K \log(Kp)}{\varepsilon^3}}) \). Our quantum algorithm achieves polynomial speedup compared to its classical counterpart.

4. Conclusion

In practical application scenarios of anomaly detection, due to the huge amount of data and the difficulty of collecting abnormal labels or normal label samples, the data is often unlabeled. This requires unsupervised anomaly detection to identify anomalies. The cost of performing classical ADPAAD is too much when dealing with large-scale sequences. Therefore, in this paper we proposed a quantum algorithm for ADPAAD, which achieves polynomial speedup compared to its classical counterpart.

Furthermore, our proposed quantum algorithm for the PAAD representation of each subsequence can be reused as a subroutine for other quantum dimensionality reduction algorithms. Our approach opens a new avenue for the quantum algorithm of unsupervised anomaly detection. Designing more quantum algorithms for unsupervised anomaly detection is a goal worth considering in the future. We hope our algorithm can inspire more efficient quantum anomaly detection algorithms.

Appendix A: Detailed Analysis of the General Case of Step 1.5

In the general case \( 0 < p < 1 \) of step 1.5, according to ref. [40], we perform amplitude amplification to obtain the quantum state

\[
\frac{1}{\sqrt{K}} \sum_{i=1}^{K} | i \rangle \frac{1}{\sqrt{q}} \sum_{i=1}^{q} | t \rangle (\sqrt{p}|\Phi^i \rangle + \sqrt{1-p}|(\Phi^i)\rangle)
\]

where \( |\Phi^i \rangle = \frac{1}{\sqrt{q}} \sum_{j=1}^{q} | j \rangle | c_i \rangle | (\phi^i) \rangle \), \( \sqrt{p} \) represents the amplitude of \( |\Phi^i \rangle \) and \( 0 < p < 1 \), \( |(\Phi^i)\rangle \) is the quantum state corresponding to \( p_i \). Then we perform a controlled operator in step 1.6, which satisfies that when \( p_i \leq 0 \) in the seventh register, a controlled rotation is performed, and when \( p_i > 0 \), an XOR gate is performed on the ancilla register. We can get

\[
\frac{1}{\sqrt{qK}} \sum_{i=1}^{K} | i \rangle \sum_{i=1}^{q} | t \rangle \times \left[ \sqrt{p}|\Phi^i \rangle \left( \sqrt{\frac{x_i}{C}} |0 \rangle + \sqrt{1-\frac{x_i}{C}} |1 \rangle \right) + \sqrt{1-p}|(\Phi^i)\rangle |1 \rangle \right] = \frac{1}{\sqrt{qK}} \sum_{i=1}^{K} \sum_{j=1}^{q} | j \rangle | 2 | (\psi^j) \rangle
\]

The \( |\psi^j \rangle \) can be rewritten as \( \sin \theta^j |(\psi^j)\rangle + \cos \theta^j |(\psi^j)\rangle \), where \( |(\psi^j)\rangle \) denotes the normalized quantum state of \( \{|\Phi^i \rangle \sqrt{\frac{x_i}{C}} |0 \rangle + |(\Phi^i)\rangle |1 \rangle \} \) is the quantum state that is orthogonal to \( |(\psi^j)\rangle \). We can use amplitude estimation to estimate \( \sin^2 \theta^j = \frac{p_i}{2} \sum_{j \in E} \frac{k_i}{2} p_i = \frac{1}{2} \frac{p_i}{C} \).

By performing the operators of steps 1.7–1.8, we can obtain

\[
\frac{1}{\sqrt{qK}} \sum_{i=1}^{K} \sum_{j=1}^{q} | j \rangle | \tilde{\rho}^j \rangle = \frac{1}{\sqrt{qK}} \sum_{i=1}^{K} \sum_{j=1}^{q} | j \rangle | p \cdot (\rho_i) \rangle
\]

where \( \tilde{\rho}^j = p \cdot \mu_i^j = C \cdot \sin^2 \theta^j \).

