КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 3 (23), 2024
ISSN 2663 - 4023
DOI 10.28925/2663-4023.2024.23.97109
UDC 004.056

Jiang Xue
PhD candidate
Bengbu University, Bengbu City, Anhui Province, China
National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine
ORCID 0009-0000-1676-2331
jx1283@163.com

Valerii Lakhno
Doctor of Technical Sciences, Professor, Department of Computer Systems, Networks and Cybersecurity
National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine
ORCID 0000-0001-9695-4543
lva964@nubip.edu.ua

Andrii Sahun
Candidate of Technical Sciences, Associate Professor, Department of Computer Systems, Networks and Cybersecurity
National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine
ORCID 0000-0002-5151-9203
a.sagun@nubip.edu.ua

RESEARCH ON DIFFERENTIAL CRYPTANALYSIS BASED ON DEEP LEARNING

Abstract. In the age of pervasive connectivity, cryptography is a vital defensive measure for information security, and the security of cryptographic protection is of critical importance. Deep learning technology has recently made significant strides in areas like image classification and natural language processing, garnering considerable interest. Compared with classic cryptographic algorithms, modern block ciphers are more intricate, and the mappings between plaintext and ciphertext are less distinct, rendering the extraction of plaintext features from ciphertexts by neural networks as almost infeasible. However, the symbiosis of deep learning and traditional differential cryptanalysis holds promise for enhancing crypto-attack performance. Thus, the integration of deep learning theory and methods into the field of cryptography is becoming a significant trend in technological advancement. In this context, cryptanalysis is progressively developing in the direction of intelligence and automation, with an increasing number of researchers employing deep learning to assist in cryptanalytic tasks. This review aims to delve into the current research trends surrounding deep learning-supported differential cryptanalysis. It commences with a thorough recapitulation of differential analysis in cryptography and introduces common models in deep learning, along with their characteristics. Moreover, it encapsulates the design of differential classifiers powered by deep learning, inclusive of various optimization techniques utilized within these algorithms. The paper also posits directions for future research focus. Despite challenges, deep learning possesses vast potential in reinforcing conventional differential cryptanalysis, providing deeper insights for security analysis and response strategies, and serving as a valuable tool and perspective for the design and appraisal of future cryptographic solutions.

Keywords: deep learning; differential cryptanalysis; differential classifiers; convolutional neural network.

INTRODUCTION

With the development of computer technology and communication technology, information security has become an important factor affecting the development of a country

and society. Cryptography plays a pivotal role in ensuring information security, which is widely used in personal privacy protection, business trade, national defense security and other fields [1]. Cryptanalysis is an important branch of cryptography and an important part of intelligence analysis for military activities. The most representative of these statistical analysis methods is differential analysis.

CONTRIBUTION OF THIS INVESTIGATION

1) We present a comprehensive review of differential analysis in cryptography.
2) We outline deep learning models, such as Convolutional Neural Networks (CNNs) and differential perceptrons.
3) We provide an overview of the design of differential distinguisher models based on deep learning and discuss various optimization strategies adopted by these algorithms.
4) We explore different application areas of cryptographic analysis and identify key lessons for future research exploration.

BASIC PRINCIPLES OF DIFFERENTIAL CRYPTANALYSIS

Differential cryptanalysis is an effective cryptanalysis method for block ciphers, and whether a block cipher can successfully resist differential cryptanalysis has become an important index to measure the security of this cipher algorithm. The main idea of differential cryptanalysis is to obtain some guess information of the key by analyzing the probability non-uniformity of the ciphertext generated by the fixed input differential in the differential propagation, and then reduce the candidate key selection space.

In 1990, Israeli cryptographers Biham and Shamir [2] first proposed differential cryptanalysis, which belongs to the plaintext attack method and is often used to distinguish encrypted ciphertext from random data. Its basic idea is to find a differential path with high probability by analyzing the possible defects in the cryptographic algorithm, and use the differential path to build a differential distinguizer. Because of its characteristics, differential cryptanalysis is very effective in breaking iterative cryptosystems. Therefore, differential cryptanalysis is usually used as the breaking algorithm of iterative block ciphers, and it is also one of the important indicators to measure the security of ciphers, which plays an important role [3] in cryptanalysis and related fields of cryptographic security.

Definition 3.1: For random differential ciphertext pair data, its probability distribution is $P(\alpha, \beta) = \frac{1}{2^m}$.

Definition 3.2: When the input differential and output differential satisfy the given differential path, the input ciphertext is called a correct pair, otherwise it is called an error pair.

According to the above definition, when a $\gamma - 1$round difference is found and the probability is greater than $\frac{1}{2^m}$, the $\gamma - 1$ round fixed differential ciphertext pair can be distinguished from the random differential ciphertext pair. Using the differential distinguizer, the attack steps are summarized as follows:

Step 1: According to the block cipher that needs to be attacked, a differential distinguizer is designed to find $\gamma - 1$ the high probability differential characteristics of the round block cipher algorithm(α, β).

