mathematics

IMPACT FACTOR 2.592

CITESCORE 3.5

Article

Denial-of-Service Attack Defense Strategy for Continuous Variable Quantum Key Distribution via Deep Learning

Wenhao Yin, Yuhan Zhou and Duan Huang

MDPI

https://doi.org/10.3390/math11122681

mathematics

Article

Denial-of-Service Attack Defense Strategy for Continuous Variable Quantum Key Distribution via Deep Learning

Wenhao Yin [1], Yuhan Zhou [2] and Duan Huang [1,*]

[1] School of Computer Science and Engineering, Central South University, Changsha 410083, China; yinwenhao@csu.edu.cn
[2] School of Mechanical and Manufacturing Engineering, The University of New South Wales, Sydney, NSW 2052, Australia; z5141747@zmail.unsw.edu.au
* Correspondence: duanhuang@csu.edu.cn

Abstract: In the practical Continuous Variable Quantum Key Distribution (CVQKD) system, there is a large gap between the ideal theoretical model and the actual physical system. There are still some inevitable flaws, which give quantum hackers the opportunity to manipulate the channel in complex communication environments and launch Denial of Service attacks on the quantum channel. Therefore, a DoS attack-aware defense scheme for the CVQKD system based on convolutional neural networks (CNN) is proposed. The simulation results show that the proposed model can effectively detect DoS attacks launched by quantum hackers in CVQKD system in a complex communication environment, and the model has strong robustness due to the addition of the attention mechanism module. In addition, multiple sets of comparative experiments show that compared with the existing artificial neural network model, the CNN-based model has higher accuracy and stability.

Keywords: continuous variable quantum key distribution; denial-of-service attacks; deep learning model; attention mechanism

MSC: 81P94; 81P45

check for updates

Citation: Yin, W.; Zhou, Y.; Huang, D. Denial-of-Service Attack Defense Strategy for Continuous Variable Quantum Key Distribution via Deep Learning. *Mathematics* **2023**, *11*, 2681. https://doi.org/10.3390/math11122681

Academic Editor: Antanas Cenys

Received: 17 May 2023
Revised: 6 June 2023
Accepted: 7 June 2023
Published: 13 June 2023

1. Introduction

In recent years, it has been observed that quantum key distribution (QKD) exhibits excellent information-theoretical security, and is employed for the secure transmission of cryptographic keys between authorized sender (Alice) and receiver (Bob). QKD is not derived from mathematical principles but is based on fundamental laws of quantum physics [1]. The theoretical unconditional security of QKD is guaranteed by the Heisenberg uncertainty principle [2] and the no-cloning principle of quantum states [3], which also ensures the detectability of Eve. Currently, discrete-variable (DV)QKD [4] and continuous-variable (CV)QKD [5–7] are the two main models of QKD systems. Compared to the DVQKD system, the CVQKD system has many advantages, such as its ease of implementation, compatibility with existing optical communication systems, and better application prospects.

The Gaussian Modulated Coherent State (GMCS) protocol is currently one of the most mature Continuous Variable Quantum Key Distribution (CVQKD) protocols. The system based on the GMCS protocol has the characteristics of classical light field, convenient preparation, and long-distance secure transmission, which can better meet the application requirements [8,9]. In addition, theoretically, continuous variable quantum key distribution has been proven to be secure against arbitrary collective attacks [10–12]. However, due to the difference between the idealized theoretical model and the actual physical system, Eve exploits the existing practical security loopholes to launch attacks on the CVQKD system or protocol, such as wavelength attack [13], calibration attack [14], local oscillator intensity (ILO) attack [15], and saturation attack [16]. Fortunately, a large number of researchers have proposed corresponding defense measures against possible attack strategies [17–20]. Moreover, in recent years, researchers have found that machine learning technology has significant advantages as a detection method in practical CVQKD systems. D. Huang and S. Liu summarized the work of using machine learning as an attack and defense detection method in CVQKD systems in recent years [21]. It is also mentioned that researchers now use machine learning to reduce the complexity of CVQKD systems to a certain extent while improving performance and ensuring system security. Furthermore, it is believed that neural network-based solutions require very few additional devices and can be easily implemented in practical CVQKD systems.

In fiber-based CVQKD systems, the transmittance is usually considered a constant and plays an important role in the subsequent parameter estimation process. In practical communication, the channel transmittance is highly likely to be controlled by Eve. By slightly changing the channel transmittance, Eve can cause significant parameter estimation errors in the subsequent process. As a result, the secure key rate is underestimated, then the estimated secure transmission distance could be reduced significantly, and the secure channel is insecure. Ultimately, communication interruption will be caused. This attack is similar in performance to denial-of-service attacks in classical channel communication. Y.Li et al. proposed a denial-of-service (DoS) attack strategy targeting the parameter estimation method in the communication process [22]. Dai et al. proposed a detection method based on the spectral estimation method, which effectively detected the low-rate denial-of-service (LDoS) attacks in CVQKD systems [23]. However, in practical implementations, the transmission of legitimate optical signals and the presence of real detectors and electronic components introduce inherent fluctuations, which pose challenges for accurate estimation. To address these challenges, Alice and Bob have to engage in multiple iterative computations to achieve precise estimation. Moreover, the estimation process usually occurs after the completion of the key transmission. In the unfortunate event of detecting an attack, the entire key data must be discarded, leading to substantial time and resource wastage. Hence, an efficient quantum denial-of-service (DoS) attack defense solution is needed to effectively resist such attacks.

This paper proposes an efficient detection scheme based on deep learning to detect DoS attacks targeting the parameter estimation method in CVQKD. The typical characteristics of two kinds of DoS attacks are studied, and the changes in these characteristics under attack are analyzed by referring to the normal state of not being attacked. A set of feature vectors labeled by different DoS attack types are constructed to train the neural network, and the data are preprocessed and input into the proposed neural network model. The trained neural network model can automatically detect abnormal feature vectors and classify them. Thus, a DoS attack detection model has been established, capable of recognizing DoS attacks launched by Eve, solely by employing a single forward propagation computation process. Bob can input the received keys into the model sequentially, and in the event of detecting any anomalous data, the transmission process will be immediately terminated. Consequently, Bob is alleviated from the need to wait for the completion of the key transmission process before assessing the system's vulnerability to attacks, which will greatly improve the efficiency of the system. Using the proposed detection model, the corresponding experiments and simulations are carried out on the CVQKD system based on Gaussian modulation.

2. Principle

2.1. GMCS CVQKD System Description

In the ideal condition of GMCS-CVQKD modulation protocol, the sender Alice sends the eigenstate of the complementary observation quantities X_A and Z_A following zero-center Gaussian distribution with a variance of $V_A N_0$ to the receiver node Bob, and Bob adaptively performs one of the regular measurements with the null difference detection system. After several rounds of exchange, Alice and Bob will share a series of relative Gaussian variables and gain the key with further post-processing approaches which include parameter estimation, reserve correction, and amplification. In the parameter estimation of CVQKD systems, the traditional statistical method is commonly used to estimate the channel parameters. In a theoretical CVQKD system, Alice and Bob can share a set of interrelated data $x = \{x_1, x_2 \ldots x_n\}$ and $y = \{y_1, y_2 \ldots y_n\}$, where x represents the canonical component modulated by Alice (X_A or P_A),y represents the canonical component value measured by Bob (X_B or P_B), and n represents the total amount of shared data. In the parameter estimation method, T and ε are used to indicate channel transmittance and channel excess noise, respectively. Therefore, the data transmission of Alice and Bob is usually assumed to be a normal linear transmission model, which can be expressed as

$$y = tx + z \tag{1}$$

where $t = \sqrt{T} \in R$ and z follows zero-centered Gaussian distribution with variance $\sigma^2 = 1 + T\varepsilon$.

