Adversarial Neural Cryptography

Michael Guarino
Marist College Department of Computer Science
MSCS 630: Security Algorithms and Protocols

February 7, 2018

Cryptography is concerned with algorithms and protocols that protect the secrecy of source information as it is transferred from it's point of origin to it's intended destination. Encryption describes the method used to convert source data or plaintext from an intelligible form to an encoded form by which can only be made intelligible through decoding the encoded form with access to a decryption key. An encryption algorithm is said to be secure if information about plaintext cannot be extracted from cipher texts.

Neural networks have been proven to be universal function approximates therefore they can learn to represent a great variety of functions given appropriate parameters. Therefore neural networks are excellent candidates as encryptions algorithms. There has been some work in the area of using deep learning architectures to learn to encrypt source data. Furthermore, neural networks can also learn what to encrypt in order to achieve a desired secrecy property while maximizing utility. Thus, when we wish to prevent an adversary from seeing a fragment of a plaintext, or from estimating a function of the plaintext, encryption can be selective, hiding the plaintext only partly.

Generative adversarial networks (GAN) are a generative model devised by Goodfellow et al. in 2014. The GAN architecture is characterized by two differentiable functions, neural networks, that play different roles in refining the system. One differentiable function is known as the generator and the other is the discriminator. The generator learns to produce data from a learned probability distribution. The discriminator determines if the data that was produced by the generator is valid by determining if the input comes from the generator or from the actual data set. In this work we will

validate that a generative adversarial network can be used to learn symmetric encryption specifically shared-key encryption.

1 Bibliography

Abadi, Martin, et al. "Learning to Protect Communications with Adversarial Neural Cryptography". arxiv:1610.06918, October 2016.

Goodfellow, Ian, et al. "Generative Adversarial Networks". arxiv:1406.2661, June 2014.