计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.200300177

对抗网络上的可认证加密安全通信

吴少乾　李西明
华南农业大学数学与信息学院　广州 510000
(wu_shaoqian@163.com)

摘　要　生成对抗网络(Generative Adversarial Networks,GANs)自提出以来被广泛应用于各个领域。虽然在信息安全领域中对其的应用研究日益深入,但利用 GANs 实现公钥密码体制下的安全通信问题还未见公开报道。鉴于通信双方和敌手的对抗性质,文中利用 GANs 的对抗学习机制,在公钥密码体制场景下,将密钥生成器、通信双方的加解密和敌手的破译过程均作为神经网络,利用认证保密性来增强公私钥的联系,再利用对抗学习机制训练通信双方和敌手,以此实现通信双方在公开信道上的可认证加密安全通信(Authenticable Encrypted secure Communication based on Adversarial Network,AEC-AN)。实验采用了 16 bit,32 bit,64 bit 和 128 bit 长度的 4 种密钥进行训练,结果表明,Bob 的正确率在 91%～94%之间,Eve 的错误率在 43%～57%之间,该值接近 Eve 随机猜测的概率,从而证明了所提方法能够实现通信双方在敌手窃听环境下的安全通信。

关键词:对抗网络;可认证加密;公钥密码体制;安全通信
中图法分类号　TP183

Authenticable Encrypted Secure Communication Based on Adversarial Network

WU Shao-qian and LI Xi-ming
College of Mathematics and Informatics,South China Agricultural University,Guangzhou 510000,China

Abstract　Since GANs(Generative Adversarial Networks) has been put forward,it has been widely used in various fields,and its application in the field of information security is getting more and more deeply. However,using GANs to realize secure communication under public key cryptosystem has not been reported publicly. Therefore,in view of the adversarial nature of both communication sides and their adversary,this paper proposes an adversarial learning mechanism of GANs. In the public key cryptosystems scenarios,the key generator,encryption and decryption of both communication sides,and the decipher process of adversary are regarded as neural networks,then we use the certification confidentiality to strengthen public-private key linkage. Afterwards,by using the adversarial learning mechanism to train both communication sides and their adversary,we realize the authenticable encrypted secure communication (AEC-AN) between both communication sides on the open channel. In the experiment,4 keys with lengths of 16 bit,32 bit,64 bit and 128 bit have been used for training. The experiment result shows that Bob's accuracy rate is between 91%～94%,and Eve's error rate is between 43%～57%,which is close to the probability of Eve's random guess,thus proving that the proposed mechanism of GANs achieves the secure communication between both communication sides under the environment of adversary eavesdropping.

Keywords　Adversarial network,Authenticable encrypted,Public key cryptosystem,Secure communication

1　引言

Goodfellow 等[1]提出的生成对抗网络(GANs)为解决高维度概率密度分布中的采样及训练问题提供了极大的帮助。GANs 由于对抗学习机制的特性,迅速成为一个热门的研究方向,并在各个领域都得到了迅速的发展,在信息安全领域也不例外。近年来,GANs 在密码学上的研究也取得了一定的进展,Abadi 等[2]用神经网络代替对称加密体系中的通信双方及敌手,利用 GANs 的对抗学习机制,实现了通信双方在公开信道上的安全通信。Coutinho 等[3]结合选择明文攻击的概念和文献[2]的思想,证明了神经网络在适当的环境下可以学习一次性密码本。除了保护通信之外,GANs 也被应用于破译之中。Gomez 等[4]基于 CycleGAN[5]提出了 Cipher-GAN,用于推断给定的未成对明文和密文库的底层密码映

到稿日期:2020-03-30　返修日期:2020-06-26　本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。
基金项目:国家自然科学基金(61872152);广东省特支计划科技创新青年拔尖人才项目(2015TQ01X79);2018 年广东省农业厅省级乡村振兴战略专项项目(粤农计[2018]54 号);广东省科技计划重大专项课题(2016B010110005)
This work was supported by the National Natural Science Foundation of China(61872152),Program for Special Support of Top-Notch Young Professionals of Guangdong Province (2015TQ01X79),Provincial Rural Revitalization Strategy Special Project of Guangdong Provincial Department of Agriculture of 2018(54) and Science and Technology Major Project of Guangdong Province of China(2016B010110005).
通信作者:李西明(liximing@scau.edu.com)

射，并破译了移位密码[6]和维吉尼亚密码[7]。Hitaj 等[8]提出的 PassGAN 通过在泄露密码列表中训练 GANs 来增强密码破译性能，这为 GANs 在密码学上的应用提供了更广阔的前景。在对称密钥场景下，文献[9]利用改进的 GANs 模型实现了在模糊密钥条件下的安全通信，文献[10]则实现了具有抗泄漏[11]功能的安全通信模型。