However, in the CVQKD systems working situations, extra noise will be introduced into the system via complicated channel conditions and additional instruments [24]. To evaluate the security performance of the communication channel, a detailed channel and equipment parameters should be introduced to acquire calculations and evaluations with more precision. In general situations, characteristics such as the shot noise N_0, the technical excess noise ξ, the electrical noise V_{el}, and the efficiency of zero-difference detector η should be extra considered. Moreover, in the evaluation of the system's theoretical security key rate, all these parameters related to key extraction are calculated in the unit of N_0. For example, the detector's electronic noise is expressed as $V_{el} = v_{el}N_0$ and the technical excess noise of the system is expressed as $\eta = \varepsilon N_0$. Therefore, the aforementioned Formula (1) can be rewritten into a normal linear transmission model under complex channels in actual communication

$$y_B = t_R x_A + z_R \tag{2}$$

where, $t_R = \sqrt{\eta T} \in R$, and z_R follows a Gaussian distribution with a mean of 0 and a variance of $\sigma_R^2 = v_{el} + N_0 + \eta T \xi$.

In the subsequent parameter estimation, Alice and Bob randomly extract m $(m < n)$ sets of data from the total amount of shared data n to evaluate the parameters T and ε of the quantum communication channel. According to Formula (2), the maximum likelihood estimation method can be used to calculate

$$\begin{aligned} \hat{t}_R &= \frac{\sum_{i=1}^{m} x_{A,i} y_{B,i}}{\sum_{i=1}^{m} x_{A,i}^2} \\ \hat{\sigma}_R^2 &= \frac{1}{m} \sum_{i=1}^{m} \left(y_{B,i} - \hat{t}_R x_{B,i} \right)^2 \end{aligned} \tag{3}$$

where $\hat{t}_R$ and $\hat{\sigma}_R^2$ are two separated estimators who obey normal distribution and chi-square distribution, respectively.

$$\begin{aligned} \hat{t}_R &\sim N\left(t_R, \frac{\sigma_R^2}{\sum_{i=1}^{m} x_{A,i}} \right) \\ \frac{m\hat{\sigma}_R^2}{\sigma_R^2} &\sim \chi^2(m-1) \end{aligned} \tag{4}$$

where $\hat{t}_R$ and $\hat{\sigma}_R^2$ are $\sqrt{\eta T}$ and $v_{el} + N_0 + \eta T \xi$, respectively.

Therefore, we can use Formula (3) to fully estimate the transmittance T of the channel and the channel excess noise ε without considering the finite-size effect [25]. Meanwhile, when the total amount of data extracted is large enough, the chi-square distribution in Formula (4) can converge to an expressible normal distribution and the corresponding central interval can be calculated, which has been proved.

The confidence intervals for $\hat{t}_R$ and $\hat{\sigma}_R^2$ can be expressed as follows

$$\begin{aligned} \hat{t}_R &\in [t_R - \Delta t_R, t_R + \Delta t_R] \\ \hat{\sigma}_R^2 &\in \left[\sigma_R^2 - \Delta\sigma_R^2, \sigma_R^2 + \Delta\sigma_R^2\right] \end{aligned} \tag{5}$$

where $\Delta t_R = Z_{\epsilon_{PE}/2}\sqrt{\left(\frac{\sigma^2}{mV_{x_A}}\right)}, \Delta\sigma_R^2 = z_{\epsilon_{PE}/2}\frac{\sigma^2\sqrt{2}}{\sqrt{m}}$ and $z_{\in_{PE}/2}$ is a coefficient which obeys $1 - \frac{1}{2}\,\mathrm{erf}\left(z_{\epsilon_{PE}/2}/\sqrt{2}\right) = \frac{1}{2}\epsilon_{PE}$. Additionally, $1 - \epsilon_{PE}$ represents the confidence level. Moreover, $\mathrm{erf}(\cdot)$ denotes the error function:

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt \tag{6}$$

Since considering the finite-size effect [25] will only misguide the analysis process without changing the final result, the finite-size effect can be ignored. Next, the impact of Eve's perturbation of quantum channel parameters on subsequent channel parameter estimates will be discussed.

2.2. Estimation of Quantum Channel Parameters in Complex Communication

In the communication process, the modulated data of Alice terminal are X_A, and the measured data of Bob terminal are Y_B. Both X_A and Y_B are Gaussian variables. Next, X_A and Y_B are used to evaluate the parameters of the quantum channel, and then to investigate the estimation deviation of the channel transmittance T and channel excess noise ε under the attack. Their expression in the case of complex communication can be assumed as:

$$X_A = |A|\cos(\theta) \tag{7}$$

$$Y_B = \sqrt{\eta T}[|A|\cos(\theta + \Delta\varphi) + X_\varepsilon] + X_N \tag{8}$$

In the Formula (7), $|A|$ represents the amplitude of the coherent state for AM modulation, while θ represents the phase of the coherent state for PM modulation. In the Formula (8), T represents the channel transmittance and η represents the efficiency of the homodyne detector. Notably, $\Delta\varphi$ represents the phase drift of these quantum states during transmission, and it should be noted that these parameters are completely independent of each other. Additionally, X_ε represents the increase in noise due to channel excess noise ε, while X_N represents the increase in noise due to scattered noise N_0 and detector electronic noise V_{el}. Moreover, as the existing phase compensation technology can effectively suppress phase noise, it is reasonable to set the parameter A to 0 in the above formula. It is also important to consider that noise causes the mean of the amplitude to be 0. Additionally, the mean value $\mathrm{E}\left(\mathrm{x}_\mathrm{A}^2\right) = \mathrm{E}\left(|\mathrm{A}|^2\cos^2\theta\right)$ is also equal to the modulation variance V_a. Therefore, based on Formulas (7) and (8), the new formula can be easily obtained as:

$$E(X_A Y_B) = E\left(\sqrt{\eta T}|A|^2\cos^2\theta\right) = E(\sqrt{\eta T})V_a \tag{9}$$

$$E\left(Y_B^2\right) = \eta V_a E(T) + \eta\xi E(T) + N_0 + V_{el} \tag{10}$$

Next, by utilizing X and Y and neglecting the finite-size effect of the parameter estimation block size m, the formulas below can be estimated by combining X_A, Y_B, and Formula (3):

$$\begin{aligned} \hat{t}_R &= \frac{E(X_A Y_B)}{E\left(X_A^2\right)} \\ \sigma_R^2 &= E\left[\left(Y_B - \hat{t}_R X_A\right)^2\right] \end{aligned} \tag{11}$$

At the same time, the quantum channel transmittance T and channel over-noise ε can be obtained by transforming the aforementioned estimators and the parameters in the normal linear transmission model in the case of complex channels in actual communication.

$$\begin{aligned} T &= \frac{\hat{t}_R^2}{\eta} \\ \varepsilon &= \frac{\hat{\sigma}_R^2 - N_0 - V_{el}}{N_0 \hat{t}_R^2} \end{aligned} \tag{12}$$

Combining Formulas (7)–(12), the estimated value of quantum channel parameters in the complex communication environment can be obtained as:

$$\begin{aligned} \hat{T} &= [E(\sqrt{T})]^2 \\ \hat{\varepsilon} &= \frac{E(T)}{[E(\sqrt{T})]^2}(V_a + \varepsilon) - V_a \end{aligned} \tag{13}$$

Through the analysis of the Formula (13), it is clear that if the channel transmittance remains unchanged, the parameters of the channel excess noise can be estimated accurately. However, if Eve attacks the quantum communication channel and changes the transmittance of the channel, it will lead to inaccurate excess noise estimation of the channel which could cause a large deviation in the estimation of the channel security key rate. Then, the parameter change of the channel under Eve attack is analyzed, and the attack detection module will be proposed.