Our analysis of the general case of step 1.5 does not affect the subsequent process and final result of the quantum algorithm, that is, we can perform the operators of steps 2 and 3 to get

\[
\frac{1}{\sqrt{K}} \sum_{i=1}^{K} \sum_{j=1}^{q} \left| \frac{1}{\sqrt{K}} \sum_{i=1}^{K} | i \rangle | j \rangle \langle \tilde{S}_p(X_i, X_j) \rangle \right| = \frac{1}{\sqrt{K}} \sum_{i=1}^{K} \sum_{j=1}^{q} | i \rangle | h_i \rangle
\]

\[
\tilde{S}_p(X_i, X_j) = \frac{1}{\sqrt{K}} \sum_{i=1}^{K} \left( \frac{\mu_i - \mu_j}{2C} \right)^2 = \frac{p_i}{\sqrt{K}} \sum_{j=1}^{q} \left( \frac{\mu_i - \mu_j}{2C} \right)^2
\]
\[
\frac{1}{K} \sum_{k=1}^{K} \mathcal{S}_F (X_i, X_k) = \frac{1}{\sqrt{K} \sum_{i=1}^{K} \sum_{k=1}^{K} \mathcal{S}_F (X_i, X_k)} \sqrt{\frac{1}{K} \sum_{i=1}^{K} \left( \frac{\mu^2(X_i) - \mu^2(X_k)}{2\varepsilon} \right)^2} = \frac{\sum_{k=1}^{K} \mathcal{S}_F (X_i, X_k) / K}{\sum_{i,k=1}^{K} \mathcal{S}_F (X_i, X_k) / K^2} = h_i
\]

Therefore, for the general case of step 1.5, we can still get anomaly score \( h_i \) of the subsequences.

Acknowledgements

The authors thank Di Zhang, Yongmei Li, Jing Li, and Zhenqiang Li for useful discussions on the subject. This work is supported by NSFC (Grant Nos. 61976024, 61972048, 62072051).