上述均是在对称密码体制下的研究，而对于对抗网络在公钥密码体制中的研究至今还未见公开报道。相比对称密码体制[12]，公钥密码体制[13]除了要保证敌手无法破译通信内容外，还需要保证公钥能够起到安全的加密作用且私钥能够起到准确的解密作用。在利用对抗网络直接实现公钥密码体制下的安全通信时，模型无法准确“告诉”神经网络公钥和私钥的作用，因此无法抵御敌手的窃听。本文基于 GANs 的对抗学习机制，利用 GANs 能生成更好样本的优势，同时融合签名加密[14]，在公钥密码体制场景下，实现了对抗网络上的可认证加密安全通信。本文贡献主要有以下两点：

(1)利用密码学中的认证保密性来间接“告知”系统公私钥的作用，即只有拥有私钥的人才能够正确解密用其公钥进行加密的密文，以此增强公私钥之间的联系，使得公私钥能够在神经网络中发挥正确加解密的作用。

(2)以神经网络替代密钥生成器、签名算法、加密算法和解密算法，使用密钥生成网络生成有相关联系的公私钥，利用对抗学习机制进行训练，最终在公钥密码体制场景下实现了通信双方在公开信道上的可认证加密安全通信。

2 背景

2.1 生成对抗网络

生成对抗网络源自双人博弈论，是一种深度生成模型，由生成模型 G 和判别模型 D 组成，其中，G 通过学习真实样本的潜在分布来生成以假乱真的生成样本，而 D 则努力正确分辨训练样本的真实来源。通过不断的对抗训练，G 最大化生成样本与真实样本的相似度，D 最小化判别错误的概率。因此，D 和 G 的训练可以表示为关于值函数 $V(G,D)$ 的极小化极大的双方博弈问题，即：

$$\min_G \max_D V(G,D)=\mathbb{E}_{x\sim p_{data}(x)}[\log D(x)]+\mathbb{E}_{z\sim p_z(z)}[\log(1-D(G(z)))]$$

在对抗训练过程中，G 和 D 交替迭代，先固定 G，训练 D，更新 D 的最优参数，反之亦然，最终达到模型稳定，即当且仅当噪声分布 $p_z=p_{data}$ 时，达到纳什均衡，此时 G 学会了真实样本 p_{data} 的潜在分布，使得 D 只能对训练样本在真或假之间进行随机猜测，即 D 的准确率稳定在 0.5。

2.2 对抗网络在对称密码体制上的应用

如图 1 所示，Abadi 等[2]将对称密码体制中的通信双方的加、解密算法和敌手的破译算法均作为神经网络，在对称密码体制场景下，通信双方 Alice 和 Bob 进行通信，Alice 利用密钥 K 将明文 P 加密为密文 C 并在公开信道上进行传输，之后密文 C 被 Bob 和 Eve 所获取，Bob 利用相同的密钥 K 对 C 进行解密得到 P_{Bob}，而 Eve 只能依靠神经网络对 C 进行破译得到 P_{Eve}。Alice 和 Bob 组成的加解密模型与 Eve 的破译模型形成对抗关系，若利用 GANs 的对抗学习机制进行训练，使得 $d(P,P_{Bob})=0$ 且 $d(P,P_{Eve})=0.5$，则表示通信双方能够在公开信道上进行安全通信。

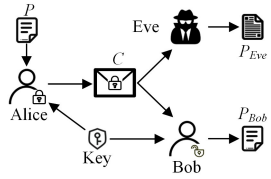

图 1 对称密码体制的保密框架

Fig. 1 Security framework of symmetric cryptography

除此之外，Comez 等[4]以 CycleGAN[5]的架构为基础进行改进，提出了 CipherGAN 模型，用于推断给定的非配对密文和明文库的密码映射关系，通过移位加密和 Vigenere 加密来说明 CipherGAN 的密码破译能力，并展示了生成对抗网络和离散数据的兼容性。Hitaj 等[8]提出一种基于 GANs 的密码猜测技术——PassGAN，其与当前基于规则的密码猜测工具相反，不依赖任何附加信息，通过 GANs 来学习密码泄漏数据中的真实密码分布信息，然后在不需要用户干预的情况下利用学习到的分布信息生成高质量的密码猜测。

3 传统可认证加密安全通信模型

我们将 Alice 作为发送方，Bob 作为接收方，Eve 作为窃听者。当 Alice 和 Bob 想要进行安全通信时，Bob 首先通过密钥源生成具有联系关系的公钥 PK_B(用于加密)和私钥 SK_B(用于解密)，然后 Bob 公开公钥 PK_B，Alice 则利用 PK_B 将想要保密的消息进行加密，得到密文 C，接着在公开信道上进行传递，同时被 Bob 和 Eve 所获取，最后 Bob 利用自己的私钥 SK_B 进行解密，而 Eve 只能依靠自己的手段进行破译，这就是公钥密码体制。