3. DoS Attack Launched by Eve in Complex Channel Environment

The purpose of this chapter is to study the security impact caused by Eve's denial of service attacks on CVQKD system. Then provide a detection method in Bob to detect DoS attacks on the system. Finally, adjust the parameter estimation in the subsequent stages according to the detected attack types. Additionally, this chapter is only focusing on the detection scheme that can detect DoS attacks launched by Eve, not the subsequent parameter estimation adjustment scheme.

In the absence of an attack, the CVQKD system operates in a familiar pattern. Alice was a string of coherent state $|X_A + iP_A\rangle$, coherent state of regular X_a and P_a variance component values for $V_a N_0$ bivariate Gaussian distribution, and the key information encoded in the above. Alice sends it to the receiver node Bob, where n_0 is scatter shot noise, V_a represents the modulation variance V_m, and the intersection uncertainty variance V_s. The quantum channel has the full characteristics of quantum channel transmission T, modulation variance Va, detector electronic noise v_{el}, and technical excess noise ε introduced by the actual channel. All of these key extraction-related parameters are in units of granular noise n_0. When the transmission is completed, the mean value and variance of the output of the two-mode state measured by Bob can be expressed as

$$\begin{aligned} \bar{y}_b &= 0 \\ V_b &= \eta T V_a N_0 + \eta T \varepsilon N_0 + N_0 + V_{el} \end{aligned} \tag{14}$$

where η is the efficiency of the heterodyne detection or homodyne detector, $V_{el} = v_{el} N_0$.

From the above section, once the channel is attacked by Eve, the channel transmittance T will be changed, which will result in the deviation of the system channel parameter estimation. Finally, this lead to the deviation of the channel security key rate estimation, and the communication parties are forced to close the channel. Importantly, Eve can actively control the channel to interfere with the communication environment [26], but it is difficult to know the specific law of transmittance change of the quantum communication channel. According to [22], Eve can easily manipulate the channel so that the channel transmittance follows a two-point distribution or a uniform distribution. Therefore, under the assumption that the quantum channel is attacked by Eve, the channel transmittance T obeys a certain distribution. In the following sections, this paper will discuss the changes in quantum channel parameters when the transmittance T of the quantum channel follows the distribution of two points, respectively, in a complex communication environment, and then the case of uniform distribution will be derived.

3.1. Two-Point Distribution of Channel Transmittance T

In this section, it is assumed that the transmittance of the channel follows a two-point distribution. Under this condition, transmittance can be varied between 0 and T_0, which is the channel transmittance in the case of complex quantum channels. Because Eva can manipulates this channel easily. Hence, the serious biases on channel estimation could be difficult to correct.

The assumptions previously stated can be precisely articulated using the mathematical expression $\mathrm{T/T0} \sim \mathrm{B(1,p)}$, where $\mathrm{B(1,p)}$ denotes a binomial distribution with a single trial and success probability p. Subsequently, it can be deduced that $E(T)$ and $E(\sqrt{T})$ are equivalent to PT_0 and $P\sqrt{T_0}$, respectively. In particular, it is assumed in this context that the evaluation of the quantum channel parameter is ideal and regarded as its true value, based on which the following formula can be calculated using Formula (13).

$$\begin{aligned} T_{DoS,P} &= p^2 T_0 \\ \varepsilon_{DoS,P} &= \frac{1}{p}(V_a + \varepsilon) - V_a \end{aligned} \tag{15}$$

where the parameter P represents the probability of channel transmission rate T being equal to T_0, where P equals 1 in the absence of Eve attacks on the quantum communication channel. The parameter ε represents the actual excess noise in the channel. The "DoS, P" represents the situation that the transmittance T of the quantum channel in a complex communication environment obeys the distribution of two points.

Subsequently, by substituting Formula (15) into Formula (14) and rearranging, the following formula can be obtained as:

$$\begin{aligned} &\bar{y}_{DoS,P} = 0 \\ &V_{DoS,P} = \eta T_{DoS,P}(V_A N_0 + \xi_{DoS,P}) + N_0 + V_{el} \end{aligned} \tag{16}$$

where $\xi_{DoS,P} = N_0 \varepsilon_{DoS,P}$ represents excess noise.

3.2. Uniform Distribution of Channel Transmittance T

In this scenario, the channel transmittance T is a uniformly distributed random variable between λT_0 and T_0, where λ is a constant such that $0 < \lambda < 1$. That is, T follows a uniform distribution with probability density function $f(T) = 1/(T_0 - \lambda T_0)$ for $\lambda T_0 \leq T \leq T_0$. Hence, the expected values of T and $\sqrt{T}$ are given by

$$\begin{aligned} E(T) &= \int_{\lambda T_0}^{T_0} \frac{1}{T_0 - \lambda T_0} T dT = \frac{(1+\lambda)T_0}{2}, \\ E(\sqrt{T}) &= \int_{\lambda T_0}^{T_0} \frac{1}{T_0 - \lambda T_0} \sqrt{T} dT = \frac{2\left(1 - \lambda^{\frac{3}{2}}\right)\sqrt{T_0}}{3(1-\lambda)} \end{aligned} \tag{17}$$

Furthermore, by combining Formula (13), the following result can be obtained:

$$
\begin{aligned}
T_{DoS,U} &= \frac{4\left(1-\lambda^{\frac{3}{2}}\right)^2}{9(1-\lambda)^2} T_0 \\
\varepsilon_{DoS,U} &= \frac{V_a\left(1-9\lambda+16\lambda^{\frac{3}{2}}-9\lambda^2+\lambda^3\right)}{8\left(1-\lambda^{\frac{3}{2}}\right)^2} \\
&\quad + \frac{9\varepsilon\left(1-\lambda-\lambda^2+\lambda^3\right)}{8\left(1-\lambda^{\frac{3}{2}}\right)^2}
\end{aligned}
\tag{18}
$$

where the "DoS, U" represents the situation that the transmittance T of the quantum channel in a complex communication environment obeys a uniform distribution. Using it in Formula (14), the Formula (17) can be derived:

$$
\begin{aligned}
&\bar{y}_{DoS,U} = 0 \\
&V_{DoS,U} = \eta T_{DoS,U}(V_A N_0 + \xi_{DoS,U}) + N_0 + V_{el}
\end{aligned}
\tag{19}
$$

Here, the technical noise is represented as $\xi_{DoS,U} = N_0 \varepsilon_{DoS,U}$.

3.3. CVQKD System Scheme in Complex Channel with Attack Detection Module

Figure 1 illustrates a remote CVQKD system, which incorporated a monitoring device in the experiment. Coherent light at 1550 nm was generated by Alice through an external telecommunication diode at a repetition rate of 1 MHz, with the pulse train split into a weak signal path and a strong local oscillator (LO) path by an unbalanced coupler. Before passing through a set of variable attenuators, the signal pulse underwent amplitude and phase modulation to enable control of the variance within a specified range. The signal and LO pulses were then delayed through the use of delay lines and Faraday mirrors, and polarization multiplexing was achieved by employing a polarization beam splitter (PBS) along two related optical paths. Both polarization and time-multiplexed paths were transmitted from Alice to Bob through a standard single-mode fiber quantum channel, whereupon the signal was received passively and uninterrupted by Bob. Eve may exist in the quantum channel during transmission.

