大数据环境下的网络身份安全挑战及解决方案

刘文印　　广东工业大学网络身份安全粤港联合实验室 WIS Lab

2018 ISC 互联网安全大会 中国 · 北京
Internet Security Conference 2018 Beijing · China
（原中国互联网安全大会）

提纲

网络身份 & 网络身份安全

网络身份安全遇到的挑战

网络身份安全技术及解决方案

- 基于可信用户代理的多方闭环网络身份认证管理机制
  - 自动注册、自动登录、自动修改密码；防撞库、防钓鱼、解决“密码疲劳”
- 基于寄生社区分割的钓鱼网站及目标识别
  - 防钓鱼欺诈、防身份被盗

网络身份

- 区分、认定（Identify）网络空间中个体（Entity）的**唯一性和不可否认性标识**
  - 个体：个人用户、组织机构、设备、网络资源

- 网络身份信息：跟身份有关的信息
  - **用户名、口令**、身份证号、邮箱、手机号、卡号、微信号、QQ等
  - 域名、名称、组织机构代码等
  - 序列号、URI等

网络身份安全

- 准确**认证**和**管理**网络身份
- **检测**并**防止**
  - 身份造假、假冒
  - **被盗**（防拖库、防钓鱼）
  - **泄露**（防拖库、防撞库）
  - 其他未授权行为

网络身份安全遇到的挑战

- 大数据环境
  - 公开数据被利用
  - 私有数据被泄露
- 内部威胁（自身原因）
  - 内鬼、系统漏洞、管理不善（制度缺陷、犯错失职）、用户习惯
- 外部威胁（黑客攻击）
  - 钓鱼、假冒、拖库、撞库

大数据环境中的网络身份信息(PII)

- **大数据：信息社会的“数字垃圾”包含身份信息、易遭非法利用（诈骗、窃取、撞库、破解）**

- **公开数据**：
  - 个人主页、机构官网、百科、博客、微博：内容（姓名等）丰富真实!
  - 讣告、新闻、论文：真实的姓名，邮箱、单位信息
- **各网站私有数据：**
  - 注册的id，口令，姓名，生日，手机号，QQ号，邮箱等
  - 社交网络关系 （好友、粉丝、在线习惯等）