Alice 首先通过密钥源生成公钥 PK_A(用于解密)和私钥 SK_A(用于加密)，然后 Alice 公开公钥 PK_A，若 Alice 直接用 SK_A 将消息 M 加密为密文 C 并进行传输，此时，由于公钥 PK_A 是公开的，则任何人都可以对 C 进行正确解密。因为只有拥有私钥 SK_A 的 Alice 才能对 M 进行加密，所以 C 可以被看作 Alice 对 M 的签名。虽然上述过程不能保证消息的保密性，但却能够保证消息的来源和消息的完整性。

为了同时保证消息的认证性和保密性，可以利用可认证加密安全通信模型进行通信，如图 2 所示。

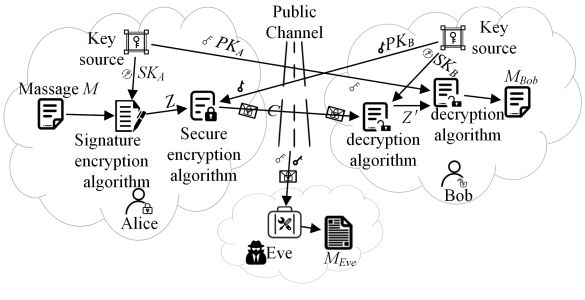

图 2 公钥密码体制的认证、保密框架

Fig. 2 Authentication and security framework of public key cryptosystem

Alice 用自己的 SK_A 对消息 M 签名加密，得到中间密文 Z，以此提供消息的认证性，再利用 Bob 的 PK_B 再次对 Z 进行保密加密得到密文 C，以此保证消息的保密性，并在公开信道

上对 C 进行传输；Bob 接收到 C 后，先用自己的 SK_B 进行一次解密，验证消息的保密性，然后利用 Alice 公开的 PK_A 进行二次解密，以验证消息的来源；而敌手 Eve 只能获取到在公开信道上传输的 C，PK_A 和 PK_B，然后利用某些手段对 C 进行破译。

4 AEC-AN 分析与设计

我们将 GANs 的对抗学习机制和公钥密码体制下的可认证加密通信相结合，将通信双方 Alice 和 Bob 的加、解密算法以及敌手 Eve 的破译过程均替换为神经网络，而非特定的加、解密算法。同时，我们设计了密钥生成网络 Gen_key1 和 Gen_key2，整体结构如图 3 所示，关于它们的网络架构，将在第 5 节进行详细阐述。Gen_key1 利用随机生成的公共参数 R_1 来生成 Alice 的密钥对 Key_A=(私钥，公钥)=(SK_A, PK_A)，Gen_key2 利用随机生成的公共参数 R_2 来生成 Bob 的密钥对 Key_B=(私钥，公钥)=(SK_B, PK_B)，其中，随机串 R_1 和 R_2 都是由 -1 和 1 组成的二元组。由于单次加、解密很难让网络准确地认识到公钥和私钥的作用，我们首先融合签名加密，采用可认证加密的双重加、解密进行实验，并将 PK_B 作为网络 Decryption 的输入，以此来增强公钥和私钥之间的联系，间接“告知”网络公钥和私钥的各自作用，然后利用对抗学习机制训练通信双方和敌手，使通信双方之间的重构误差趋于 0，敌手的破解误差趋于随机猜测概率 0.5，最后以此实现在公钥密码体制场景下的可认证加密安全通信。

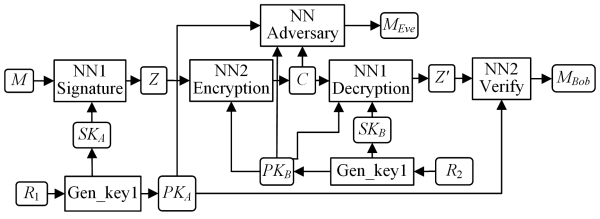

图 3 AEC-AN 网络模型

Fig. 3 AEC-AN network model

模型中，通信双方 Alice 和 Bob 希望在抵御 Eve 窃听的同时，能够在公开信道上进行正确的安全通信；相反，Eve 则希望能够准确重建 M，以得到加密的消息内容。即 Alice 和 Bob 希望最小化 M 和 M_{Bob} 之间的误差，而 Eve 希望最大化 M 和 M_{Eve} 之间的误差，这就是一种对抗性思维。根据密码学的定义，Alice 和 Bob 作为通信双方，共同承担着抵御敌手 Eve 窃听的责任，因此，我们将 Alice 和 Bob 的加、解密过程作为一个整体进行训练，以此抵抗 Eve 的破译。