When receiving a signal, Bob passively mitigates interference between the signal and the LO by using an active polarization controller. Subsequently, the signal and LO pulse timing were superimposed through another delay line, and measurements were generated by using a scattered-noise limited zero-difference detector. Random real-time quadrature selection was made by Bob via the control of the measured phase using a phase modulator in the LO path. Additionally, a small proportion of the pulse within the LO optical path was extracted for real-time measurement of particle noise and LO power. This monitoring method was employed to avoid any interference with the actual key rate and standard communication of the system. Additionally, the utilization of three isolation modulators within the system served to prevent potential Trojan attacks [27,28]. Post-processing entailed the transmission of relevant parameters to the processing module (DPC) for security analysis.

It is noteworthy that the original processing module was augmented by additional data preprocessing programs and attack detection programs during the post-processing phase. Subsequent to detector analysis, the results were imported into further processes through a small interface program. In our past high repetition-rate CVQKD system experiment, real-time data collection and computation were facilitated by a computer equipped with 64 GB of memory, an Intel Core 5.0 GHz CPU, and a GeForce RTX 3090 GPU. This experimental process was considered to be optimally suited for systems operating within the 1–100 Mbps range.

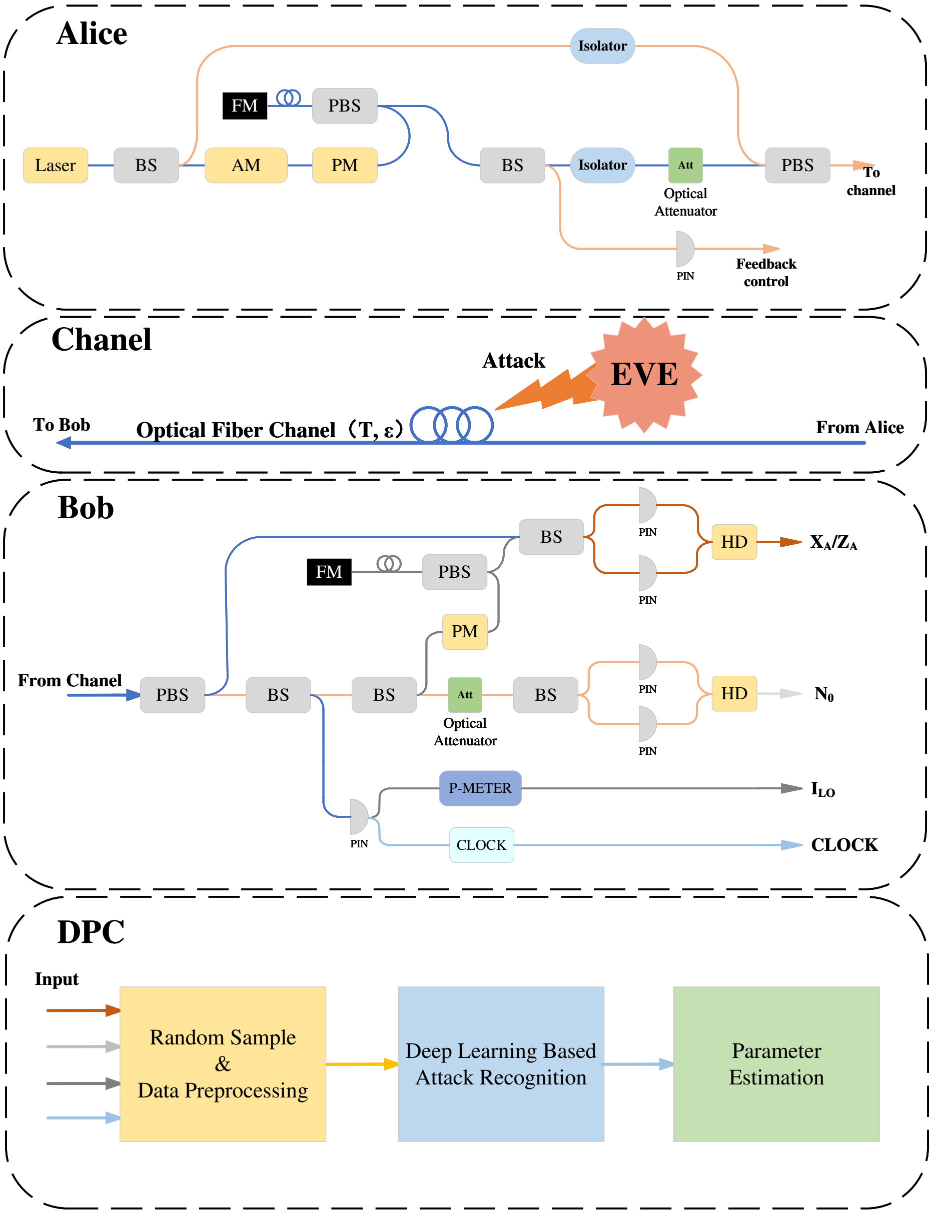


Figure 1. Schematic diagram and experimental setup for acquiring training data. Laser: laser produce model; BS: beam splitter; AM: amplitude modulator; PM: phase modulator; PBS: polarizing beam splitter; Optical attenuator: makes the signal conform to the quantum signal level; HD: homodyne detection; FM: Faraday mirror; PIN: PIN photodiode. The data $[X_A/Z_A, I_{LO}, N_0, CLOCK]$ collected by the QKD system will be transmitted to the DPC. DPC: data processing center used to sample analog signal, attack detection, and parameter estimation.

4. Signal Detection Model

There are limited attack perception methods currently available for detecting DoS attacks on CVQKD systems, and the performance of the proposed traditional detection methods is not ideal. In this section, a neural network detection model with a neural network structure is introduced as a defense detection program for the CVQKD system. The network model is used to detect the scenario where Eve controls the channel and launches a DoS attack. This network enables existing CVQKD quantum communication systems to achieve a higher level of security without changing the original protocol. Different from traditional detection methods, this deep learning classifier considers classification and detection based on multiple coupling parameter details, rather than applying fixed thresholds to individual measurement parameters. It can extract potential abnormal data features from the actual system data without any prior assumptions.

4.1. Deep Learning Method

Deep learning has emerged as one of the most rapidly advancing and widely adopted subfields of machine learning [29], offering unprecedented computational capabilities for processing large volumes of data. Deep learning has become increasingly popular due to its ability to learn from large, complex datasets, and achieve state-of-the-art performance on a wide range of tasks, including image classification, object detection, and natural language processing. One of the key advantages of deep learning over traditional machine learning approaches is its ability to learn hierarchical representations of data through the use of multiple layers of artificial neural networks. Convolutional Neural Networks (CNNs) are a prominent neural network model that has found wide-ranging applications in tasks such as image processing [30–32], pattern recognition [33,34], natural language processing [35–37], and quantum communication [38].

CNN typically consists of several layers, including an input layer, one or more convolutional layers, a pooling layer, and one or more fully connected layers. Before being convolved in the convolutional layer, input data are typically normalized through the input layer, which can help to improve the convergence speed and overall performance of the CNN during training.

In the convolutional layer, a kernel is convolved with the output from the previous layer, followed by the activation function that applies a non-linear transformation to the resulting feature map. The output of each convolutional layer is the convolution of multiple input feature maps, which can be mathematically described as a dot product between the input and a set of learnable filters, which is mathematically described as [39]:

$$h_{i,j} = \sigma(b + \sum_{k=1}^{K}\sum_{l=1}^{L} w_{k,l}x_{i+k-1,j+l-1}) \tag{20}$$

In the Formula (20), $x_{i,j}$ represents the pixel value of the input data, $w_{k,l}$ represents the weight of the convolution kernel, b represents the bias term, K and L represent the size of the convolution kernel, $h_{i,j}$ represents an element of the output feature vector, and $\sigma(\cdot)$ represents the non-linear activation function.

