一种改进的椭圆曲线数字签名方案

栗亚敏 张 平
(河南科技大学数学与统计学院 河南 洛阳 471023)

摘 要 数字签名是当前网络安全领域的研究热点之一。现有的椭圆曲线数字签名方案可以通过替换消息来伪造签名。针对此风险提出了一种改进方案。对改进方案进行了正确性证明和安全性证明。安全性分析可知,改进方案可以抵抗随机数攻击、不知明文密文对攻击、替换消息伪造签名攻击。将改进方案与改进前方案以及经典的 ECDSA 方案进行效率比较,结果表明:改进方案在安全性上有了很大的提高。

关键词 椭圆曲线 伪造签名 随机数攻击 随机预言模型

中图分类号 TP309 **文献标志码** A DOI:10.3969/j.issn.1000-386x.2020.12.048

AN IMPROVED ELLIPTIC CURVE DIGITAL SIGNATURE SCHEME

Li Yamin Zhang Ping
(*School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang 471023, Henan, China*)

Abstract Digital signature is one of the research hotspots in the field of network security. The existing elliptic curve digital signature scheme can forge signature by replacing message. For this security risk, an improved scheme is proposed. The correctness and security of the improved scheme were proved. Security analysis shows that the improved scheme can resist three attacks: random number attack, unknown plaintext ciphertext attack and replacement message forgery signature attack. This paper compares the efficiency of the improved scheme with that of existing scheme and the classical ECDSA scheme. The result shows that the security of the improved scheme has been greatly improved.

Keywords Elliptic curve Forged signature Random number attack Random oracle model

0 引 言

随着网络应用的蓬勃发展和广泛普及,计算机病毒、黑客、电子犯罪和电子窃听事件层出不穷,给人们的生活造成极大的隐患,因此,必须要加强网络安全意识,尽量减少安全漏洞,最大限度地降低网络安全造成的损失[1]。

数字签名是当前网络安全领域的研究热点之一,数字签名机制是保障网络信息安全的手段之一,它可以解决签名的伪造、抵赖、冒充和篡改问题[2]。数字签名在实现身份认证、数据完整性、不可抵赖性等功能方面都有着重要的应用。最初提出它的目的是在网络环境中模拟日常生活中的手工签名或印章[1]。在数字签名中,基于椭圆曲线密码系统的数字签名具有更高的安全性,椭圆曲线密码系统是离散对数密码系统在椭圆曲线上的移植[2]。椭圆曲线密码体制 ECC(Elliptic Curve Cryptography)由 Koblitz[3] 和 Miller[4] 分别独立提出,它是利用有限域上的椭圆曲线有限群代替基于离散对数问题密码体制中的有限循环群所得到的一类密码体制[5-6],它的安全性是基于椭圆曲线离散对数问题(ECDLP)的求解困难性基础之上。因此,严格地说,它不是一种新的密码体制,它只是已有密码体制的椭圆曲线型的翻版。椭圆曲线密码体制的研究历史并不太长,但由于它自身突出的优点,得到了密码学界的重视与广泛推广[7]。Johnson 等[7] 在 1992 年第一次提出椭圆曲线密码的数字签名算法(ECDSA),这一算法被国际化标准组织定义为标准数字签名算法。2007

收稿日期:2019-06-28。国家自然科学基金项目(11401172);河南科技大学研究生创新基金项目(CXJJ-2018-KJ17)。
栗亚敏,硕士生,主研领域:密码学。**张平**,副教授。

