口令安全

摘要

身份认证是保障信息系统安全的第一道防线，口令具有简单易用、成本低廉、容易更改等特性，在现阶段以及未来依然是主要的身份认证方法。在大多数系统中，口令认证是唯一的一道防线。面对口令泄露事件的频繁发生，攻击者的方法不断增强，研究口令猜测方法对口令安全有重要意义。本文总结了用户的脆弱口令行为，介绍口令安全的研究背景及研究意义，口令分布及不同的口令猜测的方法，最后对该研究方向进行总结和并对未来进行展望。

Identity authentication is the first line of defense to ensure the security of information systems, and passwords are easy to use, low cost, easy to change, etc., and will still be the main identity authentication method at this stage and in the future. In most systems, password authentication is the only line of defense. In the face of frequent password leakage events, the attacker's method is constantly increasing, and it is of great significance to study the password guessing method for password security. This paper summarizes the user's vulnerable password behavior, introduces the research background and significance of password security, password distribution and different password guessing methods, and finally summarizes the research direction and looks forward to the future.

背景及研究意义

上世纪60年代，科尔巴托参与一个名为兼容分时系统（CTSS， Compatible Time-Sharing System）的项目[[[1]](#endnote-0)]，该系统可让多位用户在不同地点，通过电话线同时访问一台计算机。当时距离第一台计算机问世才10年多，人们对计算机使用非常的简单，并且当时计算系统非常缓慢。当时的CTSS系统为每个用户提供了一组私有文件，但由于没有密码，用户可以随便阅读其他人的文件。为解决问题，科尔巴托提出为每个用户设置一个密码的解决方案。这也被认为是计算机密码的诞生。

新的认证方式如图形口令，生物认证等已经得到广泛应用，但几乎替代方案在可部署性上劣于口令，以及其自身如成本高昂，隐私泄露等固有缺陷，口令现阶段以及可预见的未来都将广泛应用。如果口令出现问题，如弱口令，口令被破解等，将使得大量的信息泄露，生命财产安全将得不到保证。

近年来，数以千计的网站口令文件库遭到泄露。已经确认口令文件库泄露的网站，2022年有1999[[[2]](#endnote-1)]起，2021年有2186起[[[3]](#endnote-2)]。随着互联网的发展，用户账户数目的不断增长，网站口令管理策略的杂乱无章，用户处理安全事务的能力和精力有限，导致了解决口令安全问题变得越来越复杂。

口令分布

用户的脆弱口令行为

用户口令的偏好性

用户口令行为偏向选择流行口令，例如123456，loveyou等等。受网站口令策略以及所处文化环境的影响，用户倾向于选择更为简单，便于记忆的组合，并且具有群体特征。在12306、CSDN、126、Dodonew4个中文口令数据集上，攻击者尝试最流行的10个口令猜测成功的概率最高可达10.44%，选择最流行的10个PIN码的用户的比例达到14.77%-25.62%[[[4]](#endnote-3)]。

口令重用

面对着越来越多需要管理的账号，又由于人类大脑记忆是有限的，大多数用户都会进行口令重用。直接完全使用一个其他账户的密码，或者对其他一个账户的密码稍加修改，使用其他密码中含有的组成部分的人超过75%。平均34%的用户直接重用口令，31%的用户重用口令时仅仅进行简单的修改。

基于个人信息构造口令

为便于记忆，用户在构造口令时往往会掺入个人信息。使用各类个人可标识信息构造口令，如姓名，生日等。这也就意味着，攻击者可以结合用户的信息和口令构造模型来猜测用户口令。

口令分布理论模型Zipf

2016年，Wang[[[5]](#endnote-4)]等人证实口令服从Zipf分布。Zip分布的发现经历了先否认再确认的过程。Malone 和 Maher[[[6]](#endnote-5)]曾在2012年试图调查密码的分布，分析出Rockyou数据集实际上不太可能是Zipf分布的。Malone和Maher试图用Zipf对口令分发进行建模，并将他们的模型与统一模型进行比较，得出Zipf分布于用户选择口令的频率是相对较好的匹配。