Step 2: According to the classification results of the above differential distinguisher, for all candidate keys,g_i, $0 \leq i \leq 2^l - 1$, count from 0 (is the length of the key).

Step 3: Randomly select the plaintext X and $X^* = X \oplus \alpha$,encrypt it with the same candidate key k to obtain the ciphertext the response Y and Y^*.

Step 4: Filter the obtained ciphertext pair, retain the filtered ciphertext pair(Y, Y^*), decrypt it with the keyg_i, and calculate the difference Δ, if $\Delta = \beta$ the candidate key count of the shadow is added by 1.

Step 5: Sort all the counters according to the size of the value, and select the corresponding key with the larger counter value as the candidate key value after screening.

Differential cryptanalysis is a chosen plaintext attack method. A differential distinguizer is designed to find the high probability differential feature in the encryption algorithm, and the fixed differential ciphertext pair is distinguished from the random ciphertext pair in the block cipher, and the candidate key is screened on this basis. Namely a cop partition can found a $\gamma - 1$ round high probability difference, it can be $\gamma - 1$ round of fixed difference cipher encryption algorithm and random ciphertext to separate, using the differential partition, the block cipher can be a candidate key filtering attacks. Differential cryptanalysis of the flow diagram is shown in Fig. 1.

According to the flow chart and basic steps of differential cryptanalysis, all ciphertexts obtained can be filtered once according to the output results of the differential distinguizer during differential cryptanalysis, which reduces the candidate key space, greatly reduces the number of keys that need to be guessed in the subsequent key recovery attack, and reduces the complexity of differential cryptanalysis. Therefore, how to design an effective differential distinguisher is the core step in differential cryptanalysis. The function of the differential distinguisher is to distinguish the fixed differential pair from the random differential ciphertext pair, which corresponds to the binary classification task in machine learning. Therefore, the deep learning method can be used to design a classifier to replace the traditional differential distinguisher, so as to further improve the performance of the differential distinguisher by taking advantage of the advantages of neural network in feature extraction and other aspects.

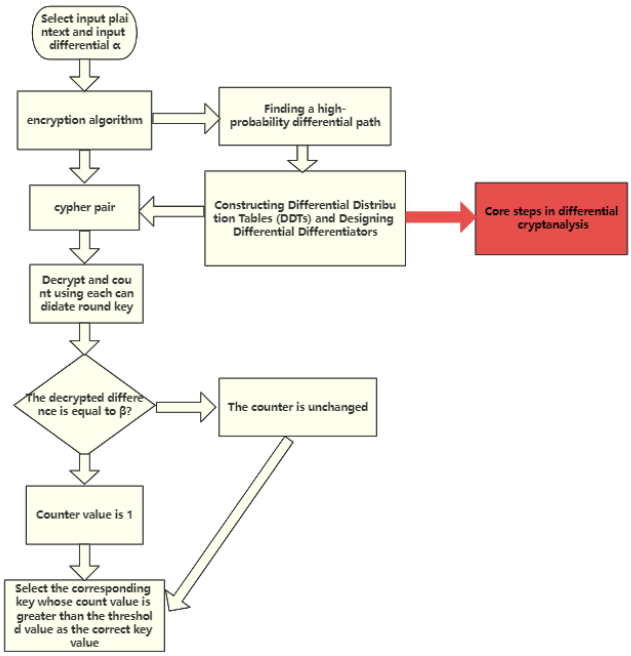

Fig. 1. Schematic diagram of differential cryptanalysis process

In all current cryptanalysis methods, the core idea of cryptanalysis is to design an effective cipher distinguisher, and use the classification results of the distinguisher to reduce the candidate key space, so as to reduce the difficulty of further cryptanalysis. It is often used to distinguish between plaintext and ciphertext to assist cryptanalysis. In the traditional differential analysis, the first thing is to find a high probability differential feature, and then construct the differential distinguisher through the high probability differential feature. The construction of the differential distinguisher depends more on the possible defects of the algorithm itself, and the construction process relies heavily on manual derivation, which greatly slows down the cryptanalysis process. In recent years, relying on automated search technology to find differential distinguishers has gradually become the mainstream method [4] of differential distinguisher construction.

In order to improve the accuracy of neural discriminators, researchers have explored two main directions. One of the popular directions is changing the format of the neural discriminator's input data, another ones — is using deep learning to build different neural networks.

DESIGN OF DIFFERENTIAL CLASSIFIER BASED ON DEEP LEARNING

There are a variety of machine learning algorithms, such as support vector machine algorithm, naive Bayes algorithm, decision tree algorithm, expectation maximization algorithm, artificial neural network algorithm, and so on. Now deep learning has become a research hotspot in machine learning. The common models include multilayer perceptron, deep neural network, convolutional neural network, recurrent neural network, long short-term memory network, etc [5]. Convolutional neural network is suitable for many fields such as natural language processing, speech processing and computer vision. Recurrent neural networks have great advantages in processing sequential information and speech.