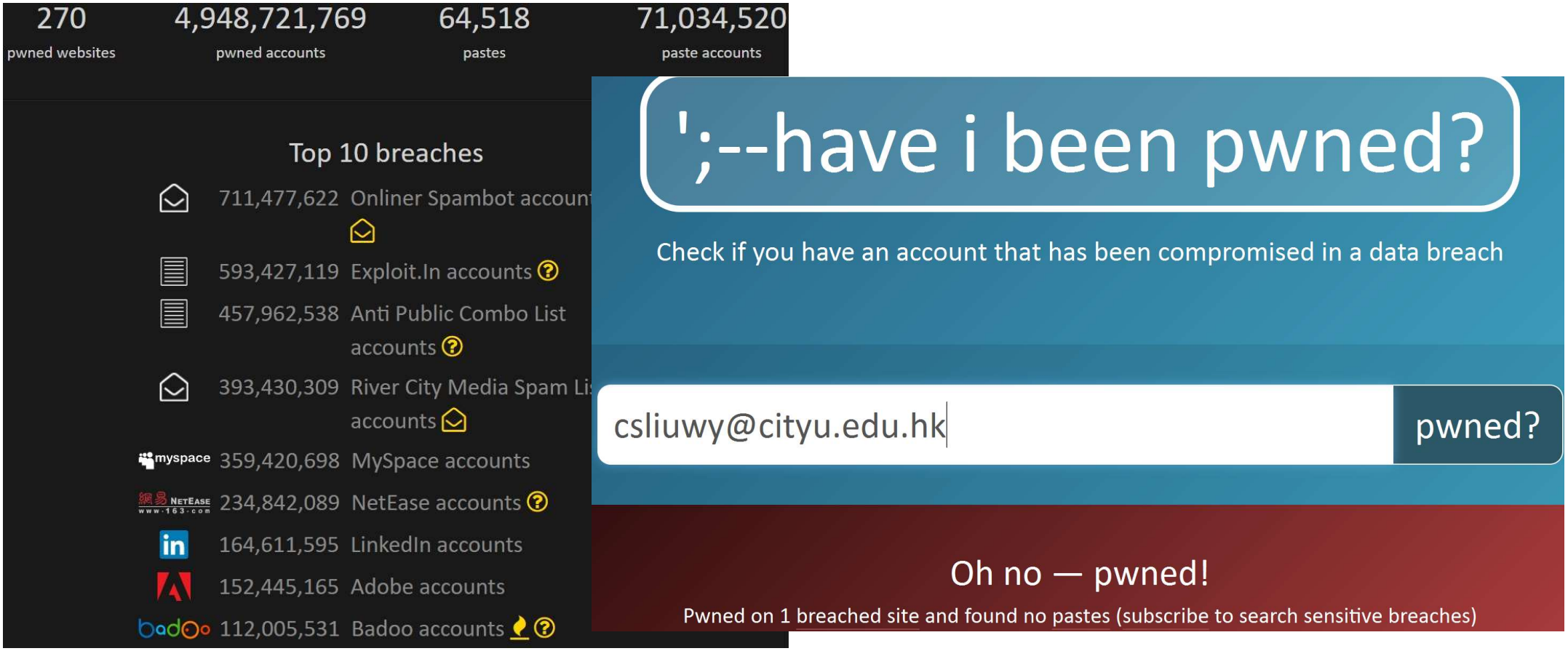
大数据环境中的网络身份安全威胁

- **公开数据蕴含着隐私，被用于破解、撞库**
  - 关联关系：**liuwy@fastersoft.com.cn；liuwy@gdut.edu.cn**
  - 社交网络：奶茶妹妹 – 刘强东
- **私有数据被窃取、泄露，形成黑色产业链：**
  - 拖库：窃取整个身份数据库（账号、密码、其他身份信息）
  - 洗库：解密、整理，分类，出售里边的信息
  - 撞库：用其中发现的密码去猜测登录别的系统

内部威胁 – 数据泄露

- **管理员使用弱密码、多人分享密码、失误**
  - 被撞库、被拖库、被接管、留后门
- **漏洞不及时修补、留后门**
  - 被拖库
- **加密措施不强（甚至明文存储密码）**
  - 拖库后被破解，94%MD5; 74%哈希加盐，0.1%Bcrypt
- **内鬼（80%）**

数据泄露的结果—社工库：haveibeenpwned.com



密码疲劳问题导致的用户习惯 – 不安全

- 密码/口令模式有限：被破解
- 弱密码（常用密码库）：被破解
- 长期不换：被破解
- 账号关联、密码/口令重用：被撞库
- 经验不足，不够警惕，被恐吓、诱惑而轻信：被钓鱼

常用模式 – 利用人性弱点精准分析破解个人密码

- 利用个人信息，猜解各种组合
- 暴力定向破解密码（**汪定**）
  - 7种模式
- PassWord->PassPhrase
  - NIST的Paul Grassi提出
  - 仅仅“长”没有用，仍是有限模式

感谢泰格实验室
姜栋提供：

账号关联、密码/口令重用 - 撞库越来越容易

- **简单撞库针对单个账号，多次猜测密码**
- **“分布式集体轮番撞库” （Credential Stuffing）**
  - 公开获取或非法购买被泄露的大量的账号/密码对
  - 针对同一个网站的多个账号
  - 轮番从多个代理或僵尸网络发起登录请求
  - 辅助自动工具通过图灵测试
  - 1-2%的成功率

网络身份欺诈 / 网络钓鱼攻击

- **网络身份欺诈**
  - 未经他人授权，**假冒**其网络**身份**，特别是用于非法目的
- **网络钓鱼攻击**
  - 通过模仿真实网页，**假冒**其他个人或组织的网络**身份**，**骗取用户的网络身份**等私隐信息

    如：用户名、密码、卡号等
  - 二维码钓鱼

天下网络，无坚不破！天下密码，唯“强”不破！

- **人是最薄弱的环节！**
- **何为“强”密码？**
  - 随机、无规律（无模式）
  - 足够长 （仅仅长还不行，要记住就得有规律，有规律就能破！）
  - 未泄露过、经常换
  - 记不住，输入也不方便，可用性很差！

