차분 분석을 위한 딥러닝 기반의 신경망 구별자 연구 동향

김현지*, 임세진*, 강예준*, 김원웅*, 서화정*†

*한성대학교 (대학원생)
*†한성대학교 (교수)

Deep Learning-based neural distinguisher research trend for differential cryptanalysis

Hyunji Kim*, Sejin Lim*, Yeajun Kang*, Wonwoong Kim*, Hwajeong Seo*†

*Hansung University(Graduate student)
*†Hansung University(Professor)

요 약

차분 분석은 블록암호에 대한 대표적인 암호 분석 기법이다. 차분 분석의 데이터 복잡도를 줄이기 위해 구별자 (Distinguisher)가 사용되며, 인공 신경망은 이를 위한 좋은 솔루션이 될 수 있다. 딥러닝 기반의 구별자를 신경망 구별자라고 하며, CRYPTO 2019에서 제안된 Gohr[1]의 연구를 기반으로 하여 다양한 암호 알고리즘과 입력 차분, 그리고 신경망 구별자 모델 구조에 대한 연구들이 진행되고 있다. 본 논문에서는 이러한 신경망 구별자에 대한 연구 동향에 대해 소개하였다. 향후에는 현재 딥러닝 기반의 신경망 구별자가 갖는 메모리 제한 문제와 라운드 축소된 암호에 대한 분석만 가능하다는 한계점을 극복하기 위한 방법론에 대한 연구가 필요할 것으로 생각된다.

I. 서론

차분 분석은 블록 암호에 대한 암호 분석 기법 중 하나이다. 암호 알고리즘을 안전하지 않게 설계한 경우, 입력 차분에 따른 출력 차분이 존재하며, 이를 활용하면 사용된 키를 유추할 수 있다. 만약 차분 분석의 과정 중 하나인 1라운드에 대한 복호화에서 무차별 대입에 사용 되는 키가 올바른 키라면 입력 차분에 대한 출력 차분이 높은 확률로 만족될 것이고, 잘못된 키를 사용하면 출력 차분을 만족하지 못하게 된다. 이를 통해 옳은 라운드 키를 사용하였는지 확률적으로 알아낼 수 있다. 이때, 차분 분석에 사용되는 데이터 복잡도를 감소시킬 수 있는 딥러닝 기반의 분류기를 신경망 구별자라고 하며, 이는 특정 차분을 갖는 암호문 데이터와 랜덤 데이터를 구별하는 작업을 수행한다.

II. 관련 연구

2.1 딥러닝 기반의 신경망 구별자

차분 분석은 블록암호에 대한 대표적인 분석 방법이다. 차분 분석은 선택된 평문 공격에서 사용할 수 있는 공격 기법이며, 입력 차분과 출력 차분을 활용한다. 입력 차분이란, 하나의 평문 쌍이 있을 때 이 두 평문을 XOR한 것이고, 출력 차분은 해당 평문 쌍을 암호화한 암호문 쌍을 XOR한 값이다. 이상적인 암호 알고리즘의 경우, 입력 차분이 있는 평문 쌍을 암호화하면 암호문 간의 특정 차분이 나오지 않고 균일하다. 그러나, 안전하지 않게 설계된 암호 알고리즘의 경우, 입력 차분을 갖는 평문 쌍을 암호화하면 특정 출력 차분을 갖게 된다. 이처럼 차분 분석은 안전하게 설계되지 않은 암호 알고리즘의 경우, 입력 차분에 대한 출력 차분이 유지된다는 성질을 이용하여 다음과 같이 수행된다. 차분분석은 특정 입력 차분을 갖는 평문 쌍을 암호화한 r-1라운드 암호문쌍의 출력 차분을 활용한다. R라운드의 암호문을 얻은 후, r라운드의 라운드키를 무차별 대입하여 r-1라운드의 암호문

을 얻는다. 즉, 1라운드 복호화를 수행한다. 만약 옳은 라운드키를 사용했다면 높은 확률로 r-1라운드의 출력 차분을 만족할 것이고 틀린 키인 경우 출력 차분을 만족하지 않게 된다. 이 때, 무작위 암호문 데이터와 출력 차분을 갖는 암호문 데이터를 분류할 수 있다면, 데이터의 복잡도를 감소시킬 수 있다. 인공지능은 이러한 작업에 대한 좋은 솔루션이다. 기본적인 딥러닝 기반의 신경망 구별자는 다음과 같이 동작한다. 먼저, 특정 입력 차분을 갖는 평문 쌍을 암호화 한 암호문 쌍을 신경망에 입력한다. 이때, 암호문 쌍은 랜덤 평문 쌍을 암호화한 랜덤 데이터와 차분을 갖는 평문 쌍을 암호화한 암호문 데이터이며, 랜덤 데이터는 0으로 라벨링 되고 암호문 데이터는 1으로 라벨링한다. 이후, 해당 데이터를 신경망에 입력하면 신경망은 이를 랜덤과 암호문 데이터로 분류하는 이진 분류 문제를 해결한다. 해당 구조를 응용하여 여러 입력 차분을 갖는 신경망 구별자를 설계할 수 있으며 이에 대한 다양한 연구들이 수행되고 있다.