Deep Learning models

1) Convolutional Neural Network (CNN). At the end of the 20th century, convolutional neural networks began to appear in people's field of vision. With the concept of deep learning proposed, its related applications have been developed rapidly, and significant results have been achieved in many fields [6]. Fig. 2 shows the specific composition of a convolutional neural network. It mainly consists of five parts, and the corresponding explanations are as follows: at the Input layer the data samples processes; at Convolutional layer extracts features from the data and scans the entire sample vector space through the convolution kernel, which is a smaller matrix. At Pooling layer: after the Convolutional layer, it mainly selects the features obtained in the previous step and filters the information. At Fully connected layer: belongs to the most terminal in the network, and further performs nonlinear combination of the extracted features to obtain the output. At Output layer solves different problems, the output is not the same.

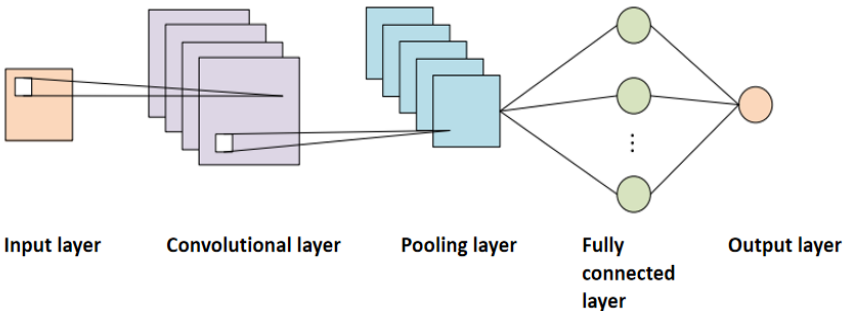

Fig. 2. Convolutional neural network

ISSN 2663 - 4023

2) Multilayer Perceptron (MLP) (classifier). Multilayer perceptron consists of three parts: input layer, hidden layer and output layer. The number of hidden layers can be one or more. The simplest structure is only one hidden layer (as shown in fig. 3). Each layer of the multilayer perceptron is composed of one or more perceptron units, and each layer is fully connected, that is, all neurons in each layer are connected to all neurons in the next layer. Each perceptron is connected by the weight and output signal, and as the input of the next layer of network perceptron, so the multilayer perceptron is composed of multiple perceptron units. The input of MLP is a vector (array) through the form of full connection to each element of the overall array layer by layer to give weight and obtain the final classification. This method is a rough learning method, directly learn all elements of the direct linear or nonlinear correlation, but did not go to the depth of the array of better performance features, classification effect is not good.

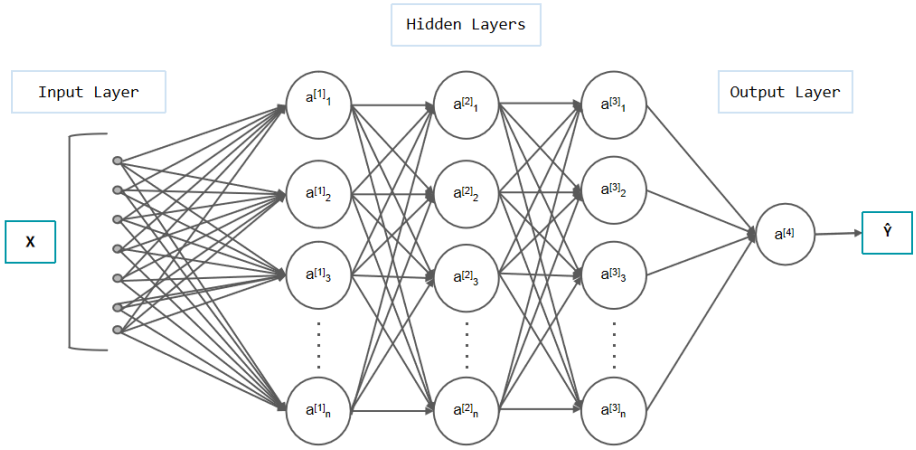

Fig. 3. Multilayer perceptron diagram

3) Residual Network (ResNet). In previous work, it has been believed that the deeper the network, the more things the network can learn. However, after a large number of experiments, when the number of layers of the convolutional neural network increases to a certain level, its accuracy decreases, which is called network degradation problem [7]. Therefore, in 2015, He Kai-ming et al. proposed the residual network. In order to solve the degradation problem in deep networks, some layers of the neural network can be artificially made to skip the connection of the next layer of the neural network, connect between layers, and weaken the strong relationship between each layer. It solves the problem that the deep CNN model is difficult to achieve in the training process. Fig. 4 shows the structure diagram of the residual network:

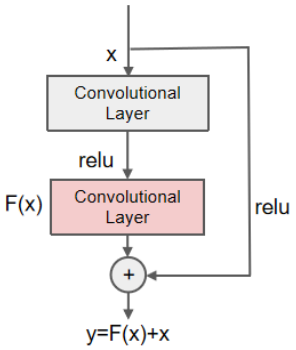

Fig. 4. Residual network

Deep learning-based differential classifier

Traditional machine learning algorithms are usually difficult to achieve good performance in solving complex problems due to computational bottlenecks, expert knowledge limitations and other reasons. The emergence of deep learning provides solutions for cryptographic researchers, using the advantages of different neural network structures to build effective differentiators. The successfully trained distinguisher model can effectively distinguish between random data and encrypted data, and the neural network can give full play to its own advantages when dealing with large-scale data. Therefore, the construction of an effective neural distinguisher model has significant practical significance and research value for the security of cryptanalysis algorithms.

The training of neural network differential classifier is a process of mining and extracting the features of plaintext data and classifying them, so it is necessary to generate training data in advance. When performing differential cryptanalysis, all the ciphertexts obtained can be filtered once according to the output of the differential classifier, so as to reduce the space of candidate keys. So that the number of keys that need to be guessed in the subsequent key recovery attack is greatly reduced, and the complexity of differential cryptanalysis is reduced. This way, the designing of an effective differential distinguisher is the core step in differential cryptanalysis. The function of differential distinguisher is to distinguish the fixed differential pair from the random differential ciphertext pair, which corresponds to the binary classification task in machine learning. Therefore, the deep learning method can be used to design a classifier to replace the traditional differential distinguisher. The using a neural network for future feature extraction and in some other aspects allows improving the efficiency of the classifier.

In recent years, based on the goal of improving the accuracy of neural distinguisher, scholars have proposed a series of improvements. These improvements include using more ciphertext pairs as input and using more complex and powerful neural networks. Gohr proposed a neural discriminator based on residual neural network for Speck32/64. Combined with Bayesian optimization, Gohr further proposed a deep learning-based key recovery attack, which was successfully applied to 11-round and 12-round Speck32/64. Compared with the traditional cryptanalysis method, this new attack includes two additional operations in addition to decryption. The first operation is to send the decrypted ciphertext pair to the neural distinguisher and obtain the output of the neural distinguisher. The second operation relies on Bayesian optimization to recommend a batch of key guesses that are most likely to be correct for verification at each iteration. In addition, Gohr adopts a reinforcement learning mechanism to dynamically allocate computing resources. During the attack, a batch of ciphertext structures is generated for multiple iterations.

In each iteration, a ciphertext structure with the highest probability of matching the correct ciphertext structure is selected to verify the guessed key. Once the maximum number of iterations is reached, a batch of new ciphertext structures are generated and iterated again. It is not difficult to find that the key recovery attack based on deep learning proposed by Gohr is very different from the classical cryptanalysis method, and its time complexity is also affected by more factors.

The neural network used by Baksi [8] is an MLP network, which was found to provide the best accuracy and can be tuned in a very fast time. The best architecture is an MLP that uses three hidden layers with 1,024 neurons each. The activation function is either LeakyReLU or ReLU. The paper notes that using LeakyReLU as the activation function for a neuron allows a small positive gradient when the neuron is inactive (for example, when the input is negative). This is considered more balanced and is an advantage over traditional activation functions like

ReLU for smaller networks. However, this function still performs slightly worse for networks with many parameters.

Aayush Jain [9] simply changed the architecture of the MLP to two layers based on Baksi and claimed that using only 2 hidden layers in an MLP network can produce results in less time and has a better chance to avoid overfitting the data. The results are also compared with Baksi and indeed improved.

AdrienBenamira [10] argued that the MLP module in Gower's network model was not necessary. In this paper, the LGBM (Lightweight Gradient Boosting decision Tree) model was replaced with the MLP module. The specific operation is to retain the first layer of MLP and use its output as the input of LGBM. The table of experiments shows that the accuracy of LGBM is close to that of Gower's model, and is better than that of random forest, support vector machine, etc.

The neural network in YiChen [11] (in addition to the input and output layers) is divided into two large modules. The first module is used to extract features for ciphertext pairs. The structure is a convolutional layer with 32 filters and a kernel size of 1x1, followed by a batch normalization layer and a ReLU activation function (called module1). The second module is used for probability estimation. The second module is divided into 4 small blocks: 1) module2 consisting of several initial residual blocks (here, only one convolutional layer with 32 filters and 3×3 kernels); 2) module3 consisting of a fully connected layer followed by a batch normalization layer and ReLU activation function; 3) module4 with 643 neurons and a fully connected layer followed by a relu activation function; 4) module5 with 644 neurons and a fully connected layer followed by a sigmoid activation function. The experiments show that under different *k*-settings, our neural discriminator can always obtain the improvement of the discrimination accuracy. When the *k*-samples misclassified by the baseline discriminator are combined into one group, our neural discriminator can still correctly distinguish the probabilities with a non-negligible probability. This indicates that the neural discriminator successfully captures these features, and the improvement in accuracy of discrimination accuracy also comes from the derived features.