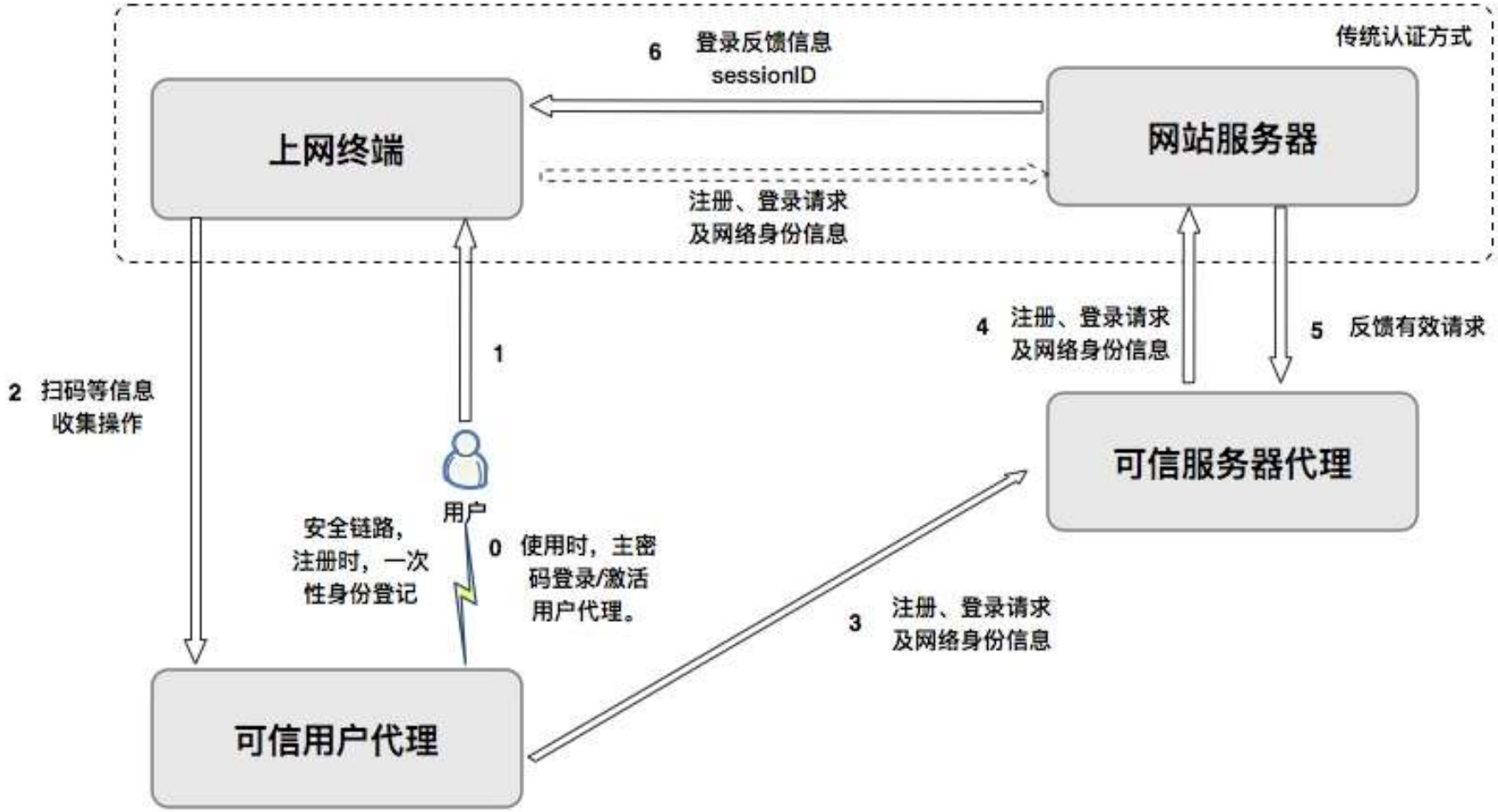
    “能记住的密码都是弱密码”