After convolution, the activation function performs a non-linear operation on the output data of each convolution kernel. Non-linear activation functions can improve the non-linearity of neural network models and enhance the robustness of neural networks, thereby establishing complex non-linear mapping relationships between input and output data. $ReLU(x) = \max\{x, 0\}$ is a classical activation function. When the input data are greater than 0, the derivative of the activation function remains 1, which is beneficial for overcoming the gradient dispersion problem during the backpropagation process. It can be expressed as

$$a_i^{l+1}(j) = f\left(y_i^{l+1}(j)\right) = \max\left\{0, y_i^{l+1}(j)\right\} \tag{21}$$

where $y_i^{l+1}(j)$ is the output of convolution kernel; $a_i^{l+1}(j)$ is the non-linear activation value of $y_i^{l+1}(j)$.

The $ReLU$ activation function has been widely used in many deep learning models due to its simplicity and effectiveness in preventing the vanishing gradient problem.

The pooling operation is another key operation in CNN, which can reduce the dimensionality of the output feature vectors, lower the computational cost, and improve the robustness of the model. Max pooling, average pooling, and overlapping pooling are three common pooling operations. In average pooling, the average value of the feature vectors in each pooling window is taken as the output feature vector. This process can be described by the following equation

$$y_{i,j} = \frac{1}{K \times L} \sum_{k=1}^{K} \sum_{l=1}^{L} h_{i+k-1,j+l-1} \tag{22}$$

where $h_{i,j}$ is the output feature vector of the convolutional layer, K and L are the size of the pooling window, and $y_{i,j}$ is the output of the pooling layer.

The Attention Mechanism Module (AMM) is presented to enhance the robustness of CNN (AMM-CNN). The proposed approach utilizes a "Squeeze-and-Excitation" attention mechanism module to improve CNN performance by emphasizing relevant features and suppressing irrelevant ones [40]. Specifically, during the squeeze phase, AMM performs global pooling on the feature map of each channel to obtain a global statistic (e.g., mean or variance). In the excitation phase, AMM calculates a gating vector using the obtained global statistic to weight each channel's feature map, thereby amplifying useful features and diminishing useless ones. The introduction of AMM results in CNN learning more representative and distinctive features, thereby improving the model's accuracy and generalization ability. Additionally, AMM enhances CNN's resistance to noise and perturbations, improving the model's robustness.

The AMM comprises three primary operations, global pooling, squeezing, and exciting. The global pooling operation generates channel descriptors, as given by the following formula:

$$z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_{i,j,c} \tag{23}$$

where z_c represents the global descriptor of the c-th channel, and $x_{i,j,c}$ represents the feature value of the c-th channel at the i-th row and j-th column. Subsequently, the squeezing operation compresses the dimensionality of the channel descriptors to a lower dimension via a fully connected layer, as given by the following formula

$$a_c = \delta(\mathrm{FC}_1(z)) \tag{24}$$

where δ is an activation function, such as ReLU or Swish. Finally, the exciting operation transforms the squeezed descriptor into attention weights via another fully connected layer, as given by the following formula

$$s_c = \sigma(\mathrm{FC}_2(a)) \tag{25}$$

where σ is the sigmoid function, and s_c represents the attention weight of the c-th channel. The output feature map $Y \in \mathbb{R}^{H\times W\times C}$ of the AMM is obtained by multiplying each channel's feature value in the input feature map by its corresponding attention weight, as given by the following formula

$$y_{i,j,c} = s_c x_{i,j,c} \tag{26}$$

where $y_{i,j,c}$ represents the output feature value of the c-th channel at the i-th row and j-th column.

The last layer of the convolutional neural network consists of a flattening layer, which transforms the output feature maps of the previous layer into a one-dimensional vector. This vector is then fed into a fully connected layer, which aggregates local information extracted by the convolutional and pooling layers to produce the final prediction. The output of the fully connected layer can be expressed as

$$Y = \sigma(WX + b) \tag{27}$$

where X is the input feature vector, W is the weight matrix, b is the bias term, Y is the predicted value, and $\sigma(\cdot)$ is the activation function.

4.2. DoS Attack Detection Implementation Details

In the previous section, it is assumed that the transmittance T of the quantum channel under DoS attacks follows either a uniform distribution or a two-point distribution. This section will conduct simulation experiments in the CVQKD system shown in (Figure 1) to collect communication data without any attack, and we will launch attacks on the system and collect experimental data under DoS attacks.

In this experiment, the Python 3.8 and the Pytorch framework based on CUDA 11.3 were used for data preparation, model training, calculation, and verification of algorithm iterations. The program ran on a computer with 64 GB of memory, an Intel Core 5.0 GHz CPU, and a GeForce RTX 3090 GPU. Moreover, the trained model was tested under two possible DoS attacks in this experiment. Typically, in a safety CVQKD system, the system parameters are kept in dynamic stability. The experiment reasonably estimated the attack capability of the third party Eve, and added some possible noise. According to the standard realistic assumption for CVQKD [14], a long-distance CVQKD system was built with the following parameters. In experiment setup, η = 0.7, V_{el} = 0.01N_0, ξ = 0.1N_0, N_0 = 0.4, V_A = 10, $T = 10^{-\alpha L/10}$, P = 0.9, U = 0.8, where the transmission distance L was 30 km, and the loss coefficient of the optical fiber α was 0.2 dB/km. The local oscillator ILO at Bob's side is set as 10^7 photons per pulse with 1% percent fluctuation. The data collected from Bob's end consists of four dimensions. To train the classifier for detecting network attacks, three datasets were generated as training sets $Y_{train} = \{y_{normal},\ y_{DoS,P},\ y_{DoS,U}\}$. The datasets y_{normal} represent normal data that have not been attacked. $y_{DoS,P}$ represents the datasets that have been attacked by Eve's DoS attack, and the quantum channel transmittance rate T follows a two-point distribution. The datasets $y_{DoS,U}$ illustrate data that have been attacked by Eve's DoS attack and the quantum channel transmittance rate T follows a uniform distribution. The one-hot encoding was performed on the collected data, and the label mapping relationship of each dataset was $\{y_{normal},\ y_{DoS,P},\ y_{DoS,U}\} \Rightarrow \{0,\ 1,\ 2\}$. Y_{train} were preprocessed by segmentation and feature vector extraction, and their categories were labeled. The length of each dataset in Y_{train} was $N = 9.216 \times 10^7$, which is divided into $M = 10,000$ groups, each containing $Q = 9216$ data blocks. Therefore, the data vector of each data group can be transformed into a 96 $\times$ 96 feature vector. Next, these feature vectors were substituted into the AMM-CNN for training to learn the features of different attack strategies. To speed up the optimal solution search of the gradient descent algorithm and eliminate the influence of different attributes with different scales, the data were min–max processed. In addition, three other datasets, which were $Y_{test} = \left\{y'_{normal},\ y'_{DoS,P},\ y'_{DoS,U}\right\}$, were also prepared as test sets and processed them in the same way. It should be noted that the data in the test set are completely independent of the data in the training set. Finally, these unlabeled feature vectors were imported into the trained AMM-CNN classifier to verify the performance of the attack classification. The activation function of the model was unified as the ReLU function, the loss function used the cross-entropy loss function, the batch-size was set to 128, and the learning rate was set to 0.01. The brief step can be sighted in Figure 2.