III.인공 신경망 기반의 구별자 연구 동향

본 논문에서는 인공 신경망 기반의 구별자에 관한 연구 동향에 살펴본다. Table 1은 관련 연구들에 대한 세부 사항이다. CRYPTO 2019에서 최초로 라운드 축소된 SPECK에 대한 딥러닝 기반의 신경망 구별자가 제안되었다[1]. 이후, 해당 연구를 기반으로 다양한 암호와 입력 차분, 신경망 구조 등에 대한 연구들이 활발히 진행되고 있으며, 현재는 모든 연구가 라운드 축소 버전에 대한 결과이고 전체 라운드에 대한 분석은 아직까지 수행되지 않았다. Gohr[1]에서는 7라운드 speck32/64에 대한 신경망 구별자가 처음으로 제안되었다. 그들은 딥러닝 기반의 분류기를 사용하여 무작위 데이터와 여러 입력 차분을 구별하였으며, 딥러닝 기반이 아닌 기존의 구별자에 비해 키 복구 공격에 대한 데이터 복잡성을 감소시켰다. 8라운드 speck에 대한 신경망 구별자의 정확도는 약 0.514이며, 0.5 이상의 정확도를 얻었기 때문에 무작위 데이터로부터 암호문 데이터를 분류할 수 있음을 확인할 수 있다. 해당 연구가 발표된 이후, Gohr가 제안한 딥러닝 기반의 신경망 구별자를 활용하여 SIMON 암호 알고리즘에 대한 차분 분석이 수행되었다[2].

[3]에서는 Gohr의 신경망 구별자와는 다른 모델을 제안하였다. [1]에서는 하나의 암호문 쌍 (두 개의 암호문)을 입력 데이터로 사용하였다면, 해당 연구에서는 k개 ($k>1$)의 암호문 쌍을 입력 데이터로 사용하였다. 즉, 특정 입력 차분을 갖는 평문 쌍 k개를 암호화 한 k개의 암호문 쌍을 신경망 구별자에 입력하고, 이를 분류하도록 한 것이다. 대상 암호는 Speck, Chaskey, Present, DES이며, Table 1에서 볼 수 있듯이 k값에 따른 실험을 진행했고 모든 경우에 대해 0.5 이상의 정확도를 얻었다.

Eurocrypt2021에서는 [1]에 대한 분석이 공개되었고[4], 해당 논문의 저자들은 5라운드 Speck에 대해 여러 입력 차분에 대해 실험을 진행하였다. Gohr에서 사용한 5라운드 Speck의 입력차분인 0x00400000 보다 더 높은 확률을 갖는 차분 특성이 있음을 분석하였고, 해당 차분과 더불어 상위 25개의 입력 차분에 대한 실험을 진행하였다. 또한, 딥러닝을 사용하지 않는 고전 구별자와의 공정한 비교를 위해 암호문 쌍을 입력데이터로 사용하지 않고, 암호문 쌍을 XOR한 출력 차분을 신경망에 입력하도록 설계하였다. 실험 결과, 대부분의 경우 75% 이상의 정확도를 얻었고 0x28000010에 대해서는 90% 이상의 정확도를 달성하였다.

Baksi et al.[5]에서는 Gohr의 연구에서 영감을 받아 딥러닝 기반의 신경망 구별자를 설계하였다. 그러나 대상 암호가 GIMLI, ASCON, KNOT, CHASKEY이며 다중 입력 차분과 단일 차분을 고려한 두 가지 모델을 제안하였다. 다중 입력 차분에 대한 신경망 구별자는 각 입력 차분들을 클래스로 설정하여 다중 분류를 수행하고 단일 입력 차분에 대한 신경망 구별자는 무작위 데이터와 암호문 데이터에 대한 이진 분류를 수행한다. 해당 연구에서는 GIMLI-Permutation에 대해 MLP, CNN, LSTM을 사용하여 실험을 수행하였고, 해당 작업에서 MLP가 가장 좋은 성능을 보였으며 CNN은 0.5보다 높은 정확도를 얻을 수 없었다. 따라서 해당 연구에서는 다른 암호들에 대해 MLP 모델만을 사용하였으며, 해당 연구 외에도 대부분 MLP가 사용됨을 Table 1에서 확인할 수 있다.