Zezhou Hou [12] chose ResNet (Residual Network) as the neural network. This residual structure uses “ReLU” as the activation layer, and “Conv1D” as the basic convolutional layer. More importantly, the built ResNet hidden layer contains a total of 5 residual towers (which are the five residual towers in Gower’s final code). The original data is first formatted and calculated by the “Conv1D” layer, then transmitted to the structure, and finally the final result is output through the output layer.

There is an approach where, based on SAT, the neural network distinguisher is extended from 9 rounds to 11 rounds. For such an approach, a last sub-key recovery attack on SIMON32 for 13 rounds is proposed using 212.5 chosen ciphertexts, with a success rate of over 90%. Compared with the traditional methods, the deep learning-based method has lower time complexity and data complexity.

In Wenqiang Tian’s paper [13], the residual structure was adjusted, that is, the activation function and the batch normalization layer were adjusted. Together with the original version, a total of five versions of residual structure were generated. According to the previous test results, the batch normalization layer was the first, the activation function was the second, and the convolution layer was finally linked with the lowest error rate and the best performance. Then, the five structures are used to construct the differential distinguisher. However, the best performing model does not show the highest accuracy in cryptanalysis. In this paper, we offer our own judgment:

1) Training accuracy does not necessarily reflect test accuracy. The model with the highest training accuracy is usually not the one with the highest validation accuracy, the training accuracy can only reflect the performance of the network on the training data and only the validation accuracy can reflect the performance of the network on new data. Therefore, the model with the highest validation accuracy should be selected. We speculate that Gower chose neural Net(c) because this Net(c) has high training accuracy for the 5-round distinguizer of SPECK32/64 block cipher [14]. Based on our experimental results, it is inferred that other networks can achieve better results on the same task.
2) Different networks have a significant impact on the training results. For the 9-round SIMON32 cipher's distinguisher, the highest network accuracy is nearly 0.022 higher than the lowest network accuracy. The result is in about 8% fewer chosen plaintexts are needed in the attack.
3) The optimal network may be different for different encrypted texts, or even for different rounds of the same encrypted segments or different input differences. This inspires us not to rush to apply a certain model to other rounds or even other ciphers just because it performs well on an individual n rounds.
4) Different random training data has little effect on the training results. For each discriminator, we randomly generated 10 experiments of different chosen plaintexts, with only negligible differences in the final accuracy.

The paper also gives some reasons why we chose resnet: in differential cryptanalysis we want the neural network to be able to learn the characteristics of the ciphertext difference obtained by XOR for ciphertext pairs; residual neural networks have been shown to perform this task well.

We tried various network models, such as fully connected networks and convolutional neural networks and residual neural networks performed well. Gower did similar work and also obtained the best results using residual neural networks.

Runlian Zhang [15] only mentioned that there are 6 hidden layers in the network, and the number of neurons in each layer is 512, 128, 64, 32, 16 and 2, respectively. The input layer needs to receive ciphertext pairs. Since the block length of TweGIFT-128 cipher is 128 bits, there are 256 neurons in the input layer. In order to improve the curve fitting ability of the model, a nonlinear activation function called LeakyReLU. This one uses to reduce the probability of gradient disappearance during training.

GaoWang [16] reported experimenting with 10 different machine learning model types, including AdaBoost (AB), Decision Tree (DT), K-Nearest Neighbor (KNN), Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), Multi-Layer Perceptron (MLP), Long Short-Term Memory (LSTM), and a series of experiments. Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN). In this paper, the effect of depth on MLP, LSTM, RNN and CNN is explored by varying the number of layers. MLP1 is an MLP model which has only input and output layers. Unlike MLP1, MLP2 has one additional hidden layer, while MLP3 has two additional hidden layers of neurons. LSTMS, CNNS, and RNNS are similar.

In machine learning, hyperparameters (hyperopt) are referred to as the parameters of a model whose values are set before the training process begins. Hyperopt is employed to find the best parameters for each type of model. We define and compare Baksi's multilayer perceptron model as MLP0. The experimental results show that the accuracy of AB, DT, KNN, LR, RF, SVM and MLP0 is lower than that of MLP, LSTM, RNN and CNN, which indicates that MLP, LSTM, RNN and CNN are more suitable for building a differential classifier. At the

same time, considering the speed issue, LSTM and RNN are not suitable, because the recurrent layer of LSTM requires high memory and computation, and the long-term dependence makes the speed of RNN too slow. Also, a model with only one input-output layer is enough to get the best results with the help of Bayesian optimization, so there is no need to consider deeper models. The deeper the model, the longer it will take to build the classifier. Therefore, CNN1 is used for experiments in this paper.