年,李复才等[8]设计了一种新的无求逆的签名算法,该算法简化了运算的复杂程度,保证了算法安全性。2008 年,张庆胜等[9]对模乘运算进行了改进,提出了一种新的椭圆曲线数字签名方案。同年,潘晓君[10]提出了一个新的基于椭圆曲线的数字签名方案,该方案不需要进行模逆操作,大大提高了签名的效率。杨青等[11]也提出了一种改进的基于椭圆曲线的数字签名方案,该方案能够有效地抵抗生日攻击,提高数字签名的安全性。2009 年,武美娜等[12]改进了椭圆曲线数字签名算法,改进算法不需要进行求逆运算,比传统算法具有更少的时间复杂度。2011 年,陈亮等[13]改进 ECDSA 签名算法,提出了一种新的椭圆曲线数字签名方案。此方案在签名和验证过程中避免了求逆运算,也减少了点乘。同年,许德武等[14]将 ElGamal 签名方案移植到椭圆曲线密码系统中,改进签名生成及验证过程,使用代数运算代替椭圆曲线上的数乘运算,得到了一种新的椭圆曲线数字签名方案。2013 年,周克元[15]设计了一种快速椭圆曲线消息恢复数字签名方案,该方案仅仅具有 2 次模乘运算,并且没有模逆运算。同年,逯玲娜等[16]提出了两个新的不需要模逆操作的基于椭圆曲线的数字签名方案。2014 年,严琳等[17]设计了一种分段快速标量乘算法,并将其运用到了 ECDSA 方案中,提高了 ECDSA 方案的效率。2015 年,陈辉焱等[18]设计了一种具有前向安全的数字签名方案,该方案有效地减少了密钥泄露带来的损失。2016 年,周克元[19]设计了一种具有消息恢复功能的椭圆曲线数字签名方案,该方案不仅能抗伪造签名攻击,还具有前向安全性。随着数字签名技术[20-23]的不断进步,近年来,许多新的椭圆曲线数字签名方案[24]被相继提出。

本文重点研究了文献[9]算法,发现该算法可被替换消息伪造签名攻击。本文分析了其原因,提出了一种新的改进方案。

1　文献[9]方案

1.1　参数选择

先选择安全的参数。$D = (q, FR, a, b, G, n, h)$,输入有限域的大小 q,随机产生 F_p 上的一条安全椭圆曲线[2] $E: y^2 = x^3 + ax + b (\mathrm{mod}p)$,在椭圆曲线上寻找一个基点 G,G 不可任意改变,取素数 $n > 2^{160}$ 且 $n > 4\sqrt{p}$,$nG = O$,$h = \dfrac{\#E(F_p)}{n} (h << n)$。$SHA_1$ 是安全的哈希函数。选择私钥 $x \in [1, n-1]$,计算公钥 $Q = dG$。若 $Y = O$,则重新选择私钥。将 $D = (q, FR, a, b, G, n, h)$,$Q$,$SHA_1$ 公开,d 保密。

1.2　签名生成

签名者 A 利用上面的参数对消息 m 进行签名:

(1) 选择随机数 $k \in [1, n-1]$;

(2) 计算 $kG = (x_1, y_1)$,$r = x_1 \mathrm{mod}n$;

(3) 计算 $e = SHA_1(m)$;

(4) 计算 $s = (er)^{-1}(k + d) \mathrm{mod}n$;

(5) 签名结果为 (r, s)。

1.3　签名验证

验证者 B 验证 (r, s) 是 A 对消息 m 的签名:

(1) 检验 $r, s \in [1, n-1]$,若不成立,返回拒绝签名;

(2) 计算 $e = SHA_1(m)$;

(3) 计算 $w = (er)s \mathrm{mod}n$,那么就有公式 $(er)s = (k + d) \mathrm{mod}n$ 成立;

(4) 计算 $wG - Q = (x_2, y_2)$;

(5) 计算 $v = x_2 \mathrm{mod}n$,若 $v = r$,则接受该签名,否则拒绝该签名。

1.4　安全性分析

在该算法中签名等式如下:

$$s = (er)^{-1}(k + d) \mathrm{mod}n \tag{1}$$

可以用另一消息 m' 替换原有消息 m 进行伪造签名。替换消息伪造签名成功的原因如下:s、e、r 已知,由:

$$s = (er)^{-1}(k + d) \mathrm{mod}n \tag{2}$$

可求得 $(k + d) \mathrm{mod}n$。然后计算 $e' = SHA_1(m')$,那么就可以计算出:

$$s' = (e'r)^{-1}(k + d) \mathrm{mod}n \tag{3}$$

从而得到伪造签名 (r, s')。

另外,高伟等[1]设计的快速签名等式如下:

$$s = k - l - rd \mathrm{mod}n \tag{4}$$

该式也存在这样的错误。

2　改进方案

2.1　参数的选择

先选择安全的参数。$D = (q, FR, a, b, G, n, h)$,输入有限域的大小 q,随机产生 F_p 上的一条安全椭圆曲线 $E: y^2 = x^3 + ax + b (\mathrm{mod}p)$,在椭圆曲线上寻找一个基点 G,G 不可任意改变,取素数 $n > 2^{160}$ 且 $n > 4\sqrt{p}$,$nG = O$,$h = \dfrac{\#E(F_p)}{n} (h << n)$。$H$ 是安全的哈希函数。

