

从

起，浅议大数据时代的隐私安全

人

922人围观，发现 10 个不明物体

观点

近
在
行。

关注我们 分享每日精选文章

都会有新的城市在公交地铁领域开通移动支付。

支付企业纷纷抢滩城市公共服务。针对地铁公交这一场景，目前已经有超过30个城市，这意味着——用户可以丢开匿名的交通卡，使用支付宝及相关应用进行刷码通行。



和此前共享单车市场相比，轨道交通等公共出行在每个人的生活中都占据了更重要的地位。用“现代支付系统”取代传统地铁充值卡”所能提供的便利，让很多用户感到迫不及待。

而在现代支付系统的接入之后，出行的数据也可能继续用在城市交通管理之中，改善更多其他服务，如改进城市的



高速发展、高吞吐量的大数据时代，绝大多数人都的所有行为逐渐都能在信息世界之中留下数据足迹，无论是用户动留下的数字脚印，还是由他人建立的关于用户的数据。

在此，笔者仍然希望“老生常谈地”与大家探讨大数据时代的个人隐私问题。

个人信息与平台的信息不对称

信
数
并



条
能
的

工具和平台，或者接入平台的服务，就会产生相应的数据。在一些观点看来，人人更透明。但是作为线上服务的使用者，即便有用户隐私条款的存在，多数的用户仍信息，被追踪了哪些数据。

，都会给普通用户的精准理解带来一点困难。除此之外，用户条款中夹杂的“可继续模糊了好不容易建立起来的一点概念。这些数据的使用状况，数据分析和挖是不得而知。

关注我们 分享每日精选文章



即便一些用户是较为谨慎的——他们在选择应用服务的时候会考虑“数据泄漏”带来的影响，但他们能够避开的更多是一些声誉差的，已经发生过数据泄漏的企业。通常情况下，规模较大的企业能够享有更好的声誉和用户口碑，获多数人的信任和青睐。

如果这家服务商的大名家喻户晓，人人都在使用它的APP，更多的用户就能放心大胆的进行注册，然后开始使用这服务。但显然，这仍然只是“一厢情愿”的信任，大的企业也可能存在安全漏洞和疏忽。仅在过去的一年中，就继续出



漏：10亿账号的用户姓名、生日、邮箱地址、密码、电话、安全问题和答案全被

数据泄漏：超过 1.43 亿的美国用户数据泄漏，包括姓名，社保号码，出生日
； 20.9 万用户的信用卡号，以及部分英国和加拿大用户信息。

移动出行 Uber 数据泄漏：5700 万乘客用户信息泄漏，5 万名司机信息泄露。
关注我们 分享每日精选文章

深度分析和机器学习在混淆隐私边界吗

积累足够全面的数据才能让有规律的随机事件，在大量重复出现的条件下，呈现出几乎必然的统计特性。

随着政府的政策支持和公共部门积极进行数据整合，冗余数据中蕴含的战略和资产方面的价值逐渐浮出水面，企业纷纷着力于进行大数据技术研发与产品化。企业在应用数据进行决策和谋利的同时，仍然会不言而喻地使用用户的个人数据，使之在整合、关联和深度分析时受到隐私侵犯。在当前数据科学和信息技术发展的过程中，隐私信息的界定在变得不严格。





关注我们 分享每日精选文章

巨大的数据集之下，即便个人提供部分数据字段，也可能被得出一些隐秘的推测，并给用户带来风险。

2016年，上海交通大学的一篇机器学习论文《基于面部图像的自动犯罪概率推断》曾引起过争议。相同遭遇的还有2017年斯坦福大学公布的论文 [《Deep neural networks are more accurate than humans at detecting sexual orientation from facial images》](#)，其中的 AI 算法已经能够通过个人的肖像照片识别同性恋，且准确率高达81%。

公开在社交网站及其他角落的相片信息中的肖像不存在隐私问题，但对于人脸图像数据的分析、存储和使用，隐私问题依旧面临挑战。如果个人用户并不在意自己在各个角落的信息披露，认为网络空间仅仅是和普通公共场所一样那么当处在隐私边界时，他们就容易遭到认知偏差的影响，比如过度低估个人信息的利用程度和数据价值。只需要些公开的信息，运用技术就可以对个人的情况进行推断、分类甚至“定价”和“特殊对待”。





关注我们 分享每日精选文章

如今企业在技术层面的数据挖掘技术的演进、机器学习和深度学习算法的发展，都让数据利用效率和程度得到显著升。特定的分析流程和算法有时很难进行描述和解释，服务运营商甚至在进行分析之前也不了解他们能够得到的结果。更多隐私问题甚至是在二次开发利用原始数据时才引发的，因此在法律监管上也存在难度。

隐私专家 John Diebold 曾前写过这样一句话，而这在不久之后的将来，这可能会成为现实。

信息时代你留下的每一个字节，都会是构成隐私的血肉。

隐私意识在觉醒，但是...