以用户为中心的身份管理

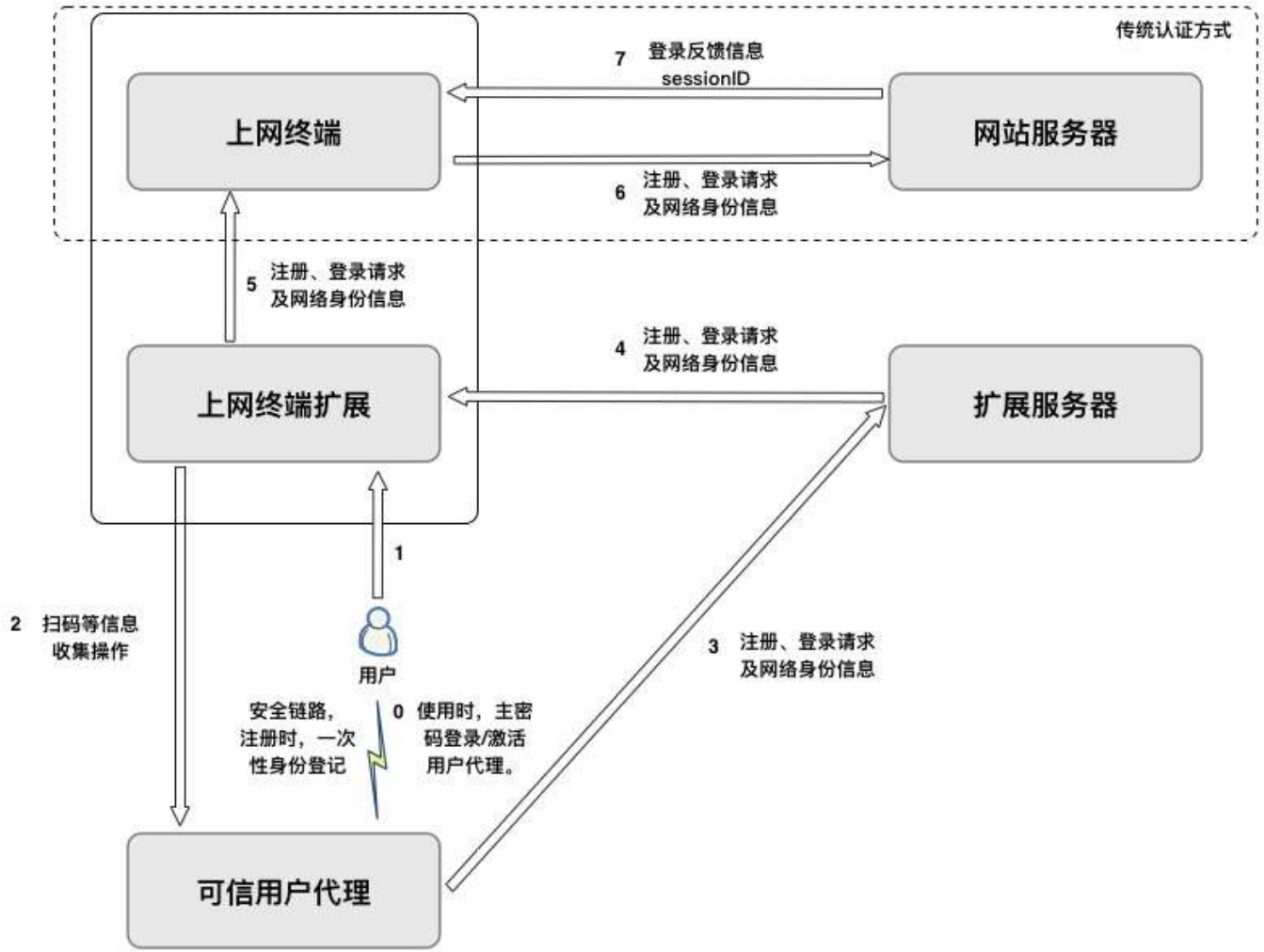
- **以网站为中心 -> 企业为中心（SSO） -> 用户为中心!**
- **基于可信用户代理的多方闭环网络身份认证机制**
  - **兼容传统密码**机制，成本低、易于部署
  - 密码绕过浏览器，直发服务器，杜绝被劫持、被钓鱼
  - 注册随机账号，做到**de-link**，防止被撞库
  - 生成并常换随机强密码，防止被破解
  - 密码不重用，防止被撞库
  - 配套密码管理器，密码再也不用记（“烧脑”）、避免“密码疲劳”