将 $D=(q,FR,a,b,G,n,h)$,H 公开。

2.2 密钥生成

(1) 选择 $x\in[1,n-1]$;

(2) 计算 $Y=xG$。若 $Y=O$,跳转至步骤(1);

(3) 公钥为 Y,私钥为 x。

2.3 签名生成

签名者 A 利用上面的参数对消息 m 进行签名:

(1) 选择随机数 $k\in[1,n-1]$;

(2) 计算 $R=kG=(x_1,y_1)$,$r=x_1\bmod n$,若 $r=0$,跳至步骤(1);

(3) 计算 $e=H(m)$;

(4) 计算 $s=ke-rxe^2\bmod n$(其中 e^2 可以通过预处理提高签名速度),如果 $s=0$,则跳至步骤(1);

(5) 签名结果为 (e,s)。

2.4 签名验证

验证者 B 验证 (e,s) 是 A 对消息 m 的签名:

(1) 检验 $r,s\in[1,n-1]$,若不成立,返回拒绝签名;

(2) 计算 $e=H(m)$;

(3) 计算 $w=xe\bmod n$;

(4) 计算 $X=w^{-1}sY+rwG=(x_2,y_2)$;

(5) 若 $X=O$,拒绝该签名;

(6) 计算 $v=x_2\bmod n$,若 $v=r$,则接受该签名;否则拒绝该签名。

2.5 方案正确性证明

若 $v=r$,则:

$$X=w^{-1}sY+rwG=(xe)^{-1}sY+rwG \tag{5}$$

由于 $s=ke-rxe^2\bmod n$,故:

$$X=(xe)^{-1}(ke-rxe^2)Y+rwG \tag{6}$$

则:

$$X=x^{-1}(k-rxe)Y+rwG \tag{7}$$

即:

$$X=x^{-1}kY-reY+rwG \tag{8}$$

由于 $Y=xG$,那么:

$$X=kG-rexG+rwG \tag{9}$$

又因为 $w=xe\bmod n$,所以:

$$X=kG \tag{10}$$

从而 $v=r$ 得证。

3 安全性证明

定理 1 在随机预言模型[25-27]中,若存在一个敌手 A 能够在 t 时间内以不可忽略的优势 ε 攻破改进方案,那么存在一个算法 B 以至少为 $\varepsilon'>\varepsilon\frac{1}{q_H}$ 的优势攻破 ECDLP 问题,其中 q_H 表示敌手 A 最多进行 q_H 次哈希查询。

证明:

该证明过程实际上是一个敌手 A 和算法 B 之间的交互式游戏。算法 B 接收一个随机的 ECDLP 问题实例 $Y=xG$,他的目标是计算出 x。算法 B 把敌手 A 作为子程序,算法 B 扮演敌手 A 的挑战者。

1) 设置阶段。游戏一开始,算法 B 将系统参数发送给敌手 A。算法 B 要维护 L_{H}、L_{S}、L_{V} 三张列表,这些表初始为空,其中:列表 L_{H} 用来跟踪敌手 A 对预言机 H 的询问;列表 L_{S} 用于模逆签名预言机;列表 L_{V} 用于模逆验证预言机。

2) 查询阶段。

(1) 哈希查询。所有之前被签名过的 (m,e) 被储存在列表 L_{H} 中,当敌手 A 对消息 m 进行哈希查询时,算法 B 首先检查消息 m 是否已经出现在列表 L_{H} 中。若已存在,算法 B 直接将 e 返回给敌手 A;否则,算法 B 从 $\{0,1\}^n$ 中任意选择 e,并将 (m,e) 存储进列表 L_{H} 中,然后将 e 返回给敌手 A。

(2) 签名查询。敌手 A 向算法 B 提交消息 m,算法 B 任意选择随机数 $k\in[1,n-1]$,然后计算 $R=kG=(x_1,y_1)$,提取 $r=x_1\bmod n$。算法 B 在列表 L_{H} 中搜索 (m,e),若列表 L_{H} 中存在 (m,e),则返回符号“$\perp$”,查询终止;否则,算法 B 计算 $s=ke-rxe^2\bmod n$;然后,将签名结果 (e,s) 返回给敌手 A。