Wang等人的发现与之相反，即用户生成的口令是Zipf分布的，并基于密码的Zipf分布提出了一种新的基于统计的密码数据集强度度量标准。Wang等人找出与之观察结果不同的原因是是否将数据集的所有口令都拟合到Zipf模型中和数据集规模。根据大数定律，那些低频次口令会极大的负面影响拟合结果。而百万级的数据集对于此问题规模太小。

密码群体服从Zipf分布，在实验的14个真实大规模数据集中得到验证。当样本足够大时，高频次事件总是占每个样本的95%以上，并且遵从Zipf定律。

2022年，Wang等人[[[7]](#endnote-6)]探索了8种不同的模型试图减小经验分布和拟合分布不可忽略的偏差，发现了CDF-Zipf的一种变体，拉伸指数模型可以很好的降低偏差，并引入了一种新的基于对数似然的拟合优度测量方法。

口令猜测

安全的定义取决于具体的攻击模型。用户脆弱的口令行为，网站脆弱的口令生成策略和口令强度评估机制是使口令存在不安全因素的原因。口令攻击主要分为两大类，猜测攻击和窃取攻击。猜测攻击与口令强度相关，如暴力猜测，字典猜测等。而窃取攻击与口令强度不相关，如钓鱼网站，木马等。

常见的口令猜测方法主要分为3类，第一类是基于字典的方法，将字典中的词汇直接或通过规则变换作为口令猜测集，但由于需要人工制定变换规则，因而猜测结果一定程度上受限于个人经验。第二类是基于统计学规律的方法，基于概率进行口令攻击，根据口令的关系和结构进行模型建立。深度学习近年来逐渐成为第三种口令猜测的方法。

基于字典的方法

HashCat[[[8]](#endnote-7)]

Hashcat号称是世界上最快的密码破解软件，可调用GPU进行密码破解,支持多种hash散列算法，支持对rar、office、pdf、windows账户、wifi等多种密码的破解。Hashcat中有一种很强大的基于规则的攻击方式，利用不同的“规则”，可产生大量用来破解的密码，大大扩充破解字典，而且这些“规则”可以交叉利用，利用一些简单的字典就可衍生出大量复杂但规律的密码以方便破解。

John the Ripper[[[9]](#endnote-8)]

John the Ripper免费的开源软件，用于在已知密文的情况下尝试破解出明文的破解密码软件，支持目前大多数的加密算法，如DES、MD4、MD5等。它支持多种不同类型的系统架构，包括Unix、Linux、Windows、DOS模式、BeOS和OpenVMS，主要目的是破解不够牢固的Unix/Linux系统密码。

John支持四种密码破解模式，字典模式:在这种模式下，用户只需要提供字典和密码列表用于破解；单一破解模式:这是john作者推荐的首选模式。John会使用登录名、全名和家庭通讯录作为候选密码；递增模式:在该模式下john会尝试所有可能的密码组合；外部模式:在这种模式下，用户可以使用john的外部破解模式。使用之前，需要创建一个名为(list.external:mode)的配置文件，其中mode有用户分配。

基于统计学规律的算法

漫步攻击算法

不关⼼攻击对象的信息，⽽只关注在允许的猜测次数内，猜测出更多的⼝令。基于PCFG的算法和Markov算法是⽬前主流的2种漫步攻击算法。

Markov

2005年，Narayanan和Shmatikov[[[10]](#endnote-9)]首次将Markov链引入口令猜测，提出Markov模型。Markov模型常用于自然语言处理中。如果被要求随机选择一系列字符，无论是不是作为口令，用户很可能会生成一个其中每个字符与它在英文文本中出现的频率大致相等分布的序列，远不是一个随机序列。Markov模型是基于Markov链提出的攻击算法，此模型假设用户在构造口令时是按照从前向后构造的。Markov模型是对整个口令进行训练，通过类似自然语言处理的办法从左到右计算口令的概率，分为训练和测试两个阶段。