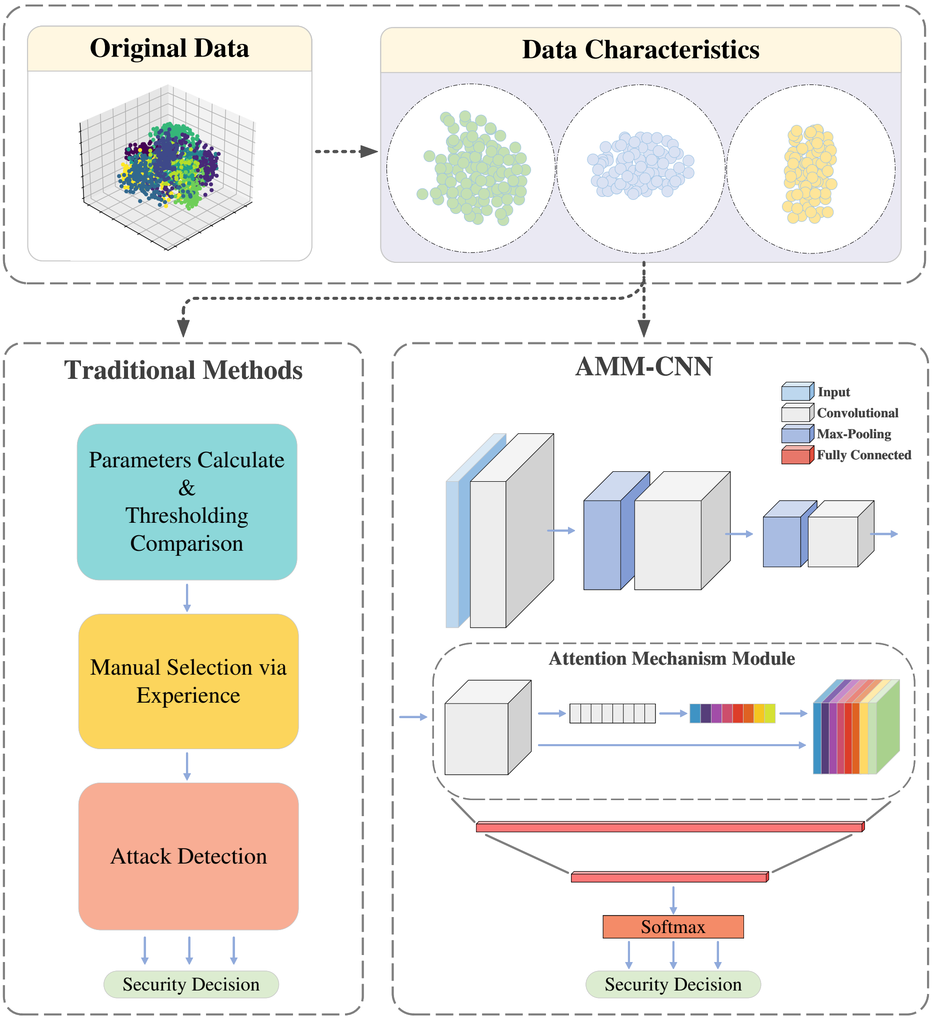
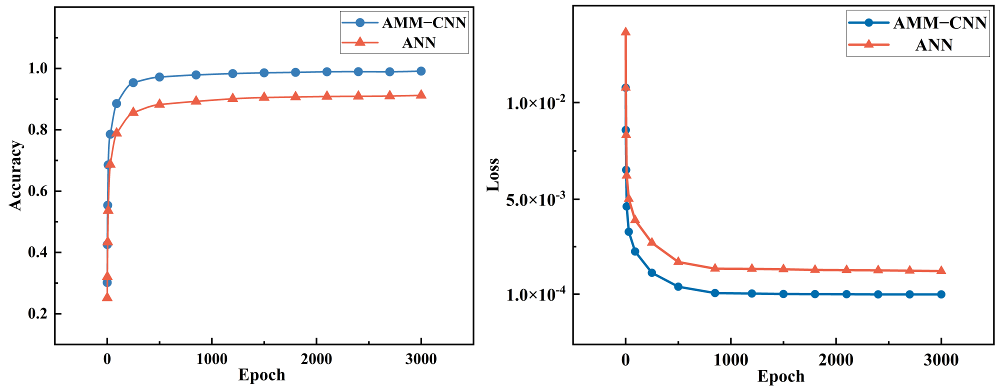


Figure 2. Traditional attack detection (left) and deep learning attack detection based on AMM-CNN (right); The data collected by Bob will be transmitted to the detection model in real time. The original data shows the distribution of the parameters under the interference condition. Data characteristics represent the characteristics of three different types of data.

5. Performance

To evaluate the detection capability of Eve launched DoS attacks, we compared our proposed AMM-CNN and the Artificial Neural Network (ANN) model that was reproduced from the ANN paper [19]. The ANN model consists of three layers, the input layer, hidden layer, and softmax layer. Specifically, the parameter n_e for the hidden layer is set to 15, which was chosen based on the recommendation in the ANN paper [19]. Furthermore, the training process and loss function of the ANN model are the same as those of the AMM-CNN model. The experiment results show that both the AMM-CNN and ANN models effectively detect DoS attacks in CVQKD systems. Considering the large amount of data processed and the complex feature mapping of the data, the AMM-CNN model has an advantage in detecting DoS attacks in a massive data feature perception. Therefore, it outperforms the ANN model in attack perception accuracy, with an accuracy rate of 98.7%. Although the ANN model has stronger interpretability, its accuracy rate of 91.6% is slightly inferior to the AMM-CNN model. The classification accuracy during the training process is shown in Figure 3a. In addition, the training loss curve of the AMM-CNN model in this experiment is shown in Figure 3b. Furthermore, the confusion matrix results of the AMM-CNN model are presented in Figure 4.



(**a**) (**b**)

Figure 3. (**a**) Accuracy function of two models with sufficient training data. (**b**) Loss function of two models with sufficient training data.

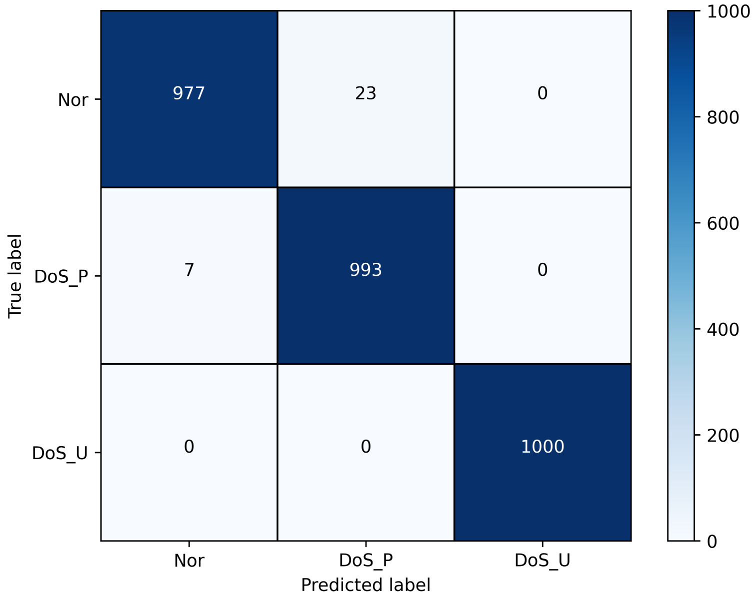
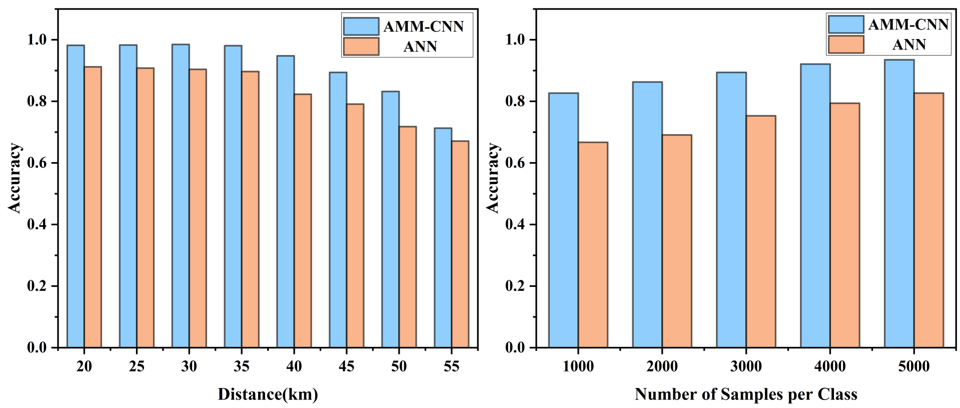


