NIST 표준 형태 보존 암호에 대한 딥러닝 기반의 신경망 구별자

김덕영*, 김현지*, 장경배*, 윤세영*, 서화정*†
*한성대학교 대학원 융합보안학과

서 론

● 차분 분석이란 암호의 분석기법 중 하나로, 입력 차분에 따라 출력 차분을 분석하여 키를 유추할 수 있다면, 암호 알고리즘이 안전하지 않다고 볼 수 있다.

● 차분 특성을 활용하여 인공 신경망 기반의 구별자를 사용하면 랜덤 데이터로 암호 데이터를 구별 할 수 있으며, 차분 공격 시 데이터 복잡도가 줄어든다는 이점이 있다.

● 신경망 구별자 관련 연구는 현재도 활발하게 진행되고 있으며, 본 논문에서는 형태보존암호인 FF1,FF3-1 32/128에 대해 차분 특성을 고려하여 최초의 신경망 구별자를 제안하였다.

모델의 데이터 셋, 모델 구성

● 임의의 랜덤 평문 P0, P1을 생성 후 입력차분을 만족하는 평문 쌍을 생성하기 위해 P0에 입력차분을 XOR하여 평문 P2를 구한다.

● 각 평문 P0, P1, P2를 암호화하여 암호문 C0, C1, C2를 구하고 이때, C0와 C1은 차분 관계가 아닌 랜덤 평문을 암호화한 결과이므로 두 값을 연접한 결과를 0으로 라벨링 한다.

● 이때 C0와 C2는 입력차분을 만족하는 평문의 암호문으로 특정 확률로 출력 차분을 만족하는 암호 데이터이므로 연접한 값을 1로 라벨링 한다.

● 암호화 과정에서 사용되는 평문 및 암호문은 숫자 (0 ~ 9) 또는 소문자 (a ~ z) 도메인에서 선택되고, 실제 데이터셋에는 C0, C1, C2의 비트 값이 저장된다.

● 모델의 구성은 랜덤 또는 차분 암호문 쌍의 각 비트는 입력 레이어의 각 뉴런에 할당되어 그후 히든 레이어를 거치고 출력 레이어에서 sigmoid 활성화 함수를 거쳐 0 ~ 1 사이의 값을 도출하여 해당 값과 실제 정답의 손실을 계산한다.

실험 결과

< FF1, FF3-1의 하이퍼파라미터>

Format-Preserving Encryption	FF1	FF3
Epoch	20	15
Hidden layers	5 hidden layers with 64 units	4 hidden layers with 128 units
parameters	173,956	74,497
Batch size	32	
Activation	ReLu (Hidden). Sigmoid (Output)	
Optimizer (Learning rate)	Adam(lr = 0.0001~0.001)	
Loss function	binary_crossentropy	

< FF1의 정확도 표>

	Number (10-round)				Lowercase (2-round)			
	Tr	*Val*	*Ts*	*Rel*	*Tr*	*Val*	*Ts*	*Rel*
01	0.73	0.74	0.73	0.23	0.50	0.50	0.50	0.00
02	0.74	0.75	0.74	0.24	0.51	0.51	0.51	0.01
03	0.71	0.71	0.71	0.21	0.52	0.51	0.52	0.02
04	0.75	0.75	0.75	0.25	0.51	0.51	0.51	0.01
05	0.75	0.75	0.75	0.25	0.51	0.51	0.51	0.01
06	0.75	0.75	0.75	0.25	0.51	0.51	0.51	0.01
07	0.75	0.75	0.75	0.25	0.51	0.51	0.51	0.01
08	0.80	0.80	0.80	0.30	0.51	0.50	0.51	0.01
09	0.84	0.83	0.84	0.34	0.52	0.52	0.52	0.02
0A	0.84	0.84	0.84	0.34	0.50	0.50	0.50	0.00
0B	0.82	0.82	0.82	0.32	0.51	0.51	0.51	0.01
0C	0.85	0.84	0.85	0.35	0.5	0.5	0.5	0.00
0D	0.78	0.78	0.78	0.28	0.51	0.51	0.51	0.01
0E	0.81	0.81	0.81	0.31	0.52	0.52	0.52	0.02
0F	0.85	0.85	0.85	0.35	0.52	0.52	0.52	0.02

< FF3-1의 정확도 표>

	Number (8-round)				Lowercase (2-round)			
	Tr	*Val*	*Ts*	*Rel*	*Tr*	*Val*	*Ts*	*Rel*
01	0.62	0.62	0.62	0.12	0.54	0.54	0.54	0.04
02	0.82	0.82	0.82	0.32	0.55	0.54	0.54	0.04
03	0.78	0.76	0.77	0.27	0.52	0.51	0.51	0.01
04	0.76	0.75	0.75	0.25	0.52	0.52	0.51	0.01
05	0.77	0.75	0.74	0.24	0.53	0.53	0.53	0.03
06	0.75	0.74	0.75	0.25	0.52	0.51	0.52	0.02
07	0.75	0.73	0.74	0.24	0.53	0.52	0.52	0.02
08	0.98	0.97	0.97	0.47	0.55	0.55	0.55	0.05
09	0.96	0.94	0.94	0.44	0.54	0.54	0.54	0.04
0A	0.96	0.95	0.95	0.45	0.53	0.53	0.53	0.03
0B	0.97	0.96	0.96	0.46	0.53	0.52	0.52	0.02
0C	0.97	0.95	0.95	0.45	0.53	0.53	0.53	0.03
0D	0.96	0.96	0.96	0.46	0.53	0.52	0.51	0.01
0E	0.96	0.96	0.96	0.46	0.54	0.54	0.55	0.05
0F	0.96	0.93	0.94	0.44	0.52	0.52	0.52	0.02

● 본 논문에서 FF1, FF3-1에 대한 최초의 딥러닝 신경망 구별자를 제안하였다. 우리는 단일 차분 모델에 대한 실험을 진행하였으며 모델에서 FF1일 때 10라운드와 0F 차분을 사용하였을 때 숫자 및 소문자 도메인 모두에서 가장 높은 정확도를 보였으며, FF3-1에서는 8라운드와 08차분에서 숫자 및 소문자 도메인 모두에서 가장 높은 정확도를 보였다. 본 실험을 통해 FF1에서는 0F 차분일때 정확도 및 신뢰도가 높고, FF1에서는 08차분을 사용하였을 때 정확도 및 신뢰도가 높다는 것을 확인하였다. 이처럼 각 암호마다 그에 맞는 좋은 차분이 있다는 것을 확인하였다.