阶是Markov模型中的一个重要概念，n阶Markov模型需要统计长度为n的字符串之后紧跟着的字符频数。例如，口令Mar567需要统计的值有：首字母是M的频数、M后面是a的频数、Ma后面是r的频数、Mar后面是5的频数、Mar5后面是6的频数，ar56后面是7的频数。按照这种方式，每个字符串口令在经过n阶Markov训练后都能得到一个概率值。从前到后将长度为n的子字符串放到结果中查询，将所有概率想成便可得到目标口令的概率值。上述例子的概率计算为：

按照概率递减排序就可以得到一个猜测集。在测试阶段用得到的猜测集对口令集进行破解测试。在150个真实用户密码数据库上进行验证，超过三分之二的密码被恢复。

在此基础上，许多研究者陆续出新的改进方法，如2014年Ma,Yang,Luo[[[11]](#endnote-10)]等人首次对Markov模型进行正则化和平滑处理。正则化技术是为了使得攻击算法生成的猜测的概率总和始终为1；平滑技术是为了消除数据集中过拟合的问题。

PCFG

2009年，Weir,Shmatikov[[[12]](#endnote-11)]等人首次提出一个基于概率上下文无关文法（PCFG）的口令猜测模型。该算法的核心思想是将口令按字母（L）、数字（D）以及特殊字符（S）进行独立分段，并对每个段的长度今次那个计数表示。比如口令Mar&567，在PCFG模型中按照前3中字符类型可切分成L3，D3，S1。想要得到口令的概率值，需要所有口令中以L3D3S1为模式的口令频率，Mar在长度为3的字母串中的频率，567在长度为3的数字串中的频率，以及&在长度为1的特殊字符串中的频率。所以上述例子的概率计算为：



统计出各自部分对应的概率即可进行概率的计算。通过这样的方法可以计算每个口令的概率值，按照概率递减排序即可获得一个猜测集。在测试阶段用得到的猜测集对口令集进行破解测试，便可评估准确率。

NLP

2014年，Veras等人[[[13]](#endnote-12)]将NLP应用于密码样本的分割和分类，提出了首个密码分割、语义分类和语义融合的框架和一个捕获密码样本语义和语法的模型。利用提出的能够捕获密码语义和语法的模型，将分割的密码进行语义分类。建立一个英语语料库，基于语料库，用自然语言处理的方法，首先将口令进行分割，然后进行词性标注、语义分类，对口令的含义进行分析。猜测生成阶段与PCFG相同。在限制30亿次的猜测中，该方法可从LinkedIn泄露中猜测超过67%的密码，从MySpace泄露中猜测超过32%的密码。

定向攻击算法

攻击者会利用攻击对象相关的个人信息如姓名，生日，年龄，其他网站使用过的旧口令以及泄露的口令等尽可能快的猜测出用户在指定平台上的口令。

Personal-PCFG

2016年，Li[[[14]](#endnote-13)]等人提出了基于PCFG的定向攻击猜测算法。将口令按照字符类型长度进行切分，除 了 PCFG的“L”、“D”和“S”符号，还增加了更多的语义符号，包括“B”代表生日，“N”代表姓名，“E”代表电子邮件地址，“A”代表帐户名，“C”代表手机号码，“I”代表身份证号码。在原有的PCFG方法中增加了一个额外的个人信息匹配阶段和一个自适应替换阶段。例如，口令“helloAlice816！”，如果已知Alice在1988年8月16日出生，在个人信息匹配阶段和预处理阶段后，该口令将被切分为L5N5B3S1。猜测生成阶段，与PCFG类似，以递减的概率，生成包括实际猜测和部分Personal-PCFG语义符号的部分猜测，例如在此步可将L5N5B3S1实例化为“helloN5B3!”。在自适应替换阶段，将第三步猜测得到的信息进一步用个人信息来实例化，例如，“helloN5B3！”，“N5”会直接换成“Alice”。然而，由于“B3”有许多候选段，“19880816”中任何长度为 3 的子串都可能是候选项 ， 猜 测 包 括 所 有 子 串 ，如“helloalice198！”，“helloalice988！”，“helloalice816！”。然后逐一尝试，直到我们发现有一个候选人与 Alice 的密码完全匹配。