In this paper [17] Emanuele Bellini reviews two distinguishers. The first one is called time distributed distinguisher, and there is no innovation in input/output and loss function. The innovation is that the hidden layer is divided into two parts: the first part is the time distributed network of, and the second part is the multilayer perceptron in the classical definition. In the first part, the input is split into four 32-bit blocks, each representing one of the four blocks that make up the two ciphertexts, and we pass each block separately to two 32-neuron dense layers (in our case, the perceptron). The name "temporal distributed" comes from a method that is common when dealing with temporal data, whereas in this network it is treated as if the blocks were processed separately, without letting their values influence each other. The output is four 32-bits vectors, which are flattened and concatenated into a 128-bits vector. This 128-bits vector will be the input for the second part. The second part consists of three fully connected layers of 64, 64 and 32 neurons that eventually go into the output layer. The idea of splitting the network into two parts comes from the fact that in both ciphers, the output is calculated separately as two different parts. The experimental results prove that the distinguisher outperforms the traditional discrimination by a considerable margin. We also demonstrate that we also show that these results can be achieved without excessive computational power in the round reduced version of the cipher.

Heng-Chuan Su's network [18], like Gower's, is single-bit slice convolutional with residual structure with dense connected layer. This layer functions as the main prediction structure. The innovation lies in the combination of topological structure and neural network.

The difference between topological differential neural network discriminator and differential neural network discriminator is the input data and the dense layer. Experiments show that these results can also have good results in the case of low data volume. The differential classifier based on neural network. It has good accuracy on 6-round Simon32/64. When the input difference is constant as Δ = 0x0000/0x0008, the success rate of the distinguisher gradually decreases with the increase of the number of rounds. On 10-round Simon32/64 cipher, the analyze success rate is close to 0.5.

SUMMARY OF THIS ARTICLE

The security analysis of a lightweight block cipher is mainly determined by the quality of its distinguisher model. Indeed, as deep learning technology advances, cryptographic researchers worldwide are increasingly turning their attention to employing deep learning methods for constructing distinguisher models for cryptanalyze applying. Currently, it is promising to explore the methods of constructing a differentiator based on deep learning and its influence on the parameters of the neural network model. Given the application of deep learning models in the differential cryptanalysis of block ciphers, we find it logical to choose and describe the method of constructing a differentiator based on deep learning. Following this, we will proceed with training a neural network differentiator based on CNN, MLP, and ResNet models. Furthermore, the discriminators trained by the two methods were compared.

In the experiments, it was found that different neural network models have different effects during training on the differentiator for cryptographic algorithms. Since the construction of the differential classifier is a precomputation process in cryptanalysis, once it is trained, it can be used continuously in the subsequent cryptanalysis. Therefore, even if a large amount of time is spent training the neural network differential distinguisher to improve the accuracy of the distinguisher in the early stage, it is still effective for cryptanalysis.

FURTHER RESEARCH DIRECTIONS

Although deep learning has a certain application in the field of cryptography, its application in the field of cryptanalysis is still in its infancy, and there are many problems to be solved. Summing up all the above, authors believe that the following problems need to be further studied:

1) The combination of ciphertext pair and ciphertext difference is used as the input of the neural network model, although the accuracy and the number of rounds of the distinguisher are improved to a certain extent, it can be further optimized for the data set. In the future work, the idea of multiple differential cryptanalysis can be introduced into the differential neural distinguisher model, and the ciphertext groups and ciphertext difference groups can be generated through several input differences with high probability as the input of the model, and the key recovery attack can be constructed on the basis of this work.
2) The deep learning model is often regarded as a black box, which makes it difficult for us to deeply understand its internal logic and decision-making process. Therefore, researchers may not be able to accurately interpret the output of the model when performing differential cryptanalysis, which reduces the credibility of the security assessment.
3) Current researches on deep learning-based block cipher cryptanalysis focus on neural differential cryptanalysis, and no results have been found on the combination of deep learning and other traditional cryptanalysis methods. Differential cryptanalysis is inevitably and more calculated difficult for new ciphers. In summary, the combination of deep learning with other cryptanalysis methods is a worthy research direction.
4) At this stage the data used to train the differentiator is generated by random methods then differential computation, and then encrypted to obtain the training dataset. There are many advanced dataset processing methods in the field of artificial intelligence, and whether these methods can be introduced into the construction of cryptographic differentiator datasets is a question that this paper plans to investigate in the future.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Maurer, U., et al. (2007). *Information Security and Cryptography*.
2. Biham, E., Shamir, A. (1991). Differential cryptanalysis of DES-like cryptosystems. *J. Cryptology, 4,* 3–72. https://doi.org/10.1007/BF00630563
3. Sarker, I. (2021 Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Comput. Sie. 2, 420*. https://doi.org/10.1007/s42979-021-00815-1