将 Alice 在签名加密训练和保密加密训练时的参数分别记为 θ_A^1 和 θ_A^2，并记 $\theta_A=(\theta_A^1, \theta_A^2)$，Bob 在两次解密训练时的参数分别记为 θ_B^1 和 θ_B^2，并记 $\theta_B=(\theta_B^1, \theta_B^2)$，Eve 进行破译训练时的参数记为 θ_E。由于消息 M 是由 -1 和 1 组成的二元组，我们采用 L1 距离公式 $d(M,M')=\sum_i^N |M_i-M_i'|$ 来测量 M 分别与 M_{Bob} 和 M_{Eve} 之间的误差(其中 N 为 M 的长度)。结合密码学定义，将 M 和 SK_A 输入 Alice 的 Signature 网络中，进行签名加密训练，即：

$$Z=A(\theta_A^1, M, SK_A)$$

然后将签名加密得到的 Z 和由密钥生成器 Gen_key2 生成的 PK_B 作为 Alice 的 Encryption 网络的输入，进行保密加密训练，即：

$$C=A(\theta_A^2, Z, PK_B)$$

在公开信道上将 C 发送给 Bob，然后 Bob 的 Decryption 网络将 C，SK_B 和 PK_B 作为输入进行解密训练，即：

$$Z'=B(\theta_B^1, C, SK_B, PK_B)$$

得到中间的解密结果 Z'，将其和 PK_A 作为 Bob 的 Verify 网络的输入进行验证解密训练，得到最后的解密结果 M_{Bob}：

$$M_{Bob}=B(\theta_B^2, Z', PK_A)$$

Eve 在公开信道上获取密文 C，PK_A 和 PK_B，但只能利用 PK_A 和 PK_B 对密文 C 进行破译训练，得到破译结果 M_{Eve}，即：

$$M_{Eve}=E(\theta_E, C, PK_A, PK_B)$$

利用 L1 距离公式计算 Eve 每次进行破译训练时的结果 M_{Eve} 与原消息 M 之间的差距，以此作为其每次训练的损失函数，其定义为：

$$\begin{aligned}L_E(\theta_A,\theta_E)&=d(M,E(\theta_E,C,PK_A,PK_B))\\&=d(M,E(\theta_E,A(\theta_A^2,Z,PK_B),PK_A,PK_B))\\&=d(M,E(\theta_E,A(\theta_A^2,A(\theta_A^1,M,SK_A),PK_B),\\&\quad PK_A,PK_B))\end{aligned}$$

同理，利用 L1 距离公式计算 Bob 每次训练的重构误差，即：

$$\begin{aligned}L_B(\theta_A,\theta_B)=d(M,B(\theta_B^2,B(\theta_B^1,A(\theta_A^2,A(\theta_A^1,M,SK_A),\\PK_B),SK_B),PK_A)))\end{aligned}$$

因为 Alice 和 Bob 的目的是抵御敌手 Eve 的窃听，所以 Eve 成功破译密文所造成的损失也应该是 Alice/Bob 损失函数中的一部分，因此 Alice/Bob 的损失函数由 Bob 和 Eve 两者的重构误差共同决定，即 Bob 解密消息时所造成的错误损失以及 Eve 成功破译消息所造成的损失，故将 Alice/Bob 的损失函数定义为：

$$L_{AB}(\theta_A,\theta_B)=L_B(\theta_A,\theta_B)+(1-L_E(\theta_A,\theta_E))$$

为了增强对抗网络的收敛性，Gulrajani 等[15]通过给损失函数添加梯度惩罚项，将参数与 Lipschitz 限制[16]联系起来，以此保证模型的收敛性。结合该特性，我们设计了一个软约束，即将 M 和 M_{Bob} 结合为 M_{AB}，再次经过加、解密训练并进行误差计算，然后将其梯度惩罚添加到损失函数中，因此可以得到最终的损失函数为：

$$\begin{aligned}L_{AB}(\theta_A,\theta_B)=L_B(\theta_A,\theta_B)+(1-L_E(\theta_A,\theta_E))+\lambda*\\\mathbb{E}_{M_{AB},Key_A,PK_B}[(\|\nabla L_E\|_2-1)^2]\end{aligned}\quad(1)$$

我们的最终目标是保证 Alice 和 Bob 能够在公开信道上进行正确清晰的交流，同时能够抵御 Eve 的窃听，即最小化 Bob 的重构误差，最大化 Eve 的重构误差，因此，最终的函数目标可以转换为值函数 $O(B,E)$，即：

$$\begin{aligned}\min_{Bob}\max_{Eve}O(B,E)=L_B(\theta_A,\theta_B)+(1-L_E(\theta_A,\theta_E))+\lambda*\\\mathbb{E}_{M_{AB},Key_A,PK_B}[(\|\nabla L_E\|_2-1)^2]\end{aligned}\quad(2)$$

本文将安全性定义为，在 AEC-AN 模型中，假设敌手可以观察到公开信道上 Alice 和 Bob 的所有公共信息，当敌手和通信双方具有相同的“智能”，且具有相同的计算能力时，系统被攻破的概率是可以忽略的。