3) 伪造阶段。若算法 B 在查询阶段没有终止退出,那么敌手 A 将以至少为 ε 的优势输出伪造消息 $m*$ 和有效伪造签名 $(e*,s*)$。由于敌手 A 执行哈希查询的次数大于 $\frac{1}{q_H}$,所以敌手 A 至少以 $\varepsilon'>\varepsilon\frac{1}{q_H}$ 的优势计算出 $k=s*(e*)^{-1}+rxe*$,从而 ECDLP 问题被解决。

证毕。

由于 ECDLP 问题是数学难题,目前没有算法可以解决该问题,因此不存在能够在 t 时间内以不可忽略的优势 ε 攻破改进方案的敌手。因此,改进方案是安全的。

3.1 不可伪造性

传统的椭圆曲线数字签名方案 ECDSA 方案并没有严格的安全性证明。2005 年 Brown[28] 基于离散对数难解性以及哈希函数具有抗碰撞性的假设,给出了 ECDSA 方案的不可伪造性证明。该证明同样适用于改进方案,此处省略了其证明。

3.2　抵抗随机数攻击

不同消息使用同一签名方案进行签名时,使用相同的随机数(这种概率非常小)是不行的[29],因为一旦随机数相同就可以用一个二阶的线性方程组解出私钥,从而造成密钥的泄露。

如果使用 ECDSA 方案对不同消息进行签名时用相同的随机数,那么就可以根据方程组:

$$\begin{cases} s_1 k = e_1 + xr \bmod n \\ s_2 k = e_2 + xr \bmod n \end{cases} \tag{11}$$

解出:

$$k = (s_2 - s_1)^{-1}(e_2 - e_1) \bmod n \tag{12}$$

进而解出私钥:

$$x = r^{-1}(s_1 k - e_1) \bmod n \tag{13}$$

实际上,有很多方案即使每次使用不同的随机数,也有可能被该攻击方法的推广所破解。比如,记 $u = xe + s \bmod n$,其中:x 是私钥;s 是签名结果中的 s;e 是被签名的消息或消息的哈希函数。那么 u 在区间[0,n - 1]的取值是随机的,攻击者只需计算 $eY + sG$。如果对消息 m_1、m_2 的签名中计算得:

$$e_1 Y + s_1 G = e_2 Y + s_2 G \bmod n \tag{14}$$

那么就可推出 $u_1 = u_2 \bmod n$。因此:

$$e_1 x + s_1 = e_2 x + s_2 \bmod n \tag{15}$$

从而就可以解出私钥:

$$x = (e_1 - e_2)^{-1}(s_2 - s_1) \bmod n \tag{16}$$

在通过改进方案使用相同随机数对不同消息进行签名时,有如下方程组:

$$\begin{cases} s_1 = ke_1 - rxe_1^2 \bmod n \\ s_2 = ke_2 - rxe_2^2 \bmod n \end{cases} \tag{17}$$

式中:签名 (e_1, s_1) 和 (e_2, s_2) 是已知的;k、x 和 r 是未知的,由于方程组中含有两个方程三个未知数,因此无法求解方程组。

另外,如果想通过上述攻击方法的推广来破解改进方案也是不可能的。因为破解时要求计算不同签名的某个随机变量是否取相同的数值,这种概率也是非常小的以至于破解的计算量非常大,比直接计算离散对数还难。假设 u 取值的概率在区间[0,n - 1]上是均匀分布的,那么由生日攻击[30]的结论得 a 次不同消息的签名中存在两次签名 $u_1 = u_2$ 的概率为 0.5 时,则 $a \approx 1.17\mathrm{sqrt}(n)$,这是一个非常大的数。如果没有直接的 eY 与 sG 值的话,计算难度是非常大的。所以,改进方案可防止针对随机数的攻击。

3.3　不可否认性

如果签名者想要否认自己曾对某个消息进行签名,那么接受者可以将签名提供给第三方。第三方可以通过验证公式来验证这个签名是否是签名者对该消息的签名。由于第三方在验证过程中不需要签名者的协助,所以这就可以防止签名者否认他的签名。

3.4　不知明文密文对的攻击[31]