Figure 4. The confusion matrix table for the AMM-CNN model. Nor: The "Nor" represents normal data that is not under DoS attack; DoS_P: The "DoS_P" represents the data whose quantum channel transmission rate T follows a two-point distribution after Eve's DoS attack; DoS_U: The "DoS_U" represents the data whose quantum channel transmission rate T obeys uniform distribution after Eve DoS attack.

The performance difference between the two models can be attributed to several factors. Firstly, the ANN model requires a large number of parameters to process high-dimensional CVQKD data, which may cause overfitting. In contrast, the AMM-CNN model can reduce the dimension of the training dataset and reduce the number of parameters through convolution and pooling operations, making it more suitable for large-scale data processing. Secondly, the ANN model may not have considered the spatial information in the data, only treating it as a flattened vector input to the network. Meanwhile, the AMM-CNN model retains spatial structural information during the training process through convolution and pooling operations. Finally, compared with the ANN model, the AMM-CNN model requires less input data preprocessing. This is because the AMM-CNN model can extract features from the data through convolution, pooling operations, and the introduction of an attention mechanism module, while the ANN model requires more preprocessing to achieve better results.

To recap, the research results show that the AMM-CNN model has better detection performance for DoS attack data in the experiment's CVQKD dataset, while the ANN model is slightly inferior but has better interpretability. The choice of model depends on specific application needs and available resources. Future research can explore more neural network models, combine the advantages of multiple models, and develop hybrid models to achieve better performance in detecting DoS attacks in CVQKD systems.

Moreover, the impact of different transmission distances on the evaluation metrics of the proposed attack detection model was investigated using the optimal parameters obtained from previous experiments. The same ANN model was used as a comparative experiment. As shown in Figure 5a, when the transmission distance L is 20 km to 55 km, it can be seen that the detection accuracy of the AMM-CNN model and the ANN model show a negative correlation with the transmission distance. That is, the model accuracy gradually decreased with an increase in the transmission distance. When L was less than 35 km, the proposed AMM-CNN model demonstrated a good detection performance for Eve's DoS attacks, with the model accuracy hovering around 98.0%. However, when L equaled 40 km and greater, the model accuracy began to decrease slightly. This may be due to the influence of system noise and channel noise, which cause a deterioration in data quality as the transmission distance increases. As a result, the attack detection model gradually becomes less effective at extracting relevant features, thereby reducing its performance. Finally, when L equaled 55 km, the AMM-CNN accuracy was 71.3%, while the ANN accuracy was 61.7%. Overall, the results obtained from the AMM-CNN model were satisfactory.



(**a**) (**b**)

Figure 5. (**a**) The detection accuracy corresponding to each distance. (**b**) The detection accuracy corresponding to each number of samples in the small sample case.

Additionally, experiments were conducted to test the performance of the attack detection model under small sample conditions, considering the limitation of the training data volume. The experiment gradually increased the size of the training dataset from 1000 to 5000, using the same dataset $Y_{test} = \left\{ y'_{normal}, \ y'_{DoS,P}, \ y'_{DoS,U} \right\}$ for testing. It can be easily observed from Figure 5b that the accuracy of the model gradually improved with the increase in sample size. Even under small dataset conditions, the proposed AMM-CNN model demonstrated excellent performance, whereas the performance of the ANN model was far inferior under small sample experiments. This also indicates that the AMM-CNN model has good feature extraction ability and robustness under small sample conditions. The introduction of the attention mechanism module in the AMM-CNN model enhances the model's ability to focus on important features. By learning key information and context from the data, the attention mechanism guides the model to better distinguish between attacks and normal behavior, thereby improving accuracy. Furthermore, the weight allocation of attention allows the model to assign weights to features from different regions, reducing sensitivity to noise and attacks, thereby enhancing the robustness of the AMM-CNN model. This robustness enhancement enables the model to better handle variations and perturbations in the data, leading to improved system stability and reliability.

In summary, the AMM-CNN model can not only effectively learn the detailed distribution of DoS attack samples but also abstract potential time-dependent features and attribute-dependent features to detect potential threats robustly. In particular, the AMM-CNN model can achieve higher detection accuracy than other methods. Moreover, the AMM-CNN model has good robustness under conditions where only a few samples are available for learning.

6. Conclusions

In conclusion, the article presents a novel and unique defense strategy against quantum attacks on CVQKD systems in complex communication scenarios. Our proposed methodology stands out from existing approaches due to its incorporation of CNN deep learning algorithms and its comprehensive consideration of the impact of potential DoS attacks in realistic CVQKD systems. By harnessing the capabilities of deep learning algorithms, we trained an attack-aware model that exhibits an impressive detection accuracy of 98.7% against DoS attacks. This high accuracy in detecting DoS attacks demonstrates the effectiveness and reliability of our approach in addressing these specific threats. Moreover, our research includes comprehensive experiments conducted under various scenarios, such as long-distance and small-sample situations. By exploring previously uncharted territories, we have demonstrated the robustness and applicability of our method across diverse communication environments. Additionally, our work provides a fresh perspective within the realm of CVQKD systems. The seamless integration of deep learning and quantum cryptography not only enhances the security of CVQKD systems against quantum DoS attacks but also opens up unprecedented avenues for further advancements in quantum communication. This amalgamation paves the way for the development of more robust and efficient quantum communication protocols, with potential implications across a wide array of industries reliant on secure information exchange. To summarize, our research presents a distinctive and pioneering defense strategy against DoS attacks on CVQKD systems.

Author Contributions: Conceptualization, W.Y.; methodology, W.Y.; resources, D.H.; data curation, W.Y.; writing—original draft preparation, W.Y. and Y.Z.; software, W.Y. and Y.Z.; writing—review and editing, D.H.; supervision, D.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Scarani, V.; Bechmann-Pasquinucci, H.; Cerf, N.J.; Dušek, M.; Lütkenhaus, N.; Peev, M. The security of practical quantum key distribution. *Rev. Mod. Phys.* **2009**, *81*, 1301. [CrossRef]
2. Gisin, N.; Ribordy, G.; Tittel, W.; Zbinden, H. Quantum cryptography. *Rev. Mod. Phys.* **2002**, *74*, 145. [CrossRef]
3. Weedbrook, C.; Pirandola, S.; García-Patrón, R.; Cerf, N.J.; Ralph, T.C.; Shapiro, J.H.; Lloyd, S. Gaussian quantum information. *Rev. Mod. Phys.* **2012**, *84*, 621. [CrossRef]
4. Ekert, A.K. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.* **1991**, *67*, 661. [CrossRef]
5. Grosshans, F.; Grangier, P. Continuous variable quantum cryptography using coherent states. *Phys. Rev. Lett.* **2002**, *88*, 057902. [CrossRef]
6. Gong, L.H.; Song, H.C.; He, C.S.; Liu, Y.; Zhou, N.R. A continuous variable quantum deterministic key distribution based on two-mode squeezed states. *Phys. Scr.* **2014**, *89*, 035101. [CrossRef]
7. Lance, A.M.; Symul, T.; Sharma, V.; Weedbrook, C.; Ralph, T.C.; Lam, P.K. No-switching quantum key distribution using broadband modulated coherent light. *Phys. Rev. Lett.* **2005**, *95*, 180503. [CrossRef]
8. Grosshans, F.; Van Assche, G.; Wenger, J.; Brouri, R.; Cerf, N.J.; Grangier, P. Quantum key distribution using gaussian-modulated coherent states. *Nature* **2003**, *421*, 238–241. [CrossRef]
9. Navascués, M.; Grosshans, F.; Acin, A. Optimality of Gaussian attacks in continuous-variable quantum cryptography. *Phys. Rev. Lett.* **2006**, *97*, 190502. [CrossRef]