传统B/S架构双向网络身份认证机制

多方闭环网络身份认证机制（可信服务器代理版）



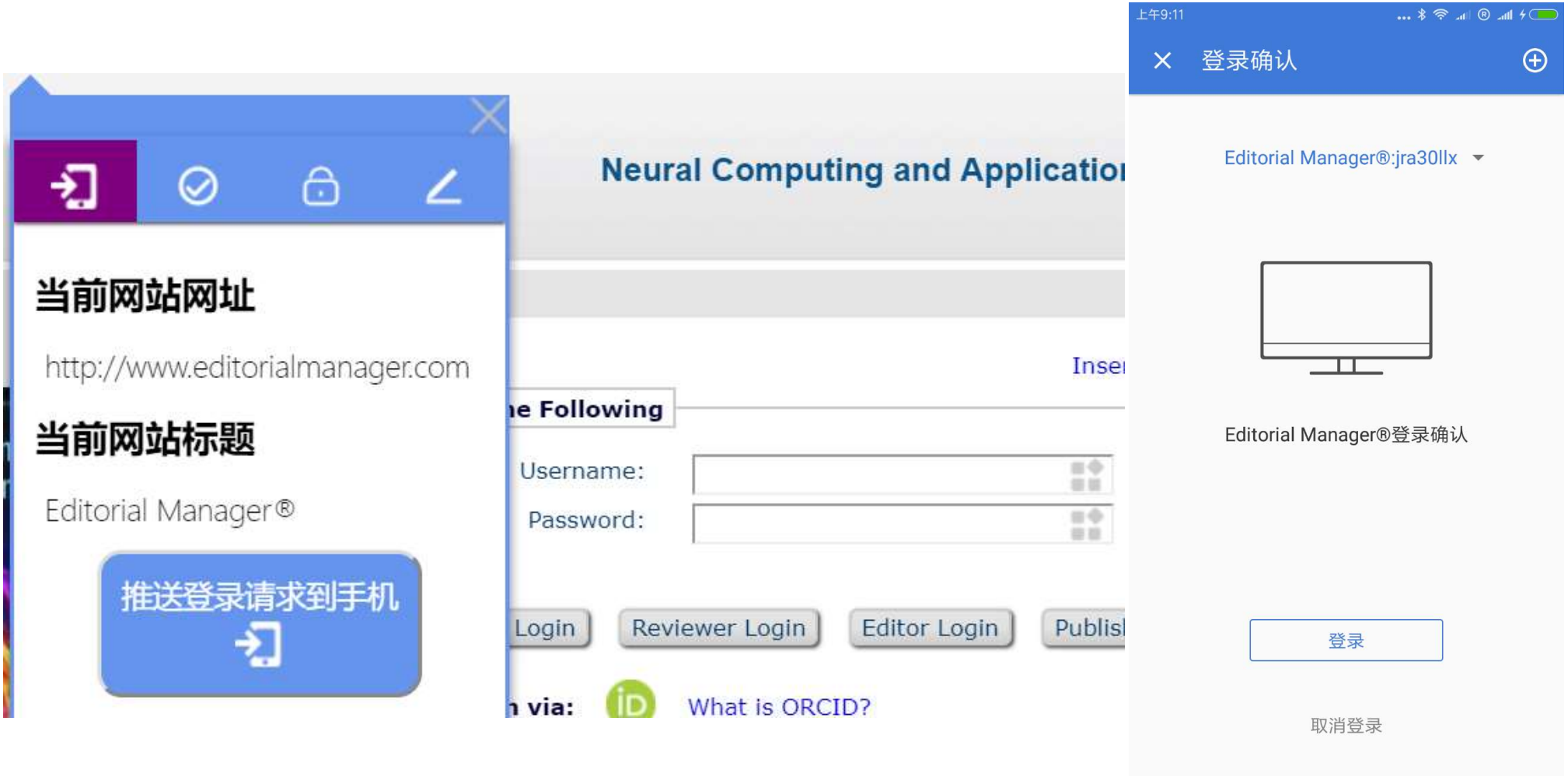
多方闭环网络身份认证机制（插件/扩展版）



自动注册

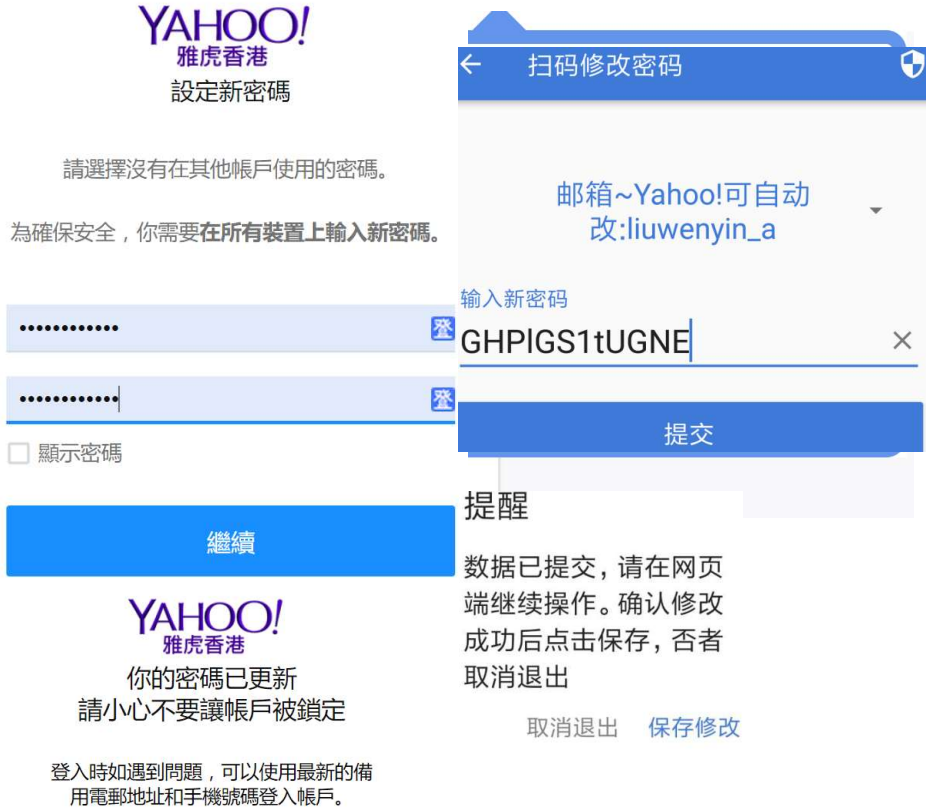
- 首次扫码会自动注册一个随机账户并生成强密码。
- 以后再扫码就自动登录。
- 希望可以取消验证码，降低上网门槛。

自动登录功能


当前网站网址
http://www.editorialmanager.com
当前网站标题
Editorial Manager®
推送登录请求到手机
Neural Computing and Applicatio
Inse
e Following
Username:
Password:
Login
Reviewer Login
Editor Login
Publis
h via:
What is ORCID?
上午9:11
登录确认
Editorial Manager®:jra30llx
Editorial Manager®登录确认
登录
取消登录