(1) 攻击者想要直接通过 $Y = xG$ 解出 x 是不可实现的,因为这需要求解椭圆曲线上的离散对数的数学难题。

(2) 攻击者想要通过验证式(18)伪造 m' 的签名 (e', s') 是不可实现的,因为攻击者需要先确定一个 r'(或 s'),再去求解 s'(或 r'),这都需要求解椭圆曲线上离散对数这个数学难题。

$$kG = w^{-1}sY + rwG \tag{18}$$

3.5　抵抗替换消息伪造签名攻击

攻击者想通过加减乘除将式(19)中的 e 替换成 e' 是做不到的,因为该签名方程涉及 e^2 项。另外,在替换的同时难以保证 $r' = x \bmod n$(其中 $k'G = (x, y)$)。

$$s = ke - rxe^2 \bmod n \tag{19}$$

由改进方案的安全性分析可知,改进方案的安全性应该不低于 ECDSA 方案的安全性。与文献[9]方案相比,改进方案涉及 e 和 e^2,大大增加了其对替换消息攻击的抵抗。

4　效率分析

从算法运算量角度分析,设模乘运算的数据规模是 n,1 次倍点运算复杂度是 $O(n^2)$,1 次模逆运算复杂度是 $O(9n^2)$(相当于 9 次倍点运算),1 次模乘运算复杂度是 $O(n^2 \log_2 n)$[32]。将本文方案的运算量与文献[9]方案、ECDSA 方案的运算量进行对比,结果如表 1 所示。

表 1　方案效率比较

方案	签名阶段			验证阶段			签名长度
	模乘	模逆	倍点	模乘	模逆	倍点	
ECDSA	2	1	1	2	1	2	2|n|
文献[9]方案	2	1	1	2	0	1	2|n|
本文方案	3	0	1	3	1	2	2|n|

由表 1 可知,ECDSA 的总运算量为:

$$N_1 = O[(4\log_2 n + 22)n^2] \tag{20}$$

文献[9]方案的总运算量为:

$$N_2 = O[(4\log_2 n + 12)n^2] \tag{21}$$

本文方案的总运算量为:

$$N_3 = O[(6\log_2 n + 13)n^2] \tag{22}$$

本文方案在签名验证上虽然无法与文献[9]方案

比较,但是其签名验证效率不低于 ECDSA 方案。影响复杂度的主要运算是模乘运算和模逆运算。本文方案在密钥生成时与另外两种方案的复杂度相同。在签名阶段,改进方案比另外两种方案多了 1 次模乘运算,少了 1 次模逆运算;在签名验证阶段,本文方案虽然比文献[9]方案多了 1 次模乘运算、1 次模逆运算和 1 次倍点运算,但是本文方案不仅能防止消息伪造签名攻击,还能防止随机数攻击。总体来说,本文方案加强了安全性,适当牺牲了速度。

5 结 语

本文首先对文献[9]方案进行重点研究和分析,发现该方案存在替换消息伪造签名的安全隐患。针对此安全隐患,本文提出了一个新的改进方案,并对其进行安全性证明。本文方案虽然在效率上有待提高,但是其不仅能防止消息伪造签名攻击,还能防止针对随机数的攻击以及不知明文密文对的攻击。总的来说,改进后方案在安全性上有了很大的提高。因此,本文方案适用于对效率要求较低、对安全性要求高的应用场合。

参 考 文 献

[1] 高伟,张国印,王欣萍. 一种改进的椭圆曲线数字签名算法[J]. 黑龙江大学自然科学学报,2010,27(3):396 - 402.

[2] 任中岗,翟东海. 有限域 GF(q)上安全椭圆曲线的选取[J]. 信息与电子工程,2009,7(5):493 - 496.

[3] Koblitz N. Elliptic curve cryptosystems[J]. Mathematics of Computation,1987,48(177):203 - 209.

[4] Miller V S. Uses of elliptic curves in cryptography[C]//Proceedings of Conference on the Theory and Application of Cryptographic Techniques,1985. Springer,1986:417 - 426.

[5] 周克元. 对一个椭圆曲线数字签名方案的攻击分析与改进[J]. 计算机应用与软件,2013,30(10):331 - 333.

[6] 周克元. 基于椭圆曲线和因子分解双难题的数字签名方案[J]. 计算机科学,2014,41(S1):366 - 368.

[7] Johnson D,Menezes A,Vanstone S. The elliptic curve digital signature algorithm (ECDSA)[J]. International Journal of Information Security,2001,1(1):36 - 63.

[8] 李复才,张永平,孙宁. 椭圆曲线数字签名方案的研究与改进[J]. 计算机工程与设计,2007,28(21):5241 - 5242,5263.