在中等规模500000次猜测的离线攻击中，Personal-PCFG达到了PCFG超2亿次猜测才能达到的成功率。在线猜测中，对每个目标账户的猜测次数限制为100次，Personal-PCFG能够比PCFG多破解309%-634%的密码。

TarGuess猜测框架

2016年，Wang，Zhang[[[15]](#endnote-14)]等人提出了TarGuess猜测框架。TarGuess分为3个阶段，准备，训练，猜测。

将个人信息分为三类，第一类是半公共的个人信息PII，对熟人公开，对陌生人保密，如姓名，性别等；第二类是用户标识凭证，一部分是公共的，如用户名，而另一部分是完全私有的，如密码。第三类是剩余的用户个人数据。将用户的半公共个人信息PII进一步分成两类。第一类PII-1是包括姓名，生日等可以成为密码构建模块的个人信息，第二类PII-2包括性别受教育程度等，会影响用户密码的行为但并不能直接用于密码的组成部分。

TarGuess捕获了四种最具代表性的目标在线猜测场景，并运用PCFG，Markov，Bayesian理论等。第一种是只有PII-1的信息，第二种是利用公开信息和一个从其他应用泄露的姊妹密码，第三种是PII-1和一个从其他应用泄露的姊妹密码，第四种是PII-1，PII-2和一个从其他应用泄露的姊妹密码。

TarGuess-I仿照了Personal-PCFG中的6种PII标签（姓名N，生日B，电子邮件前缀E，账户名称A，电话号码P，I表示NID）。使用12306数据集进行评估，结果表明电子邮箱前缀、账户名称、姓名、生日四个信息对攻击者很有价值，电话号码和NID只是略微提高了猜测成功率。

TarGuess-II使用数据驱动的方法，利用公开信息和姊妹口令，通过观察重用口令对，对口令进行结构级和字段级变换。相同的PW对被消除，建立新的不相同的密码对{PWA，PWB}。分析PWA如何调整成PWB，或者PWB是否是高频次密码。

TarGuess-III需要处理更多的信息。引入基于PII的标签（TarGuess-I）到TarGuess-II，可构建一个PII更丰富的，基于密码重用的模型。TarGuess-III只尝试100次猜测，仍然可以达到23.48%的成功率。在未破解的PW对中，超过80%有显著性差异。

TarGuess-IV因为PII-2的信息不能使用基于PII的PCFG或Markov模型直接测量，通过证明“分治”并引入贝叶斯理论解决。TarGuess-IV可比TarGuess-III提高4.38%。

基于深度学习的方法

RNN

2016年，Melicher等人[[[16]](#endnote-15)]首次提出使用循环神经网络（Recurrent Neural Network,RNN）进行口令破解。现有的通过模拟对抗密码猜测来评估密码强度的方法，是不准确的或对于实时客户端密码检查数量级太大、太慢。

神经网络是一种用于逼近高维函数的机器学习技术。它们被设计用来模拟人类神经，尤其擅长模糊分类问题和生成新序列。神经网络可以被高度压缩到几百千字节，而不会大大降低猜测的有效性，可以基于前面的元素生成字符串中下一个元素的概率，可以在比马尔可夫模型小得多的空间内模拟自然语言。

Melicher还试验了一种利用迁移学习优势训练神经网络的特殊方法，在这种方法中，神经网络的不

同部分在训练过程中学习识别不同的现象。当使用迁移学习时，首先对训练集中的所有密码训练模型。然后，模型的下层被冻结。最后，只对训练集中合适的密码重新训练该模型。

对于测试的所有密码集，神经网络从大约1010次猜测开始就优于其他模型，并在此之前匹配或击败了其他最有效的方法。

2018年，Zhou等人[[[17]](#endnote-16)]将个人信息与RNN结合，提出了定向口令猜测模型TPGXNN，猜测成功率高于同样场景下的PCFG和Markov模型。