10. Furrer, F.; Franz, T.; Berta, M.; Leverrier, A.; Scholz, V.B.; Tomamichel, M.; Werner, R.F. Continuous variable quantum key distribution: Finite-key analysis of composable security against coherent attacks. *Phys. Rev. Lett.* **2012**, *109*, 100502. [CrossRef]
11. García-Patrón, R.; Cerf, N.J. Unconditional optimality of Gaussian attacks against continuous-variable quantum key distribution. *Phys. Rev. Lett.* **2006**, *97*, 190503. [CrossRef] [PubMed]
12. Leverrier, A. Composable security proof for continuous-variable quantum key distribution with coherent states. *Phys. Rev. Lett.* **2015**, *114*, 070501. [CrossRef] [PubMed]
13. Huang, J.Z.; Kunz-Jacques, S.; Jouguet, P.; Weedbrook, C.; Yin, Z.Q.; Wang, S.; Chen, W.; Guo, G.C.; Han, Z.F. Quantum hacking on quantum key distribution using homodyne detection. *Phys. Rev. A* **2014**, *89*, 032304. [CrossRef]
14. Jouguet, P.; Kunz-Jacques, S.; Diamanti, E. Preventing calibration attacks on the local oscillator in continuous-variable quantum key distribution. *Phys. Rev. A* **2013**, *87*, 062313. [CrossRef]
15. Ma, X.C.; Sun, S.H.; Jiang, M.S.; Liang, L.M. Local oscillator fluctuation opens a loophole for Eve in practical continuous-variable quantum-key-distribution systems. *Phys. Rev. A* **2013**, *88*, 022339. [CrossRef]
16. Qin, H.; Kumar, R.; Alléaume, R. Quantum hacking: Saturation attack on practical continuous-variable quantum key distribution. *Phys. Rev. A* **2016**, *94*, 012325. [CrossRef]
17. Luo, H.; Zhang, L.; Qin, H.; Sun, S.; Huang, P.; Wang, Y.; Wu, Z.; Guo, Y.; Huang, D. Beyond universal attack detection for continuous-variable quantum key distribution via deep learning. *Phys. Rev. A* **2022**, *105*, 042411. [CrossRef]
18. Du, H.; Huang, D. Multi-Attack Detection: General Defense Strategy Based on Neural Networks for CV-QKD. *Photonics* **2022**, *9*, 177. [CrossRef]
19. Mao, Y.; Huang, W.; Zhong, H.; Wang, Y.; Qin, H.; Guo, Y.; Huang, D. Detecting quantum attacks: A machine learning based defense strategy for practical continuous-variable quantum key distribution. *New J. Phys.* **2020**, *22*, 083073. [CrossRef]
20. Wu, Z.; Wang, Y.; Zhang, L.; Mao, Y.; Luo, H.; Guo, Y.; Huang, D. Sifting scheme for continuous-variable quantum key distribution with short samples. *JOSA B* **2022**, *39*, 694–704. [CrossRef]
21. Huang, D.; Liu, S.; Zhang, L. Secure Continuous-Variable Quantum Key Distribution with Machine Learning. *Photonics* **2021**, *8*, 511. [CrossRef]
22. Li, Y.; Huang, P.; Wang, S.; Wang, T.; Li, D.; Zeng, G. A denial-of-service attack on fiber-based continuous-variable quantum key distribution. *Phys. Lett. A* **2018**, *382*, 3253–3261. [CrossRef]
23. Dai, E.; Huang, D.; Zhang, L. Low-rate denial-of-service attack detection: Defense strategy based on spectral estimation for CV-QKD. *Photonics* **2022**, *9*, 365. [CrossRef]
24. Leverrier, A.; García-Patrón, R.; Renner, R.; Cerf, N.J. Security of continuous-variable quantum key distribution against general attacks. *Phys. Rev. Lett.* **2013**, *110*, 030502. [CrossRef] [PubMed]
25. Leverrier, A.; Grosshans, F.; Grangier, P. Finite-size analysis of a continuous-variable quantum key distribution. *Phys. Rev. A* **2010**, *81*, 062343. [CrossRef]
26. Grosshans, F.; Cerf, N.J. Continuous-variable quantum cryptography is secure against non-Gaussian attacks. *Phys. Rev. Lett.* **2004**, *92*, 047905. [CrossRef]
27. Gisin, N.; Fasel, S.; Kraus, B.; Zbinden, H.; Ribordy, G. Trojan-horse attacks on quantum-key-distribution systems. *Phys. Rev. A* **2006**, *73*, 022320. [CrossRef]
28. Jain, N.; Anisimova, E.; Khan, I.; Makarov, V.; Marquardt, C.; Leuchs, G. Trojan-horse attacks threaten the security of practical quantum cryptography. *New J. Phys.* **2014**, *16*, 123030. [CrossRef]
29. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
30. Liu, F.; Lin, G.; Shen, C. CRF learning with CNN features for image segmentation. *Pattern Recognit.* **2015**, *48*, 2983–2992. [CrossRef]
31. Han, F.; Yao, J.; Zhu, H.; Wang, C. Underwater image processing and object detection based on deep CNN method. *J. Sensors* **2020**, *2020*, 6707328. [CrossRef]
32. Huang, L.; He, M.; Tan, C.; Jiang, D.; Li, G.; Yu, H. Retracted: Jointly network image processing: Multi-task image semantic segmentation of indoor scene based on CNN. *IET Image Process.* **2020**, *14*, 3689–3697. [CrossRef]
33. Wu, H.; Huang, Q.; Wang, D.; Gao, L. A CNN-SVM combined model for pattern recognition of knee motion using mechanomyography signals. *J. Electromyogr. Kinesiol.* **2018**, *42*, 136–142. [CrossRef] [PubMed]
34. Kwon, M.C.; Park, G.; Choi, S. Smartwatch user interface implementation using CNN-based gesture pattern recognition. *Sensors* **2018**, *18*, 2997. [CrossRef] [PubMed]
35. Yin, W.; Kann, K.; Yu, M.; Schütze, H. Comparative study of CNN and RNN for natural language processing. *arXiv* **2017**, arXiv:1702.01923.
36. Li, H. Deep learning for natural language processing: Advantages and challenges. *Natl. Sci. Rev.* **2018**, *5*, 24–26. [CrossRef]
37. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 53. [CrossRef]
38. Xing, Z.; Li, X.; Ruan, X.; Luo, Y.; Zhang, H. Phase Compensation for Continuous Variable Quantum Key Distribution Based on Convolutional Neural Network. *Photonics* **2022**, *9*, 463. [CrossRef]

39. Wang, X.; Mao, D.; Li, X. Bearing fault diagnosis based on vibro-acoustic data fusion and 1D-CNN network. *Measurement* **2021**, *173*, 108518. [CrossRef]
40. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 7132–7141.