[9] 张庆胜,郭宝安,徐树民,等. 快速椭圆曲线签名验证算法[J]. 计算机工程与设计,2008,29(17):4425 - 4427.

[10] 潘晓君. 一种新的基于椭圆曲线的数字签名方案[J]. 计算机系统应用,2008(1):35 - 37.

[11] 杨青,辛小龙,戢伟. 基于椭圆曲线的数字签名和代理数字签名[J]. 计算机工程,2008,34(23):147 - 149.

[12] 武美娜,刘瑞芹,张凤元. 基于无线局域网的椭圆曲线数字签名改进算法[J]. 通信技术,2009,42(4):108 - 110.

[13] 陈亮,游林. 椭圆曲线数字签名算法优化与设计[J]. 电子器件,2011,34(1):89 - 93.

[14] 许德武,陈伟. 基于椭圆曲线的数字签名和加密算法[J]. 计算机工程,2011,37(4):168 - 169.

[15] 周克元. 快速椭圆曲线消息恢复数字签名方案[J]. 西北师范大学学报(自然科学版),2013,49(5):54 - 56.

[16] 逯玲娜,程磊. 两种基于椭圆曲线的数字签名及其性能分析[J]. 重庆科技学院学报(自然科学版),2013,15(5):35 - 37.

[17] 严琳,卢忱. 基于快速标量乘算法的椭圆曲线数字签名方案[J]. 电子科技,2014,27(4):23 - 26.

[18] 陈辉焱,袁勇,万宗杰,等. 一种基于椭圆曲线的前向安全数字签名[J]. 电信科学,2015,31(10):99 - 102.

[19] 周克元. 一种椭圆曲线消息恢复数字签名方案的分析与改进[J]. 西北师范大学学报(自然科学版),2016,52(4):38 - 40.

[20] 李明祥,安妮. 基于格的前向安全签名方案[J]. 密码学报,2016,3(3):249 - 257.

[21] 姚丽莎,张军委,席何文,等. 指纹 IRLRD 特征数字签名技术[J]. 计算机应用与软件,2017,34(6):322 - 327,333.

[22] 魏文燕,彭维平,李子臣,等. 一种基于 Rabin 和 Paillier 的数字签名方案[J]. 计算机应用与软件,2017,34(12):301 - 306.

[23] 周克元. 基于双难题的数字签密方案研究[J]. 计算机应用与软件,2017,34(10):316 - 319,329.

[24] 张平,栗亚敏. 前向安全的椭圆曲线数字签名方案[J]. 计算机工程与应用,2020,56(1):115 - 120.

[25] Jiang H D,Zhang Z F,Ma Z. Key encapsulation mechanism with explicit rejection in the quantum random oracle model[C]//22nd IACR International Conference on Practice and Theory of Public-Key. Springer,2019:618 - 645.

[26] Gao W,Hu Y P,Wang B C,et al. Efficient ring signature scheme without random oracle from lattices[J]. Chinese Journal of Electronics,2019,28(2):266 - 272.

[27] Kiltz E,Lyubashevsky V,Schaffner C. A concrete treatment of fiat-shamir signatures in the quantum random-oracle model[C]//37th Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer,2018:552 - 586.

[28] Brown D R L. Generic groups,collision resistance,and ECDSA[J]. Designs,Codes and Cryptography,2005,35:119 - 152.

(下转第 327 页)

5 结　语

为检测开源软件中存在的大量漏洞,本文提出了一种基于双向 LSTM 的 Java 漏洞检测方法。首先运用静态分析提取源代码语义特征并生成中间表示,然后在将生成的中间表示映射为向量的同时为其贴上“是否安全”的标签,最后运用神经网络在数据集上训练生成漏洞检测模型。静态分析主要包括抽象语法树、数据流和控制流分析,目的是提取方法完整的数据和控制依赖,以提高漏洞检测的准确性。神经网络选择双向 LSTM,目的是可以更好地学习源代码的序列信息。

实验结果表明,模型在测试集上取得了 93.8% 的准确率和 90.1% 的召回率,均优于现有的基于机器学习的漏洞检测方法,验证了本文方法的优越性。后续工作主要分为两个方面,一是进一步提升模型的精确度,减少人工参与,实现开源软件漏洞的自动化检测;二是尝试将本文方法扩展到除 Java 外的其他语言,验证其适用性。

参 考 文 献

[1] Snyk. The state of open source security report[R/OL]. [2019-07-02]. https://bit.ly/SoOSS2019.