LSTM

2017年，Xu[[[18]](#endnote-17)]等人使用长短期记忆网络（Long Short-Term Memory,LSTM）进行口令破解，根据输入的序列预测下一个字符。神经网络包括一个输入层，几个 LSTM 层，一个 softmax 层和一个输出层。LSTM 模型包含 2 个隐藏的 LSTM 层，每层有 256 个神经元，由 3000 万个 Rockyou（RY）密码训练，包括 2700 万的训练集和 300 万的验证集。LSTM模型又可以分为单词级和字符级，其生成口令的过程与Markov方法类似，两种方法均是计算字符产生概率来生成猜测口令集。在口令猜测中需要提前设置固定的序列长度L，在训练阶段根据长度L截取训练集。神经网络经过学习训练来得到生成的各字符的概率并非基于简单的已有样本统计。在测试集上测试，破解率优于PCFG和Markov的模型。

GAN

2019年，Hitaj等人[[[19]](#endnote-18)]首次提出利用生成式对抗网络（Generative Adversarial Networks,GAN）来破解口令,他们将模型命名为PassGAN。一个GAN由两个深度神经网络组成：一个学习真实样本并生成假样本的深度神经网络G和一个用于判断是真实样本还是G生成的假样本的深度神经网络D。这两个深度神经网络通过多次迭代而相互作用。在每次迭代中，G的假样本给D，然后将D的输出提供给G，G将其作为反馈，生成假样本，并分布得越来越接近真实样本。经过足够次数的迭代后，G的输出就变成了GAN的输出。PassGAN利用这种技术来生成新的密码猜测。我们的核心想法是使用一个真实的泄露的密码列表来训练D。因此，在每次迭代中，PassGAN假样本的输出更接近原始泄漏中的密码分布，因此更有可能匹配真实用户的密码。

当GAN在RockYou的不同子集上进行训练时，我们能够在从RockYou密码数据集中提取的测试集中匹配超过46%的密码。而通过将GAN的输出与best64规则输出相结合，能够匹配RockYou数据集中超过78%的密码。

其他

随着研究的深入，口令猜测研究不仅有对原有模型进行了更进一步的研究和对两个及以上模型结合进行口令猜测研究，还有基于机器学习和统计学规律的口令猜测研究等，如PCFG+LSTM[[[20]](#endnote-19)]，PCFG+RNN[[[21]](#endnote-20)]，MLSTM[[[22]](#endnote-21)]，PCFG+GAN[[[23]](#endnote-22)]，基于随机森林经典机器学习[[[24]](#endnote-23)]的口令猜测算法研究，动态字典攻击[[[25]](#endnote-24)]等等。

总结与展望

本文总结了用户脆弱口令行为、口令理论分布模型、口令猜测方法三个方面，概述了研究思路。当前在口令安全的很多领域，都需要进一步的研究，存在大量机遇。

1. 口令集的泄露检测

很多网站都存在口令泄露的风险，甚至很多网站在口令泄露多年后才察觉到口令集泄露，但此时告知用户口令泄露为时已晚。显然，及时察觉口令集的泄露非常关键。如何及时察觉口令集的泄露，如何进一步降低用户口令集泄露风险，怎么高效的对抗攻击者的攻击，需要进一步的研究。

1. 泄露口令的猜测攻击

攻击者的手段越来越多，但网站监测和防御口令泄露的方法比攻击者要少得多。口令具有使用简单，成本低廉的特点，用户和网站普遍喜爱使用口令，但绝大部分用户的口令在不知不觉中都曾被泄露过。攻击者如何利用收集到的口令，在旧有的模型上改进或者提出新的模型，更加准确的进行口令猜测？攻击者可以利用用户哪些新的信息进行有效的口令猜测？都是可以研究的方向。