自动更换密码

1. 网页上启动更换密码服务，显示二维码
2. 登录易App扫码后自定生成新密码
3. 登录易App提交新旧密码，网页收到新旧密码后提示确认继续修改密码
4. 确认后，网页提示密码已经更新
5. 手机上提示保存新密码，点击“保存修改”！



既 方便　又 安全！

✓兼容传统密码机制

可靠、易部署

✓密码不用记

✓ 自动注册、登录、修改

✓ 密码管理集中化

✓登录信息移动化

• 特别适合不带键盘的智能设备登录

✓自动备份、同步

• 可下载临时App应急

✓ 不同网站的账户**没有关联**

• **不可能遭撞库**

• **保护隐私**（无法知道是同一个人）

✓ 绕过浏览器

• 比密码管理器自动填写表单**更安全**

• **防“钓鱼”**

✓主密码保护

• 可选用指纹、其他生物信息或字符密码**验证自己　合法使用所属设备**

可信用户代理 – 检测钓鱼及钓鱼目标（Phishing Target)

- 从可疑网页中发现蛛丝马迹
- 顺藤摸瓜发现钓鱼目标 - 被假冒的真网站
- 如果可疑网页与目标不相同，则判断可疑网页为钓鱼
- IEEE Internet Computing （2012）
  - 编辑、审评员一致好评： “创新、聪明、妙趣Intriguing、扎实、实用”