5 模型设计及训练过程

5.1 模型设计

本文将密码学中的密钥源、通信双方的加解密算法和敌

手的破译算法均作为神经网络来处理，并以 DCGANs[17] 的网络架构为基础，为通信双方和敌手设计合适的网络结构，并利用可认证加密的概念和 WGAN-GP[15] 设计损失函数，以保证在公钥密码体制下的安全通信。其中，密钥生成网络 Gen_key1 和 Gen_key2 采用相同的网络架构，将随机生成的由 -1 和 1 组成的二元组作为输入，经过两层全连接神经网络后得到公钥和私钥，其中激活函数均采用 tanh。

我们为 Alice 的两个网络（NN1 Signature 和 NN2 Encryption）设计了图 4 所示的网络架构，该架构包括 1 层全连接神经网络和 3 层卷积网络。首先，将加密内容和相应的钥匙作为全连接神经网络的输入，然后经过 3 层卷积神经网络，得到最终的密文。经过简单的调试，我们认为将 1 层全连接网络层的输出神经元个数设为 $2N$ 最为合适。4 层网络均采用 tanh 激活函数，并进行了批归一化处理[18]，且 4 层网络中的初始偏置值均取 0，其中，3 层卷积网络的通道、卷积核大小和步长分别为[3,2,2]，[2,2,1]和[1,1,1]，代表了图中的每个点都是 *mb_size* 个样本的平均误差。

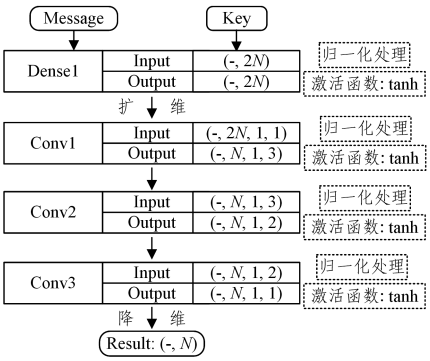

图 4 Alice 的网络架构

Fig. 4 Alice's network architecture

同理，为 Bob 的两个网络（NN1 Decryption 和 NN2 Verify）和 Eve 的破译网络（NN Adversary）设计了图 5 所示的网络架构。

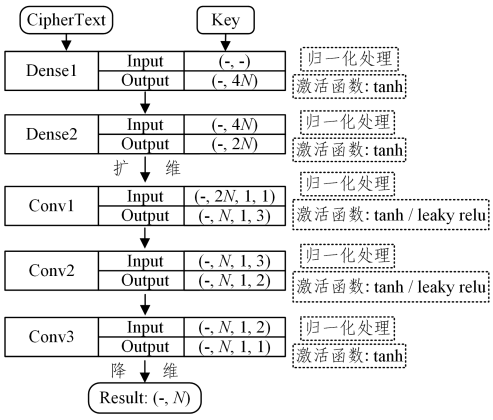

图 5 Bob 和 Eve 的网络架构

Fig. 5 Bob's and Eve's network architectures

该架构由 2 层全连接网络和 3 层卷积网络构成。2 层全连接层的神经元个数分别设为 $4N$ 和 $2N$。在激活函数设置上，Bob 的 5 层网络均采用 tanh 函数，而 Eve 的全连接层则采用了 tanh 函数，其前 2 层卷积层采用 leaky relu 作为激活函数，最后 1 层卷积网络则继续采用 tanh 函数，以保证输出的结果在$[-1,1]$之间。3 层卷积网络的通道、卷积核大小和步长分别为[3,2,2]，[2,2,1]和[1,1,1]，并将这个网络架构定义为架构一。

随着 N 的长度不断增加，上述设计的架构无法很好地实现目标，因此，在上述网络架构的基础上适当增加网络的复杂性。将 Alice 的 2 个网络的全连接层均由 1 层扩展为 2 层，即 Alice 的 2 个网络由 2 层全连接层和 3 层卷积网络层组成，除了相应的神经元数有所改变，其他模型参数不变；Bob 的 2 个网络和 Eve 的破译网络各新增 1 层卷积层，即 Bob 的 2 个网络和 Eve 的破译网络均扩展为由 2 层全连接层和 4 层卷积层组成的网络架构，其中 4 层卷积网络的通道、卷积核大小和步长分别为[4,2,2]，[3,2,2]，[2,2,1]和[1,1,1]，其他模型参数不变，并将这个网络架构定义为架构二。

5.2 训练过程

在模型训练过程中，通信方 Alice 执行加密操作，首先利用自身的私钥对原消息进行签名加密训练，然后利用通信方 Bob 的公钥对签名结果进行保密加密训练，最后将得到的结果作为最终的密文进行传输。通信方 Bob 执行解密操作，首先利用自身的公钥和密钥对密文进行解密训练，得到中间的解密结果，然后利用 Alice 的公钥对中间解密结果进行验证性解密训练，最后得到 Bob 对密文的最终解密结果。敌手 Eve 执行破解操作，利用 Alice 的公钥和 Bob 的公钥对密文进行破译训练，得到破译结果。

