计算机研究与发展'16 - 口令安全研究进展

优点：①全面介绍了口令作为最主要的身份认证方式的发展历程，从用户脆弱口令行为，猜测攻击，口令分布安全性评价指标和口令强度评测分别做介绍 ②对口令安全的未来方向做了介绍，包括在Zipf下刻画敌手优势、基于隐语义LDA模型的漫步猜测攻击、基于泄露口令的定向猜测攻击，基于定向攻击模型的口令强度评测、服务器口令集泄露的检测 ③使用真实数据集论述，可信度高

问题：①文章采用了真实数据集，应简要说明不同数据的选择理由，以及科研伦理问题 ②口令重用部分对比不同安全级网站用户口令重用的情况应更合适．③文章部分图表的介绍不够具体．

身份认证是确保信息系统安全的第一道防线，口令是应用最为广泛的身份认证方法．由于口令具有简单易用、成本低廉，容易更改等特性，在可预见的未来仍将是最主要的认证方法．

▲ 口令研究的难点：口令是人生成的，与人的行为直接相关，每个人的行为因各种因素而千差万别．

▲ 口令安全研究根据研究方法大致分为3个阶段：

① 1999年以前，采用启发式方式，没有理论体系

② 2000年～2008年，口令理论体系初现端倪，大多集中于揭示口令的弱点

③ 2009年以来，口令安全理论体系逐渐完善，形成了概率攻击理论模型，口令分布强度评价理论模型和口令分布理论模型．

用户的脆弱口令行为 → 口令无法达到理想强度的直接原因（实证分析 & 用户调查）

1> 用户往往需要管理几十上百个口令帐户，且不同网站的口令设置要求往往差异较大．

2> 用户用于处理信息安全事务的精力有限，且不会随时间推移有较大提高

▲通过8个知名真实口令集，将用户脆弱口令行为归为3类：

1> **口令构造的偏好性选择**：①语言对口令行为影响很大，高达1.01%～10.44%的用户选择最流行的10个口令 ②口令并非符合随机均匀分布（缺乏实证数据且均匀分布易于分析）根据大数定律，去除低频次口令后，发现Zipf模型可以很好地刻画口令分布，其中 $f_r = \frac{C}{r^s}$，r为排名，f_r表示排名为r的口令的频率，C和S为常数，由具体分布（数据集）决定

▲ 口令频次呈多项式下降，高频的口令和低频的口令占据整个口令集的重要部分．

③字符组成结构：当设置口令策略时，由策略决定；未设置策略时体现用户偏好（敌手挖掘）

④口令长度：也受网站策略影响，若未设置长度限制，则受网站服务类型影响（减少猜测空间）

2> **口令重用**：重用是用户理智的做法，只有跨不同安全级（或重要程度）重用口令，才应该避免的

使用文本相似度算法分析新旧口令，表明多数修改幅度较大；不同网站之间，使用Levenshtien相似度算法表明中文用户在不同网站之间的口令重用问题更严重

3> **基于个人信息构造口令**：用户使用个人信息构造口令的习惯严重降低了口令强度，定向攻击者可依此大大增强其效率．

口令猜测攻击（成功攻击该口令所需要的猜测次数）

▲漫步攻击算法（trawling attacking）：不关心具体的攻击对象是谁，唯一目标是在允许的猜测次数下，尽可能多的猜测口令．

①启发式算法：难以重现，难以相互公平比较．

②PCFG：假设口令的字母段L、数字段D和特殊字符段S是相互独立的，在训练阶段统计口令模式频率表Σ_1和字符组件（语义）频率表Σ_2．在猜测集生成阶段生成带频率猜测的集合，模拟现实中口令的概率分布．

③Markov：核心假设用户构造口令从前向后依次进行，对整个口令训练，通过从左到右的字符之间的联系计算口令概率．有阶的概率，需记录长度为n的字符串后面跟随的字符频数

▲引入了平滑技术用于消除数据集中过拟合问题；引入正规化技术使攻击算法所生成的猜测概率总和始终为1，形成一个概率模型．

④NLP：对口令中的语义做分析：对PCFG的L做分析，主要有2个核心点：分词与词性标注

▲定向攻击算法（targeted attacking）：利用与攻击对象相关的个人信息增强猜测的针对性

①Targeted Markov：将PI分成6大类，并对每个大类根据需要的粒度细分，在训练时将PI信息替换为基本字符，在猜测集生成阶段再替换回来．（猜测次数为10～1000时攻击效率更高）

②Personal-PCFG：将PI分成6大类，与原本漫步攻击的LDS同等地位，思路与Markov类似

口令分布安全性评价指标（评价口令分布安全性，比采用攻击算法评定更具有确定性）

设分布为X，样本空间大小为N，事件概率大小为 $p_1 \ge p_2 \ge \cdots \ge p_N$．

①**信息熵**：广泛用来测量一个分布的不确定性 $H_1(X) = \sum_{i=1}^{N} -p_i \times \log p_i$

②**最小熵**：用来测量一个分布中概率最大事件出现的不确定性 $H_\infty(X) = -\log p_1$

③**猜测熵**：用来测量当一个漫步攻击者按最优方式攻击时，猜测出X中任一元素需要的平均猜测数 $G(X) = \sum_{i=1}^{N} i \times p_i$

④**β-success-rate**：测量当一个漫步攻击者被限制至多猜测β次时平均成功率 $\lambda_\beta(X) = \sum_{i=1}^{\beta} p_i$

⑤**α-work-factor**：测量当一个漫步攻击者想达到至少α的成功率时对每个帐户至少发起的猜测数：$\mu_\alpha(X) = \min\{j \mid \sum_{i=1}^{j} p_i \ge \alpha\}$．

⑥**α-guesswork**：▲测量当一个漫步攻击者想达到α的成功率时，对每个帐户平均发起的猜测数目：$G_\alpha(X) = (1-\lambda_{\mu_\alpha}) \times \mu_\alpha + \sum_{i=1}^{\mu_\alpha} p_i \times i$．

②④及小成功率的⑥衡量口令分布抵抗在线猜测攻击的能力，①和大成功率的⑥衡量离线

口令强度评测

①基于规则：NIST PSM，例如口令长度和字符类型．简单但不够准确

②基于模式检测，如Zxcvbn，考虑键盘、常见语义、顺序字符等综合评分，更科学但缺乏理论性

③基于攻击算法，例如fuzzyPSM，PCFG-based PSM，Markov-based PSM．其中fuzzyPSM考虑用户在注册时可能使用原来的口令