1. 口令强度的评价

现阶段，用户创建网站口令时，网站给出的口令强度评价主要基于大小写字母，数字，特殊字符，口令长度，是否是历史的高频次口令等。但显然，这种评价方法并不是真的意味着用户口令是高强度的，用户在创建口令时，基于人的生理和心理多方面因素，创建的口令往往是易于破解的。

基于规则的口令强度评价方法比较简单，现在得到最广泛应用的NIST PSM[[[26]](#endnote-25)]评价思想，主要基于长度和字符类型，但非常明显的缺陷就是高强度和低强度口令非常不准确。基于模式检测的口令强度评价方法，主要根据键盘模式、常见语义、顺序字典模式等进行评判，相较NIST PSM评价改进许多，但依然存在判断不准确的问题。基于攻击算法的口令强度评价方法，用攻击的难易进行判定，提出了如Morkov-based PSM[[[27]](#endnote-26)]、PCFG-based PSM[[[28]](#endnote-27)]、fuzzyPSM[5]、vec-PPSM[[[29]](#endnote-28)]等口令强度评价方法。如何改进现有模型或提出新的模型，更准确的评价口令强度，基于漫步攻击及定向攻击口令强度评价模型，将机器学习用到口令强度评价模型中，都是值得进一步研究的方向。

1. 口令存储

1. [] 科普：计算机密码的诞生https://baijiahao.baidu.com/s?id=1732146348594355626 [↑](#endnote-ref-0)
2. [] 2022 Data Breach Investigations Report | Verizon [↑](#endnote-ref-1)
3. [] 2021 Data Breach Investigations Report | Verizon [↑](#endnote-ref-2)
4. [] Wang D,He D,Cheng H,et al.fuzzyPSM:A new password strength meter using fuzzy probabilistic context-free grammars//Proceedings of 2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN),(2016):595-606 [↑](#endnote-ref-3)
5. [] Wang D,Cheng H,Wang P,et al.Zipf’s law in passwords.IEEE Transactions on Information Forensics and Security,2017,12(11):2776-2791 [↑](#endnote-ref-4)
6. [] Malone D,Mather K.Investigating the Distribution of Password Choices//Proceedings of the 21st International Conference on World Wide Web.New York,USA,2012:301-310 [↑](#endnote-ref-5)
7. [] Hou Z D,Wang D,New Observations on Zipf’s Law in Passwords//Proceedings of the IEEE Transactions on Information Forensics and Security,18(2023):517-532 [↑](#endnote-ref-6)
8. [] HashCat https://hashcat.net/hashcat/ [↑](#endnote-ref-7)
9. [] John the Ripper https://www.openwall.com/john/ [↑](#endnote-ref-8)
10. []Narayanan A,Shmatikov V.Fast dictionary attacks on passwords using time-space tradeoff//Proceedings of the 12th ACM Conference on Computer and Communications Security.Alexandria,USA,2005:364-372 [↑](#endnote-ref-9)
11. [] Ma J,Yang W,Luo M,et al. A study of probabilistic password models//Proceedings of the IEEE Symposium on Security and Privacy. Oakland,USA,2014:689-704 [↑](#endnote-ref-10)
12. []Weir M,Aggarwal S,De Medeiros B,et al.Password Cracking Using Probabilistic Context-Free Grammars//Proceedings of the 30th IEEE Symposium on Security and Privacy. Oakland,USA,2009:91-105 [↑](#endnote-ref-11)
13. [] Veras R,Collins C,Thorpe J.On semantic patterns of passwords and their security impact//Proceedings of Network and Distributed System Security Symposium.San Diego,CA:Internet Society,2014:1-16 [↑](#endnote-ref-12)
14. [] Li Y, Wang H,Sun K.A study of personal information in human-chosen passwords and its security implications//Proceedings of the IEEE International Conference on Computer Communications.San Francisco,USA,2016:1-9 [↑](#endnote-ref-13)
15. [] Wang D,Zhang Z,Wang P,et al.Targeted Online Password Guessing :An Underestimated Threat//Proceedings of the 2016 ACM CCS 2016. New York:ACM,2016:1-13 [↑](#endnote-ref-14)
16. [] Melicher W,Ur B,Segreti S M,et al. Fast,lean,and accurate:Modeling password guessability using neural networks//Proceedings of the 25th USENIX Security Symposium.Austin,USA,2016:175-191 [↑](#endnote-ref-15)
17. [] Zhou Huan,Liu Qi-Xu,Cui Xiang,et al.Research on targeted password guessing using neural networks.Journal of Cyber Security,2018,3(5):25-37(in Chinese)