如图 6 所示，由 Alice 和 Bob 共同组成的加解密模型对抗由 Eve 组成的破译模型。训练 Alice 和 Bob 执行正确的加、解密操作进行通信，然后训练 Eve 对通信过程中的密文进行破解；Alice 和 Bob 发现敌手 Eve 的存在，从而“学习”更强的加解密能力再次进行通信训练，Eve 则根据 Alice 和 Bob 能力的增强而“学习”更强的破译能力，以此继续破译其通信内容。在这样不断的循环对抗学习中，Alice 和 Bob 通过不断“学习”来加强自身的加、解密能力，同时 Eve 也通过“学习”来增强自己的破译能力，直到 Alice 和 Bob 能够安全地进行正确加、解密，而 Eve 无法对其通信内容进行正确破译为止。

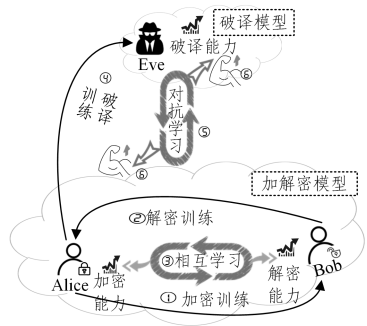

图 6 对抗训练的过程

Fig. 6 Process of confrontation training

5.3 安全性分析

本文方案的安全性主要基于以下两个方面。首先，实验均在同一台电脑上进行，这表明敌手和通信双方具有相同的计算能力。其次，在设计网络架构时，Alice 的网络层数少于

Eve的网络层数，形成敌手 Eve 的能力强于 Alice 的环境；在激活函数的设计上，将 Alice 和 Bob 的所有网络的激活函数均设置为 tanh，而 Eve 的网络的激活函数被替换为 leaky relu，以提高敌手 Eve 的收敛速度和计算速度，形成 Eve 能力强于 Ailce 和 Bob 的条件；同时设置 Alice/Bob 网络，每训练一个小批量，Eve 网络训练两个小批量，再次提高 Eve 的“智能”。本方案采用的秘钥是随机生成并保密的，而系统安全性取决于所使用的密钥的机密性，这完全符合 Kerckhoffs 准则[19]。

在本文设定的条件下，在无秘钥的前提下，当敌手 Eve 获得了公开信道上通信双方的公共信息进行密文破译时，虽然 Eve 的“智能”程度强于通信双方，但 Eve 正确破译密文的概率是可忽略的，因此保证了可认证加密通信的安全性。

6 实验与结果

实验过程中，我们设随机生成长度为 N，由 -1 和 1 组成的二元组作为需要保密的消息 M，利用 Gen_key1 和 Gen_key2 分别生成长度为 N 的公钥和私钥，使用 Tensorflow 的 Adam 优化器进行模型优化，将模型的学习速率设为 0.0008，设小批量的大小为 4096，将梯度惩罚项的系数 λ 设为 10。

实验目标是最小化 Bob 的解密结果 M_{Bob} 与 M 之间的误差，即 $d(M,M_{Bob})=0$，最大化 Eve 的破译结果 M_{Eve} 与 M 之间的误差。但 M 是由 -1 和 1 组成的二元组，如果让 M_{Eve} 与 M 完全互异，则 M_{Eve} 只需全部取反便可获得正确保密内容，因此，我们只需让 M_{Eve} 与 M 的误差比特数停留在 $N/2$，即可让 M_{Eve} 处于一种随机猜测的状态，便可以达到抵御 Eve 的目的，故令 $d(M,M_{Eve})=N/2$。为使结果有较好的可视化效果，均适用 Eve 的错误率来表达 Eve 破译效率，即将其误差转换为[0,0.5]之间的数值。例如，若当 $N=16$ bit 时，Eve 成功破译了 10 bit，则其错误率为 $1-10/16=0.375$。密码学中，密钥的长度以“位”(即比特，bit)为单位，而构成一个任意给定长度密钥的可能组合的个数可以被表示为2^n，其中 n 为密钥长度，因此，我们的实验采用 16 bit，32 bit，64 bit 和 128 bit 长度的秘钥进行实验。这 4 次实验训练的迭代次数分别为 300 000，400 000，500 000 和 600 000，因为 Alice 和 Bob 模型每训练 1 个小批量，Eve 就训练 2 个小批量，所以假设 n 为迭代次数的大小，经分析可以得到 4 次实验的时间复杂度为$O(3n)$。

实验环境为 64 位的 Windows10 系统，2.5 GHz 的 Inter i7-6500U CPU 处理器，利用 Pycharm 软件进行实现。将对本文方法与文献[20]提出的 AES 进行加密效率的比较，文献[20]中 AES 显示每字节的平均加密耗时为 0.000292511 8 s。本文实验利用神经网络来进行加密，且每字节平均加密耗时为 0.000323549 7 s。虽然本文提出的 AEC-AN 网络加密效率比 AES 的稍差，但是该研究点的发展还未完善，还具有很大的提升空间，这也是我们后续将要研究并改善的。