寻找被钓鱼攻击目标的方法示意图

I: 缩小包围圈

1. 给定的一个可疑网页

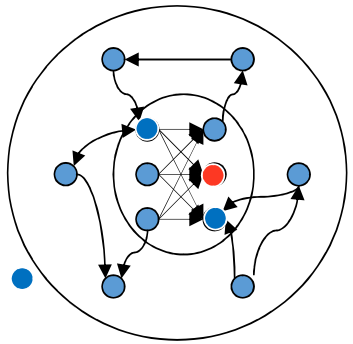
2. 寻找**有关联**的网页集合

3. 构建由关联网页组成的寄生社区

II: 锁定目标

4. 在寄生社区中锁定钓鱼攻击的目标，即被假冒的真网页

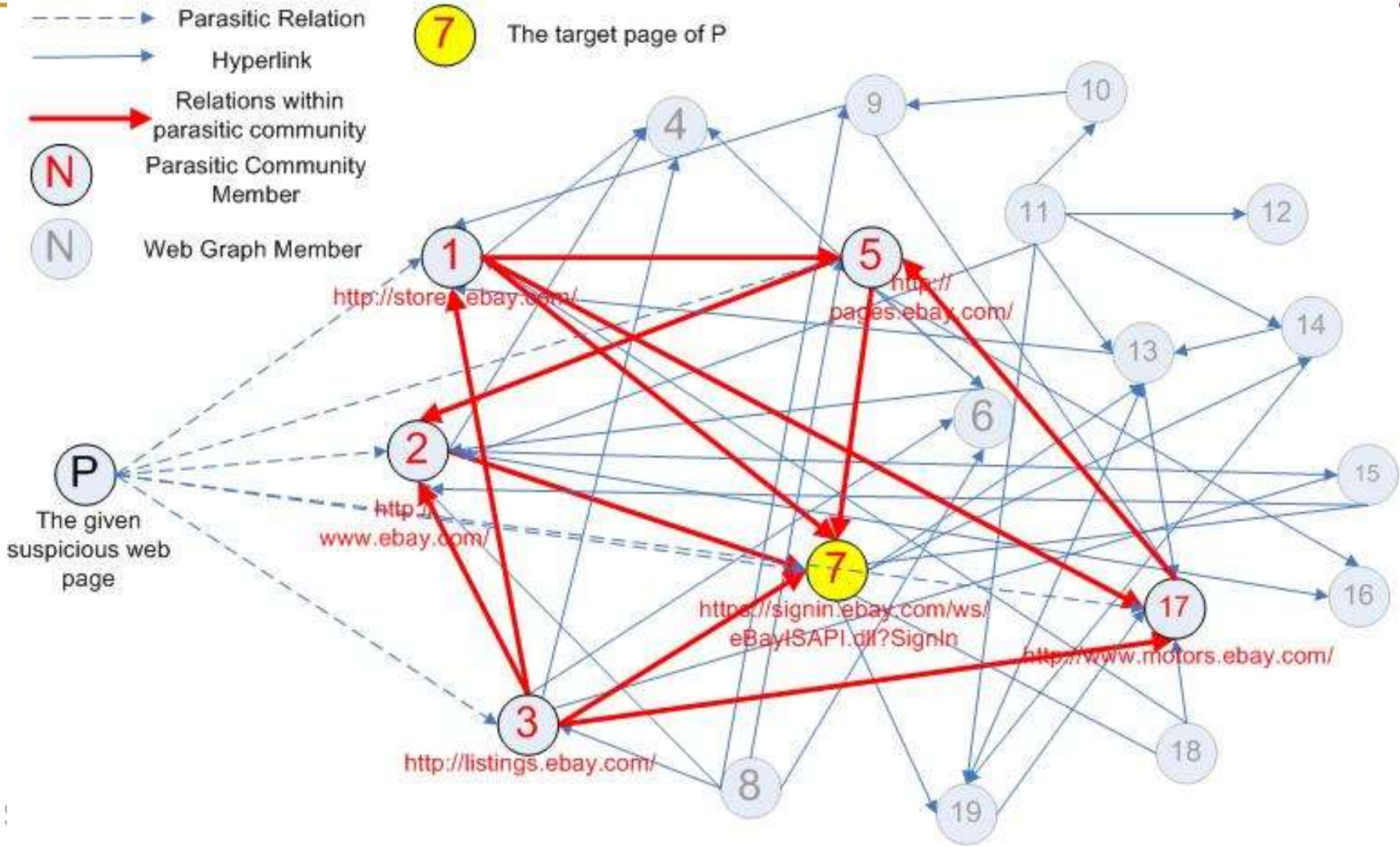
关联关系



关联关系

- **直接关联关系**
  - **链接**
    - 报告（Cova）： 42%的钓鱼网页包含指向真实网页的链接
    - 我们研究： 其中高达85.6% 包含明确指向真实网页的超链接
- **间接关联关系**
  - **搜索引擎中排序**
    - 检索词: 标题, meta标签关键词, 主体关键词
  - **文本/语义相似度**
    - 非对称性相似度 (Tversky)