4. Gohr, A. (2019). Improving Attacks on Round-Reduced Speck32/64 Using Deep Learning. *Advances in Cryptology–CRYPTO 2019: 39th Annual International Cryptology Conference, 39*, 150–179. https://doi.org/10.1007/978-3-030-26951-7_6
5. Zhao, J., et al. (2018). Differential Analysis of Lightweight Block Cipher GIFT. *Journal of Cryptologic Research, 5(4),* 5–13. https://doi.org/10.13868/j.cnki.jcr.000244
6. Kattenborn, T., et al. (2021). Review on Convolutional Neural Networks (CNN) in Vegetation Remote Sensing. *ISPRS Journal of Photogrammetry and Remote Sensing, 173*, 24–49. https://doi.org/10.1016/j.isprsjprs.2020.12.010
7. He, K., et al. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778. https://doi.org/10.48550/arXiv.1512.03385
8. Baksi, A., & Baksi, A. (2022). Machine Learning-Assisted Differential Distinguishers for Lightweight Ciphers. *Classical and Physical Security of Symmetric Key Cryptographic Algorithms*, 141–162.
9. Jain, A., Kohli, V., & Mishra, G. (2020). Deep Learning Based Differential Distinguisher for Lightweight Cipher PRESENT. *Cryptology ePrint Archive*.
10. Benamira, A., et al. (2021). A Deeper Look at Machine Learning-Based Cryptanalysis. *Cryptology–EUROCRYPT 2021: 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques*, 805–835. https://doi.org/10.1007/978-3-030-77870-5_28
11. Chen, Y., & Yu, H. (2021). A New Neural Distinguisher Model Considering Derived Features from Multiple Ciphertext Pairs. *IACR Cryptol. ePrint Arch., 310.*
12. Hou, Z., Ren, J., & Chen, S. (2021). Cryptanalysis of Round-Reduced SIMON32 Based on Deep Learning. *Cryptology ePrint Archive.*
13. Yadav, T., & Kumar, M. (2021). Differential-ml Distinguisher: Machine Learning Based Generic Extension for Differential Cryptanalysis. *International Conference on Cryptology and Information Security in Latin America*, 191–212.
14. Gohr, A. (2019). Improving Attacks on Round-Reduced Speck32/64 Using Deep Learning. Advances in Cryptology–CRYPTO 2019: 39th Annual International Cryptology Conference, 39, 150–179. https://doi.org/10.1007/978-3-030-26951-7_6
15. Zhang, R., et al. (2021). Differential Cryptanalysis of TweGIFT-128 Based on Neural Network. *2021 IEEE Sixth International Conference on Data Science in Cyberspace (DSC)*, 529–534. https://doi.org/10.1109/DSC53577.2021.00084
16. Wang, G., Wang, G., & He, Y. (2021). Improved Machine Learning Assisted (Related-key) Differential Distinguishers For Lightweight Ciphers. *2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, 164–171.
17. Bellini, E., & Rossi, M. (2021). Performance comparison between deep learning-based and conventional cryptographic distinguishers. *Intelligent Computing: Proceedings of the 2021 Computing Conference*, 3, 681–701.
18. Su, H., Zhu, X., & Ming, D. (2021). Polytopic Attack on Round-Reduced Simon32/64 Using Deep Learning. *Information Security and Cryptology: 16th International Conference*, 3–20.

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 3 (23), 2024
ISSN 2663 - 4023

Цзян Сюе
аспірант
Національний університет біоресурсів і природокористування України, Київ, Україна
ORCID 0009-0000-1676-2331
jx1283@163.com

Лахно Валерій Анатолійович
Доктор технічних наук, професор, професор кафедри комп'ютерних систем, мереж та кібербезреки
Національний університет біоресурсів і природокористування України, Київ, Україна
ORCID 0000-0001-9695-4543
lva964@nubip.edu.ua

Сагун Андрій Вікторович
Кандидат технічних наук, доцент, доцент кафедри комп'ютерних систем, мереж та кібербезреки
Національний університет біоресурсів і природокористування України, Київ, Україна
ORCID 0000-0002-5151-9203
a.sagun@nubip.edu.ua

