# 형태보존암호 FF1에 대한 딥러닝 기반 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 32/128에 대해 차분 특성을 고려하여 최초의 신경망 구별자를 제안하였다.

## 단일 차분 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.

## 다중 차분 모델의 데이터 셋, 모델 구성

● 다중 차분 신경망 구별자의 데이터셋을 생성하는 과정으로 먼저, 평문 P0를 생성한다. 입력 차분들을 만족하는 평문 쌍들을 만들어야 하므로 P0에 입력차분들을 XOR하여 평문 Pn을 구한다.

● 각 평문 Pn을 암호화하여 암호문 Cn을 구한다. 여기서 Cn들은 각기 다른 입력 차분을 갖는 랜덤 평문을 암호화한 암호문이다.

● 따라서 C0||Cn은 n-1 클래스로 라벨링한다. 다중 입력 차분을 사용하는 구별자에서도 암호화 과정에서는 숫자 도메인과 소문자 도메인을 사용하였으며 입력 차분으로 0x0||K는 0에서 F사이의 16진수를 사용하였다.

● 다중 차분 모델은 차분 특성을 만족하는 암호문 쌍을 입력 받아 사용된 입력 차분으로 분류한다.

● 모델 구성은 데이터 셋을 구성하는 암호문 쌍의 각 비트는 입력 레이어의 각 뉴런에 할당되고, 히든 레이어를 통과한다. 출력 레이어에서는 softmax 활성화 함수를 적용하여 여러 개의 클래스에 대한 확률 값을 도출한다.

## 실험 결과

< 단일 차분 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 |

< 다중 차분 데이터 셋의 세부 사항>

| Dataset | Data size | Input difference pair | Valid accuracy |
|---|---|---|---|
| I1 | $2^{16.6341}$ per class | 01, 0F | > 0.500 |
| I2 | | 01, 02, 0F | > 0.333 |
| I3 | | 01~03, 0F | > 0.250 |
| I4 | | 01~04, 0F | > 0.200 |
| I5 | | 01~05, 0F | > 0.166 |
| I6 | | 01~06, 0F | > 0.142 |
| I7 | | 01~07, 0F | > 0.125 |
| I8 | | 01~08, 0F | > 0.111 |
| I9 | | 01~09, 0F | > 0.100 |
| I10 | | 01~0A, 0F | > 0.090 |
| I11 | | 01~0B, 0F | > 0.083 |
| I12 | | 01~0C, 0F | > 0.076 |
| I13 | | 01~0D, 0F | > 0.071 |
| I14 | | 01~0F | > 0.066 |

< 다중 차분 정확도 표>

| | Number (10-round) | | | | Lowercase (2-round) | | | |
|---|---|---|---|---|---|---|---|---|
| | Tr | Val | Ts | Rel | Tr | Val | Ts | Rel |
| I1 | 0.52 | 0.52 | 0.52 | 0.02 | 0.52 | 0.52 | 0.52 | 0.02 |
| I2 | 0.34 | 0.33 | 0.34 | 0.007 | 0.36 | 0.36 | 0.36 | 0.027 |
| I3 | 0.26 | 0.26 | 0.26 | 0.01 | 0.27 | 0.27 | 0.27 | 0.02 |
| I4 | 0.21 | 0.21 | 0.21 | 0.01 | 0.20 | 0.20 | 0.20 | 0.01 |
| I5 | 0.17 | 0.17 | 0.17 | 0.004 | 0.18 | 0.18 | 0.18 | 0.004 |
| I6 | 0.15 | 0.15 | 0.15 | 0.008 | 0.15 | 0.15 | 0.15 | 0.008 |
| I7 | 0.13 | 0.13 | 0.13 | 0.005 | 0.13 | 0.13 | 0.13 | 0.005 |
| I8 | 0.12 | 0.12 | 0.12 | 0.009 | 0.12 | 0.12 | 0.12 | 0.009 |
| I9 | 0.12 | 0.11 | 0.12 | 0.02 | 0.10 | 0.10 | 0.11 | 0.010 |
| I10 | 0.10 | 0.10 | 0.10 | 0.01 | 0.10 | 0.10 | 0.10 | 0.010 |
| I11 | 0.09 | 0.09 | 0.09 | 0.007 | 0.09 | 0.09 | 0.09 | 0.007 |
| I12 | 0.08 | 0.08 | 0.08 | 0.004 | 0.08 | 0.08 | 0.08 | 0.004 |
| I13 | 0.08 | 0.08 | 0.08 | 0.009 | 0.08 | 0.08 | 0.08 | 0.009 |
| I14 | 0.07 | 0.07 | 0.07 | 0.004 | 0.07 | 0.07 | 0.07 | 0.004 |

● 본 논문에서 FF1에 대한 최초의 딥러닝 신경망 구별자를 제안하였다. 우리는 단일 차분 모델과 다중 차분 모델에 대한 실험을 진행하였으며 단일 차분모델에서는 0F 차분을 사용하였을 때 숫자 및 소문자 도메인 모두에서 가장 높은 정확도를 보였다. 다중 차분 모델에서는 모든 경우의 정확도가 유효 정확도를 초과하며 I2에서 가장 높은 정확도를 달성하였다. 본 실험을 통해 단일 차분에서는 공통적으로 0F 차분의 정확도 및 신뢰도가 높고, 다중 차분에서는 사용되는 차분 쌍에 따라 구별자의 성능이 달라질 수 있음을 확인하였다.