示例

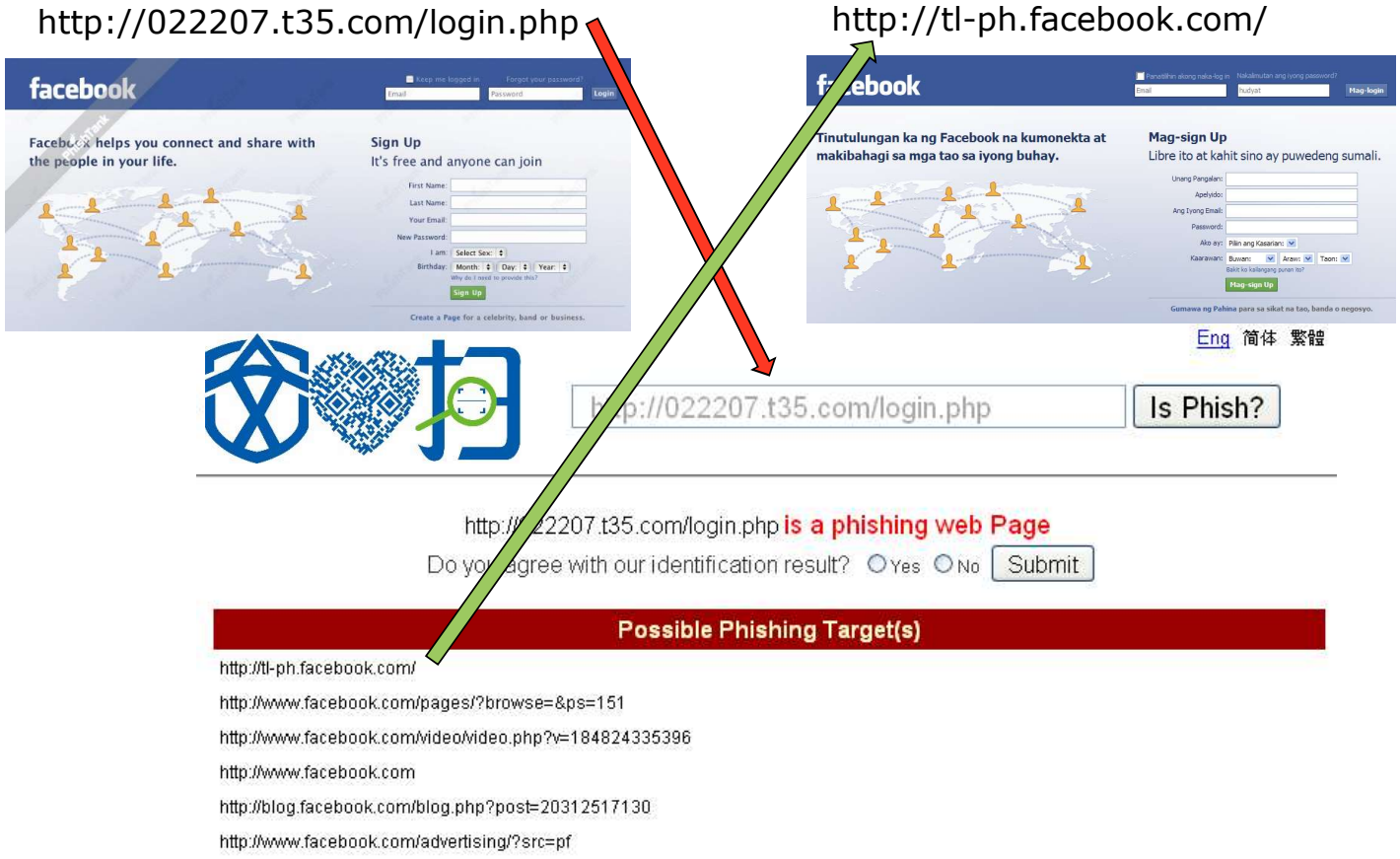

Parasitic Relation
Hyperlink
Relations within parasitic community
N
Parasitic Community Member
N
Web Graph Member
7
The target page of P
P
The given suspicious web page
1
http://stores.ebay.com/
2
www.ebay.com/
3
http://listings.ebay.com/
5
pages.ebay.com/
7
eBayISAPI.dll?SignIn
17
http://www.motors.ebay.com/
4
9
10
11
12
13
14
6
15
16
18
8
19

高准确率

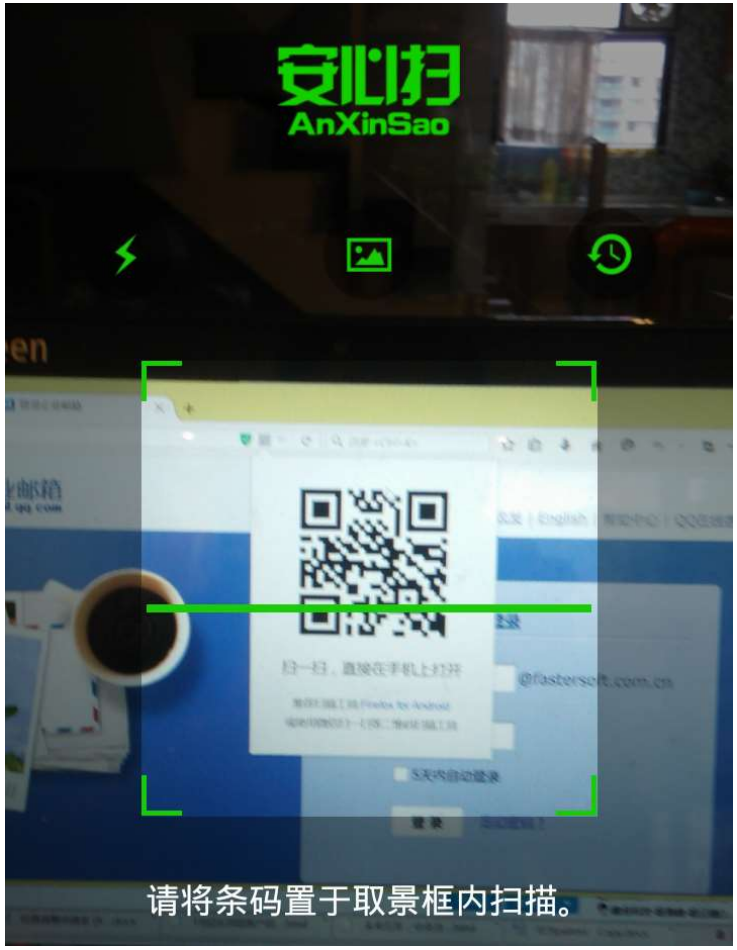
- 数据集
  - ✓ **10005** 假冒钓鱼网页
  - ✓ 3个月收集自PhishTank
- 钓鱼检测准确率
  - ✓ **99.2%**
- 被假冒目标识别率
  - ✓ **92.1%**
- 误报率
  - ✓ Alex Top 1,000,000
  - ✓ Yahoo Random Links
    - 未过滤误报率2.2%
    - 过滤后误报率~1%

反网络钓鱼方法	钓鱼检测	
	准确率	数据集
本方法	99.20%	10005 网页
Liu’s method	100%	8网页
CANTINA	90%	100 网页
Pan’s method	94.70%	279 网页
Xiang’s method	90.06%	7906 网页

云端假冒钓鱼检测引擎 - phish.anxinsao.com



二维码钓鱼 – 安全扫码器 – 请关注“安心扫” 公众号



网络身份安全的建议

- 前台匿名、后台实名
- 自动检测各种漏洞、攻击及威胁，强加密数据库
- 准备好数据泄露后的应对策略
- 使用网络身份安全措施提高个人及网站的网络身份安全
- 购买网络安全保险

谢 谢!

2018 ISC 互联网安全大会 中国 · 北京

Internet Security Conference 2018 Beijing · China

（原中国互联网安全大会）

感谢您的关注!

邮箱：liuwy@gdut.edu.cn

个人微信：csliuwy; 公众号：登录易

WIS Lab @ http://wislab.cn

登录易@ https://denglu1.cn