6.1 16 bit 模型训练

我们利用第 5 节的网络架构一进行实验，首先利用网络 Gen_key1 和 Gen_key2 生成 Alice 和 Bob 相应的公钥和私钥，其中，消息 M 的长度、Alice 和 Bob 的公钥和密钥长度均为 N，且均由 -1 和 1 组成。为了达到保护通信的目的，需要保证敌手 Eve 的破解错误率达到 -1 和 1 之间随机猜测的概率，即错误率趋近 50%，接收方 Bob 的解密正确率趋近 100%，当两者的条件同时达成时，称 Alice 和 Bob 能够在公开信道上进行安全正常通信，并抵御 Eve 的窃听。为增强公钥和私钥之间的联系，在 Bob 对密文进行二次解密时，将其自身的公钥PK_B也作为解密的钥匙之一，与私钥SK_B一同进行解密。

图 7 给出了 $N=16$ bit 时 Bob 的重构误差和 Eve 的破译误差。由图 7 可知，Bob 的解密效果和 Eve 的破译效果刚开始均为随机猜测，但随着训练步数的增加，Bob 的解密效果越来越好，Eve 的破译效果越来越差，最终在 290 000 训练步数时，Bob 解密出错的比特数为 0.98，其解密结果的准确率达到 93.875%，同时，Eve 破译出错的比特数达到 7.38，其破译结果的错误率达到了 46.125%，此时基本实现了保护通信的作用。

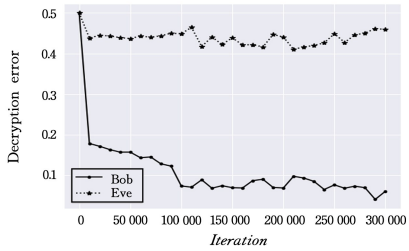

图 7 当 $N=16$ bit 时 Bob 和 Eve 的重构误差

Fig. 7 Reconstruction errors of Bob and Eve when $N=16$ bit

6.2 16 bit 以上模型训练

将 N 设置为 32，即增加消息 M、Alice 和 Bob 的公钥和密钥长度。随着 N 的长度增加，若采用与实验一相同的网络架构，经实验表明结果不太理想，因此，采用网络架构二进行实验，实验其他参数不变。

图 8 给出了当 $N=32$ bit 时，Bob 和 Eve 的重构情况。可以看出 Bob 的解密错误数在 250000 之前发生了比较明显的下降，而在后期的训练中，虽然下降的趋势变缓，但是其错误比特数仍有所下降，最终徘徊于 2.00～2.10 bit 之间，即正确率稳定在 93.75%附近，而 Eve 的破译错误比特数处于缓慢上升的趋势，最终稳定在 14.75 bit，即错误率达到了 46.09%，整体上的表现已实现了保护通信的作用。

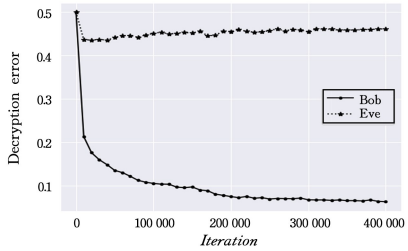

图 8 当 $N=32$ bit 时 Bob 和 Eve 的重构误差

Fig. 8 Reconstruction errors of Bob and Evewhen $N=32$ bit

同理，我们也针对 $N=64$ bit 和 $N=128$ bit 的情况进行了实验，采用网络架构二进行实验。虽然该实验结果相比 N 为 16 bit 和 32 bit 的实验结果有一定的差距，但是这两种情况下 Bob 的正确率均在 91%以上，Eve 的错误率均在 46%以上。因此，本文认为其也能较好地解决公钥密码体制场景下的安全通信问题。

对 N 为 16 bit，32 bit，64 bit 和 128 bit 这 4 种情况进行汇总，对比它们达到稳定的训练步数、Bob 正确率及 Eve 错误率的情况，具体如图 9 所示。由图 9 可知，在 4 种情形下，随着 N 的不断增大，网络需要的训练步数也随之增加，Bob 的正确率有所下降，Eve 的破译错误率基本相同。

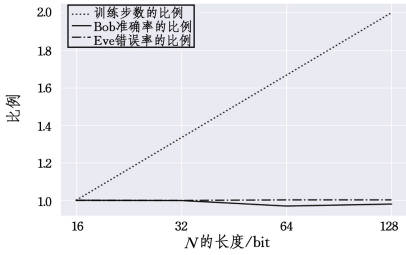

图 9 比特数对实验的影响对比

Fig. 9 Comparison of effect of bit number on experiment

结束语 本文首次将对抗网络和公钥密码体制相结合，在消息、公钥和私钥长度均为 16 bit，32 bit，64 bit 和 128 bit 这 4 种情况下分别进行实验，结果表明，与 Alice 所发送的原始消息相比，Bob 的正确率在 91%～94%之间，Eve 的错误率在 43%～57%之间，这表明通信双方能够在公开信道上进行可认证加密安全通信，同时抵御敌手的窃听。本文实验结果尚不完美，后续研究将继续改进，以使其接近理想目标，即 Bob 无误差解密，Eve 破译误差无限接近随机猜测的概率。

参 考 文 献

[1] GOODFELLOW I J，POUGET-ABADIE J，MIRA M，et al. Generative Adversarial Networks[C]// Advances in Neural Information Processing Systems 2014. MIT Press Boston，2014：2672-2680.