[2] Shankar U, Talwar K, Foster J S, et al. Detecting format string vulnerabilities with type qualifiers[C]//Proceedings of the 10th conference on USENIX Security Symposium. ACM, 2001, 10: 16.

[3] Yamaguchi F, Golde N, Arp D, et al. Modeling and discovering vulnerabilities with code property graphs[C]//2014 IEEE Symposium on Security and Privacy. IEEE, 2014: 590 - 604.

[4] Qian C X, Luo X P, Le Y, et al. VulHunter: Toward discovering vulnerabilities in android applications[J]. IEEE Micro, 2015, 35(1): 44 - 53.

[5] Eschweiler S, Yakdan K, Gerhards-Padilla E. DiscovRE: Efficient cross-architecture identification of bugs in binary code[C]//The Network and Distributed System Security Symposium(NDSS), 2016.

[6] Li Y K, Chen B H, Chandramohan M, et al. Steelix: Program-state based binary fuzzing[C]//2017 11th Joint Meeting on Foundations of Software Engineering. ACM, 2017: 627 - 637.

[7] Wang J J, Chen B H, Wei L, et al. Skyfire: Data-driven seed generation for fuzzing[C]//2017 IEEE Symposium on Security and Privacy (SP). IEEE, 2017: 579 - 594.

[8] Blanda A. Fuzzing Android: a recipe for uncovering vulnerabilities inside system components in Android[C]//Black Hat Europe, 2015.

[9] Shin Y, Williams L. Can traditional fault prediction models be used for vulnerability prediction? [J]. Empirical Software Engineering, 2013, 18(1): 25 - 59.

[10] Pan X R, Wang X Q, Duan Y, et al. Dark hazard: Learning-based, large-scale discovery of hidden sensitive operations in android apps[C]//Network and Distributed System Security Symposium, 2017.

[11] Chowdhury I, Zulkernine M. Using complexity, coupling, and cohesion metrics as early indicators of vulnerabilities[J]. Journal of Systems Architecture, 2011, 57(3): 294 - 313.

[12] The MITRE Corporation. Common weakness enumeration? [DB/OL]. [2019-07-02]. https://cwe.mitre.org.

[13] Parr T, Harwell S, Vergnaud E, et al. ANTLR4[CP/OL]. [2019-07-02]. https://github.com/antlr/antlr4.

[14] Arzt S, olhotak, mbenz89, et al. Soot[CP/OL]. [2019-07-02]. https://github.com/Sable/soot.

[15] National Institute of Standards and Technology. NIST software assurance reference dataset[DS/OL]. [2019-07-02]. https://samate.nist.gov/SARD.

[16] Li Z, Zou D Q, Xu S H, et al. VulDeePecker: A deep learning-based system for vulnerability detection[C]//The Network and Distributed System Security Symposium, 2018.

[17] 李元诚,黄戎,来风刚,等. 基于深度聚类的开源软件漏洞检测方法[J]. 计算机应用研究, 2020, 37(4): 1107 - 1110, 1114.

(上接第 308 页)

[29] 李晓峰,赵海,王家亮,等. 基于增加一个随机数的 ElGamal 数字签名算法的改进[J]. 东北大学学报(自然科学版), 2010, 31(8): 1102 - 1104, 1112.

[30] 喻秋叶. 生日悖论在密码学中的应用[D]. 武汉:华中师范大学, 2013.

[31] 张先红. 数字签名原理及技术[M]. 北京:机械工业出版社, 2004: 95 - 95.

[32] 韩益亮,杨晓元,户军茹,等. 改进的 ECDSA 签名算法[C]//第二十届全国数据库学术会议论文集(研究报告篇), 2003.

(上接第 315 页)

[18] Lin Q, Yang L, Guo Y. Proactive batch authentication: Fishing counterfeit RFID tags in muddy waters[J]. IEEE Internet of Things Journal, 2019, 6(1): 568 - 579.

[19] Rashid N, Choudhury S, Salomaa K. Localized algorithms for redundant readers elimination in RFID networks[J]. International Journal of Parallel, Emergent and Distributed Systems, 2019, 34(3): 260 - 271.

[20] Corchia L, Monti G, Tarricone L. A frequency signature RFID chipless tag for wearable applications[J]. Sensors, 2019, 19(3): 494.