ДОСЛІДЖЕННЯ ДИФЕРЕНЦІАЛЬНОГО КРИПТОАНАЛІЗУ НА ОСНОВІ ГЛИБОКОГО НАВЧАННЯ

Анотація. В епоху глобального домінування комп'ютерних систем та мереж криптографія є життєво важливим засобом захисту інформації, а безпека криптографічного захисту має вирішальне значення. Технологія глибокого навчання нещодавно досягла значних успіхів у таких сферах, як класифікація зображень і обробка природної мови, викликаючи значний інтерес у дослідників. Порівняно з класичними криптографічними алгоритмами, сучасні блокові шифри є складнішими, а відображення між відкритим текстом і зашифрованим текстом менш чіткі. Це робить вилучення функцій відкритого тексту із зашифрованих текстів нейронними мережами майже неможливим. Однак симбіоз глибокого навчання та традиційного диференційного криптоаналізу є перспективним для підвищення ефективності криптоаналізу. Таким чином, інтеграція теорії та методів глибокого навчання в область криптографії стає важливою тенденцією технологічного прогресу. У цьому контексті криптоаналіз стрімко розвивається у напрямку інтелектуалізації та автоматизації. Відповідно у цьому напрямку зростає кількість дослідників, які використовують глибоке навчання для покращення розв'язання криптоаналітичних завдань. Мета цієї оглядової роботи — заглибитися у поточні тенденції досліджень навколо диференціального криптоаналізу з підтримкою глибокого навчання. Він починається з ретельного повторення диференційного аналізу в криптографії та представляє загальні моделі глибокого навчання разом із їхніми характеристиками. Крім того, він інкапсулює дизайн диференціальних класифікаторів на основі глибокого навчання, включаючи різні методи оптимізації, що використовуються в цих алгоритмах. У документі також визначені напрямки майбутніх досліджень. Попри означені проблеми, глибоке навчання має величезний потенціал у зміцненні традиційного диференційного криптоаналізу, забезпечуючи більш глибоке розуміння для аналізу безпеки та стратегій реагування, а також слугуючи цінним та перспективним інструментом для розробки та оцінки майбутніх криптографічних рішень.

Ключові слова: глибоке навчання; диференціальний криптоаналіз; диференціальні класифікатори; згорнута нейронна мережа.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Maurer, U., et al. (2007). *Information Security and Cryptography*.
2. Biham, E., Shamir, A. (1991). Differential cryptanalysis of DES-like cryptosystems. *J. Cryptology*, *4*, 3–72. https://doi.org/10.1007/BF00630563

3. Sarker, I. (2021 Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Comput. Sie. 2, 420*. https://doi.org/10.1007/s42979-021-00815-1
4. Gohr, A. (2019). Improving Attacks on Round-Reduced Speck32/64 Using Deep Learning. *Advances in Cryptology–CRYPTO 2019: 39th Annual International Cryptology Conference, 39,* 150–179. https://doi.org/10.1007/978-3-030-26951-7_6
5. Zhao, J., et al. (2018). Differential Analysis of Lightweight Block Cipher GIFT. *Journal of Cryptologic Research, 5(4),* 5–13. https://doi.org/10.13868/j.cnki.jcr.000244
6. Kattenborn, T., et al. (2021). Review on Convolutional Neural Networks (CNN) in Vegetation Remote Sensing. *ISPRS Journal of Photogrammetry and Remote Sensing, 173,* 24–49. https://doi.org/10.1016/j.isprsjprs.2020.12.010
7. He, K., et al. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* 770–778. https://doi.org/10.48550/arXiv.1512.03385
8. Baksi, A., & Baksi, A. (2022). Machine Learning-Assisted Differential Distinguishers for Lightweight Ciphers. *Classical and Physical Security of Symmetric Key Cryptographic Algorithms,* 141–162.
9. Jain, A., Kohli, V., & Mishra, G. (2020). Deep Learning Based Differential Distinguisher for Lightweight Cipher PRESENT. *Cryptology ePrint Archive*.
10. Benamira, A., et al. (2021). A Deeper Look at Machine Learning-Based Cryptanalysis. *Cryptology–EUROCRYPT 2021: 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques,* 805–835. https://doi.org/10.1007/978-3-030-77870-5_28
11. Chen, Y., & Yu, H. (2021). A New Neural Distinguisher Model Considering Derived Features from Multiple Ciphertext Pairs. *IACR Cryptol. ePrint Arch., 310.*
12. Hou, Z., Ren, J., & Chen, S. (2021). Cryptanalysis of Round-Reduced SIMON32 Based on Deep Learning. *Cryptology ePrint Archive.*
13. Yadav, T., & Kumar, M. (2021). Differential-ml Distinguisher: Machine Learning Based Generic Extension for Differential Cryptanalysis. *International Conference on Cryptology and Information Security in Latin America,* 191–212.
14. Gohr, A. (2019). Improving Attacks on Round-Reduced Speck32/64 Using Deep Learning. Advances in Cryptology–CRYPTO 2019: 39th Annual International Cryptology Conference, 39, 150–179. https://doi.org/10.1007/978-3-030-26951-7_6
15. Zhang, R., et al. (2021). Differential Cryptanalysis of TweGIFT-128 Based on Neural Network. *2021 IEEE Sixth International Conference on Data Science in Cyberspace (DSC),* 529–534. https://doi.org/10.1109/DSC53577.2021.00084
16. Wang, G., Wang, G., & He, Y. (2021). Improved Machine Learning Assisted (Related-key) Differential Distinguishers For Lightweight Ciphers. *2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom),* 164–171.
17. Bellini, E., & Rossi, M. (2021). Performance comparison between deep learning-based and conventional cryptographic distinguishers. *Intelligent Computing: Proceedings of the 2021 Computing Conference,* 3, 681–701.
18. Su, H., Zhu, X., & Ming, D. (2021). Polytopic Attack on Round-Reduced Simon32/64 Using Deep Learning. *Information Security and Cryptology: 16th International Conference,* 3–20.