그러나 기존의 딥러닝 기반 신경망 구별자는

Table 1 Details of studies on the artificial neural network-based distinguisher

Reference	Cipher	Round	Input difference	Accuracy	Neural network
[1]	Speck (32-bit)	8	0x00400000	0.514	CNN1D + MLP
[3]	PRESENT (64-bit)	10	0x0000000000000009	0.5503, 0.5853, 0.5786, 0.5818 (K=2,4,8,16)	CNN2D + MLP
[3]	Speck (32-bit)	7	0x00400000	0.6393, 0.6861, 0.7074, 0.6694 (K=2,4,8,16)	
[3]	Chaskey (128-bit)	4	0x00008400000004000000000000000000	0.6589, 0.6981, 0.7603, 0.7712 (K=2,4,8,16)	
[3]	DES (64-bit)	6	0x4008000004000000	0.5653, 0.5568, 0.5507, 0.5532 (K=2,4,8,16)	
[4]	Speck (32-bit)	5	0x28000010, 0x802AD4A8, 0x802ED4AC etc.	0.75 이상	CNN2D
[5]	GIMLI permutation (384-bit)	8	Mask=1 (4, 12번째 바이트에 XOR)	0.5689	MLP, CNN, LSTM
[5]	ASCON permutation (320-bit)	3	Mask=1000, 10001 XOR	0.9862, 0.8314	MLP
[5]	KNOT-256 (256-bit)	10	Mask=1 (0, 1번째 바이트에 XOR)	0.5912	
[5]	KNOT-512 (512-bit)	12	Mask=1 (0, 1번째 바이트에 XOR)	0.6032	
[5]	CHASKEY permutation (128-bit)	4	0x00008400000004000000000000000000	0.6161	
[6]	Speck (32-bit)	18	0x0A604205 (Classical) 0x850A9520 (ML)	0.98~1.00	
[6]	Simon (32-bit)	12	0x04001900 (Classical) 0x1D014200 (ML)	0.95~0.98	
[6]	GIFT (64-bit)	8	0x000000000000000A (Classical) 0x0044000000110000 (ML)	0.99~1.00	

메모리 문제, 데이터 복잡성 등으로 인해 라운드 확장에 제한이 있었다. [6]에서는 이를 극복하기 위해 고전 구별자와 신경망 구별자를 결합한 새로운 신경망 구별자를 제안하였다. 즉, 0라운드 입력 차분을 고전 구별자에 입력하여 r라운드에 대한 출력 차분을 구하고, 해당 출력 차분을 신경망 구별자의 입력 차분으로 사용하여 $r+s$라운드에 대해 암호문과 무작위 데이터를 구별한다. 해당 연구에서는 Gohr의 신경망 구별자 구조와 Baksi et al.의 MLP 구조를 활용하였다. 제안 기법을 통해 기존의 신경망 구별자를 확장하여 18라운드 Speck, 12라운드 SIMON, 8라운드 GIFT에 대해 0.95 이상의 높은 정확도를 달성하였다.

IV.결론

본 논문에서는 딥러닝 기반의 신경망 구별자에 대한 연구 동향을 살펴보았다. 대부분의 연구가 [1]에서 파생되었으며, 다양한 구조의 신경망 구별자와 입력 차분들에 대해 연구되고 있다. 그러나 현재는 모두 라운드 축소된 암호에 대한 결과이며, 현재의 신경망 구별자가 갖는 메모리 문제와 라운드 축소된 암호에 대한 분석만이 가능하다는 한계점을 극복하기 위한 다양한 방법이 연구되어야 할 것으로 보인다.

V. Acknowledgment

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임(No.2018-0-00264, IoT 융합형 블록체인 플랫폼 보안 원천 기술 연구, 100%).

[참고문헌]

[1] Gohr, Aron. "Improving attacks on round-reduced speck32/64 using deep learning." Annual International Cryptology Conference. Springer, Cham, 2019.

[2] Hou, Zezhou, Jiongjiong Ren, and Shaozhen Chen. "Cryptanalysis of round-reduced SIMON32 based on deep learning." Cryptology ePrint Archive (2021).

[3] Chen, Yi, and Hongbo Yu. "A New Neural Distinguisher Model Considering Derived Features from Multiple Ciphertext Pairs." IACR Cryptol. ePrint Arch. 2021 (2021): 310.

[4] Benamira, Adrien, et al. "A deeper look at machine learning-based cryptanalysis." Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Cham, 2021.

[5] Baksi, Anubhab. "Machine learning-assisted differential distinguishers for lightweight ciphers." Classical and Physical Security of Symmetric Key Cryptographic Algorithms. Springer, Singapore, 2022. 141-162.

[6] Yadav, Tarun, and Manoj Kumar. "Differential-ml distinguisher: Machine learning based generic extension for differential cryptanalysis." International Conference on Cryptology and Information Security in Latin America. Springer, Cham, 2021.