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Keywords

anomaly detection, data mining, polynomial speedup, quantum algorithm

Received: March 28, 2023
Revised: May 12, 2023
Published online: June 25, 2023

[1] V. Chandola, A. Banerjee, V. Kumar, A. Valaba, ACM Comput. Surveys 2009, 41, 1.
[2] H. R. Ren, X. J. Liao, Z. W. Li, A. Al-Ahmari, Appl. Intell. 2017, 48, 1097.
[3] D. K. Tewatia, R. P. Tolakanahalli, B. R. Paliwal, Phys. Med. Biol. 2011, 56, 2161.
[4] J. Viinikka, H. Debar, L. Mé, A. Lehikoinen, M. Tarvainen, Inf. Fusion 2009, 10, 312.
[5] P. C. Chang, C. Y. Fan, J. L. Lin, Expert Syst. Appl. 2011, 38, 6070.
[6] M. Avazbeigi, S. Doulabi, B. Karimi, Expert Syst. Appl. 2010, 37, 5630.
[7] M. Lippi, M. Bertini, P. Frasconi, IEEE Trans. Intell. Transp. Syst. 2013, 14, 871.
[8] P. W. Shor, IEEE Computer Society Press, Los Alamitos, CA, 1994, pp. 124–134.
[9] L. K. Grover, in Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing, ACM, New York 1996, pp. 212–219.
[10] A. W. Harrow, A. Hassidim, S. Lloyd, Phys. Rev. Lett. 2009, 103, 150502.
[11] L. C. Chan, C. H. Yu, S. J. Pan, F. Gao, Q. Y. Wen, S. J. Qin, Phys. Rev. A 2018, 97, 062322.
[12] H. L. Liu, Y. S. Wu, L. C. Wan, S. J. Pan, F. Gao, S. J. Qin, Q. Y. Wen, Phys. Rev. A 2021, 104, 022418.
[13] Z. Q. Li, B. B. Cai, H. W. Sun, H. L. Liu, L. C. Wan, S. J. Qin, Q. Y. Wen, F. Gao, Sci. China Phys. Mech. Astron. 2022, 65, 290311.
[14] X. Y. Dong, Z. Li, X. Y Wang, Sci. China Inf. Sci. 2019, 62, 1.
[15] C. Y. Wei, X. Q. Cai, T. Y. Wang, S. J. Qin, F. Gao, Q. Y. Wen, IEEE J. Sel. Areas Commun. 2020, 38, 517.
[16] F. Gao, S. J. Qin, W. Huang, Q. Y. Wen, Sci. China-Phys. Mech. Astron. 2019, 62, 070301.
[17] V. Giovannetti, S. Lloyd, L. Maccone, Phys. Rev. Lett. 2008, 100, 230502.
[18] S. Lloyd, M. Mohseni, P. Rebentrost, arXiv:1307.0411, 2013.
[19] N. Wiebe, D. Braun, S. Lloyd, Phys. Rev. Lett. 2012, 109, 050505.
[20] S. C. Morampudi, B. Hsu, S. L. Sondhi, R. Moessner, Phys. Rev. A 2017, 96, 042303.
[21] P. Rebentrost, T. R. Bromley, C. Weedbrook, S. Lloyd, Phys. Rev. A 2018, 98, 042308.
[22] G. M. Wang, Phys. Rev. A 2017, 96, 012335.
[23] C. H. Yu, F. Gao, Q. Y. Wen, IEEE Trans. Knowl. Data Eng. 2019, 33, 858.
[24] C. H. Yu, F. Gao, C. Liu, D. Huynh, M. Reynolds, J. Wang, Phys. Rev. A 2019, 99, 022301.
[25] C. H. Yu, F. Gao, L. Wang, Q. Y. Wen, Phys. Rev. A 2016, 94, 042311.
[26] I. Cong, L. Duan, New J. Phys. 2016, 18, 073011.
[27] S. Lloyd, M. Mohseni, P. Rebentrost, Nat. Phys. 2014, 10, 631.
[28] S. J. Pan, L. C. Wan, H. L. Liu, F. Gao, S. J. Qin, Q. Y. Wen, Phys. Rev. A 2020, 102, 052402.
[29] C. H. Yu, F. Gao, S. Lin, J. Wang, Journal of Quantum Information Processing 2019, 18, 1.
[30] P. Rebentrost, M. Mohseni, S. Lloyd, Phys. Rev. Lett. 2014, 113, 130503.
[31] Z. K. Ye, L. Z. Li, H. Z. Sui, Y. Y. Wang, Sci. China Info. Sci. 2020, 63, 189501.
[32] N. Liu, P. Rebentrost, Phys. Rev. A 2018, 97, 042315.
[33] H. Buhrman, R. Cleve, J. Watrous, R. D. Wolf, Phys. Rev. Lett. 2001, 87, 167902.
[34] J. M. Liang, S. Q. Shen, M. Li, L. Li, Phys. Rev. A 2019, 99, 052310.
[35] M. C. Guo, H. L. Liu, W. M. Li, F. Gao, S. J. Qin, Q. Y. Wen, Phys. A: Stat. Mech. Appl. 2022, 604, 127936.
[36] G. Brassard, P. Hoyer, M. Mosca, Contemp. Math. 2002, 305, 53.
[37] L. Ruiz-Perez, J. C. Garcia-Escartín, Quantum Information Process. 2017, 16, 152.
[38] S. S. Zhou, T. Loke, J. A. Izacar, J. B. Wang, Quantum Information Process. 2017, 16, 82.
[39] V. Giovannetti, S. Lloyd, L. Maccone, Phys. Rev. Lett. 2008, 100, 160501.
[40] S. J. Pan, L. C. Wan, H. L. L, Y. S. W, S. J. Qin, Q. Y. Wen, F. Gao, Chin. Phys. B 2022, 31, 060304.
[41] K. Iordanis, J. Landman, A. Luongo, A. Prakash, in Advances in Neural Information Processing Systems, Vol. 32, CRC Press, Boca Raton 2019, p. 32.
[42] B. J. Duan, J. B. Yuan, Y. Liu, D. Li, Phys. Rev. A 2018, 98, 012308.
[43] K. Mitarai, M. Kitagawa, K. Fujii, Phys. Rev. A 2019, 99, 012301.