[2] ABADI M，ANDERSEN D G. Learning to Protect Communications with Adversarial Neural Cryptography[EB/OL]. (2016-10-21) [2020-03-20]. https://arxiv.org/abs/1610.06918.

[3] COUTINHO M，ALBUQUERQUE R D O，BORGES F，et al. Learning Perfectly Secure Cryptography to Protect Communications with Adversarial Neural Cryptography[J]. Sensors，2018，18(5)：1306.

[4] GOMEZ A N，HUANG S C，ZHANG I，et al. Unsupervised Cipher Cracking Using Discrete GANs[EB/OL]. (2018-1-15) [2020-03-20]. https://arxiv.org/abs/1801.04883.

[5] ZHU J Y，PARK T，ISOLA P，et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]// Proceedings of the 2017 IEEE International Conference on Computer Vision. IEEE New Jersey Piscataway，2017：2242-2251.

[6] CIPHER A D，BRIAN J W. Cryptanalysis of Shift Stream Generated Stream Cipher Systems-Book Review[J]. Cryptologia，1984，8(4)：360-363.

[7] SUCHITA D，FABIO D T，MARK S. Vigenère scores for malware detection[J]. Computer Virology and Hacking Techniques，2018，14(2)：157-165.

[8] HITAJ B，GASTI P，ATENIESE G，et al. PassGAN：A Deep Learning Approach for Password Guessing[C]// Applied Cryptography and Network Security-17th International Conference. Springer Berlin Heidelberg，2019：217-237.

[9] LI X M，WU J R，WU S Q，et al. Study on fuzzy key encryption based on GAN[J]. Application Research of Computers，2020，37(6)：1779-1781，1793.

[10] LI X M，WU J R，WU S Q，et al. Study on adversarial encryption based on generative adversarial networks[J]. Computer Engineering and Applications：2020，56(10)：69-74.

[11] ZHENG W C C，CHUAH C W，JANAKA A. Review on Leakage Resilient Key Exchange Security Model. [J]. IJCNIS，2019，11(1)：119.

[12] YANG X B，BOUSSAKTA S. A New Development of Symmetric Key Cryptosystem[C]// Proceedings of IEEE International Conference on Communications. IEEE New Jersey Piscataway，2008：1546-1550.

[13] DIFFIE W，HELLMAN M E. New directions in cryptography [J]. IEEE，1976，22(6)：644-654.

[14] WANG C J，XU X L，LI Y，et al. Integrating Ciphertext-Policy Attribute-Based Encryption with Identity-Based Ring Signature to Enhance Security and Privacy in Wireless Body Area Networks[C]// Information Security and Cryptology—10th International Conference. Springer Berlin Heidelberg，2014：424-442.

[15] GULRAJANI I，AHMED F，ARJOVSKY M，et al. Improved Training of Wasserstein GANs[C]// In Advances in Neural Information Processing Systems 30：Annual Conference on Neural Information Processing Systems 2017. MIT Press Boston，2017：5769-5779.

[16] ZOU D M，BALAN R，SINGH M. On Lipschitz Bounds of General Convolutional Neural Networks[J]. IEEE Transactions on Information Theory，2020，66(3)：1738-1759.

[17] RADFORF A，METZ L，CHINTALA S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks [EB/OL]. (2016-01-07) [2020-03-20]. https://arxiv.org/abs/1511.06434.

[18] YASUTAKA F，KAZUSHI I. ResNet and Batch-normalization Improve Data Separability. [C]// Proceedings of the 11th Asian Conference on Machine Learning. JMLR，2019：94-108.

[19] HENK C A V T，SUSHIL J. Encyclopedia of Cryptography and Security(2nd Edition)[M]. Berlin：Springer，2011：675.

[20] GAO J Q，LI B Y，LIAO H K，et al. Research and performance analysis of advanced encryption AES algorithm[J]. Network Security Technology & Application，2019(10)：28-30.

WU Shao-qian，born in 1994，postgraduate，is a member of China Computer Federation. His main research interests include machine learning and information security.

LI Xi-ming，born in 1974，Ph.D，associate professor，master supervisor，is a member of China Computer Federation. His main research interests include information security，intelligent image processing and machine learning.