    （周环，刘奇旭，崔翔等．基于神经网络的定向口令猜测研究．信息安全学报，2018,3（5）：25-37） [↑](#endnote-ref-16)
18. []Xu L,Ge C,Qiu W,et al.Password guessing based on LSTM recurrent neural networks.//Proceedings of the IEEE International Conference on Computational Science and Engineering(CSE)and IEEE International Conference on Embedded and Ubiquitous Computing(EUC),Guangzhou,China,2017:785-788 [↑](#endnote-ref-17)
19. []Hitaj B,Gasti P,Ateniese G,et al. PassGAN:A deep learning approach for password guessing//Proceedings of the International Conference on Applied Cryptography and Network Security.Bongotá,Colombia,2019:217-237 [↑](#endnote-ref-18)
20. [] Liu Y,Xiu Z,Yi P,et al.GENPass:A general deep learning model for password guessing with PCFG rules and adversarial generation//Proceedings of the IEEE International Conference on Communications(ICC).Missouri,USA,2018:1-6 [↑](#endnote-ref-19)
21. [] Wang D, Zou Y K, Tao Y, et al. Password Guessing

    Based on Recurrent Neural Networks and Generative

    Adversarial Networks. Chinese Journal of Computers,

    2021, 44(08):1519-1534(In Chinese)

    (汪定,邹云开,陶义,王彬.基于循环神经网络和生成式对抗网络的口令猜测模型研究.计算机学报,2021,44(08):1519-1534.) [↑](#endnote-ref-20)
22. []Chang Geng,Zhao Lan,Chen Wen. MLSTM:A password guessing method based on multi-sequence length LSTM.Computer Science,2022,49(04):354-361.

    (in Chinese)

    (常庚,赵岚,陈文.MLSTM:一种基于多序列长度LSTM的口令猜测方法.计算机科学,2022,49(04):354-361.) [↑](#endnote-ref-21)
23. [] Xia Z Y,Yi P,Liu Y Y,et al. GENPass:A Multi-Source Deep Learning Model for Password Guessing.IEEE Transactions on Multimedia,2019,22(5):1323-1332 [↑](#endnote-ref-22)
24. [] Password Guessing Using Random Forest.USENIX Security 2023 [↑](#endnote-ref-23)
25. [] 张学旺,孟磊,周印.动态字典破解用户口令与安全口令选择.计算机应用研究,2020,37(04):1166-1169.DOI:10.19734/j.issn.1001-3695.2018.09.0752. [↑](#endnote-ref-24)
26. [] Weir M,Aggarwal S,Collins M,et al.Testing metrics for password creation policies by attacking large sets of revealed passwords// Proceedings of the 17th ACM Conference on Computer and Communications Security, CCS 2010, Chicago, Illinois(2010):162-175 [↑](#endnote-ref-25)
27. []Castelluccia C, Dürmuth M,Perito D. Adaptive Password-Strength Meters from Markov Models.//Proceedings of the Network and Distributed System Security Symposium ,2012:1-15 [↑](#endnote-ref-26)
28. [] Houshmand S, Aggarwal S.Building better passwords using probabilistic techniques.//Proceedings of the Asia-Pacific Computer Systems Architecture Conference ,2012:109-118. [↑](#endnote-ref-27)
29. [] Pal B, Daniel T, Chatterjee R, Ristenpart T. Beyond credential stuffing: password similarity models using neural networks. 2019 IEEE Symposium on Security and Privacy (SP). San Francisco, CA, USA: IEEE, 2019: 417–434. [↑](#endnote-ref-28)