2018年伊始时，大家的朋友圈都曾一度遭遇被“支付宝年度账单”和“网易云音乐”统治的时刻，随后当天就有用户揭被安插在年度账单首页之中、必须签署的《用户隐私协议》。





关注我们 分享每日精选文章

细心用户发现了这行隐蔽的小字，并点开看了《用户协议》，但也许是因为个人意志还无法抗衡从众心理，又或许是败给了自己的好奇心，最终绝大多数用户还是同意将自己的数据授权给了芝麻服务，开开心心地在朋友圈晒了支付宝年度截图。

可以发现人们的隐私意识虽然存在，但依然愿意用牺牲个人数据换取更好的网络生活，或者将安全和隐私的责任完寄托在政府和网络服务商身上。2014年时，EMC在15个国家和地区进行了15000 用户的隐私保护调查，而调查结果旧让人担忧。

50%的用户遭遇过或大或小的数据泄漏

62%的用户并不会定期修改密码

33%的人不会修改社交网络上的隐私设置

39%的人不会对自己的移动设备设置密码

而 2018 年 刚刚召开的Usenix Enigma 安全峰会上，谷歌工程师也在演讲时道出了谷歌用户的安全意识现状，可以看





关注我们 分享每日精选文章

安全、便利和隐私之争始终无法停歇

大数据时代的便利已经渗透进了我们生活的每个角落，更低的商品价格、更符合实时需求的产品和社会沟通度，要使用这些服务就会创建个人数据，而通过数据信息就总有办法辨析出个人的身份。

如果用户确实关心自己的隐私状况，不希望泄漏信息还是有一些简单的措施可以采纳：

1. 避免连接公共WiFi，及时在所有设备上安全更新
2. 开启账户两步验证功能，安装密码管理器
3. 注意所有应用程序的权限设置
4. 斟酌填写真实信息，适当选择备用方式



6. 适当使用匿名互联网服务

FreeBuf 隐私安全文章推荐阅读

[科技之殇：端到端加密究竟保护了谁？](#)[用来追踪网民](#)[P可在你不知情的情况下偷偷拍照窃取隐私](#)[借鉴了一个隐私保护功能](#)[R SGX 加强联系人搜索功能的隐私安全](#)[暗中收集Android位置数据](#)

关注我们 分享每日精选文章

[iTerm2 中可能通过 DNS 请求泄漏隐私信息](#)

*本文作者Elaine，转载请注明FreeBuf.COM

上一篇：[机器学习对抗案例：愚弄Google图像识别算法](#)下一篇：[本篇已是最新文章](#)

已有 10 条评论

匿名卡用户 2018-01-22

1楼 [回](#)

如果在一个失去信任的时代，实名反而不安全

亮了

[木千之](#) (4级) 人,既无虎狼之爪牙,亦无狮象之力量,却能擒狼缚虎,驯狮猎象,... 2018-01-222楼 [回](#)

深度学习应当是机器学习中的一种算法实现方式，个人以为新时代对于多域数据间关联分析的能力要远远超过以往任何一个时代，因而必要对机器学习应用的场景和使用的数据源进行限定。

亮了

[Elaine_z](#) (8级) 2018-01-22

@ 木千之 现在可以参考的就是欧盟GDPR和DPD吧，其中制定的保护公民隐私的细化方向，包括如算法的解释权，反对分析的权利，获得自己数据的权利，数据被遗忘的权利，然而具体怎么执行还要看各国的立法...如果企业都能如约执行，代价想必会是昂贵的

亮了


bi3g0 (1级) 2018-01-22

3楼 回

各位，从16年开始到现在断断续续接到过诈骗，推销的电话不止我一个吧？请问是xxx先生（女士）吗！看视频逛网站被弹出针对性广告推广的也不在少数吧？大学投简历，工作找房子，看病咨询，登录某站点等等生活里凡是需要实名电话记录的行为随后接到过多少骚扰电话？网络攻击带来的数据泄露只是冰山一角，现在不会消除以后也不会。但其实更多的侵犯到我们生活的数据泄露来自于企业（员工）私下里的贩卖和利用。

亮了

bi3g0



所以：
权？：

良： [

@Lsgo 美国的《隐私法案》，都对这些方面有提及，我国这方面落后很多。。。 [

关注我们 分享每日精选文章

4楼 回

所言我们被采集了什么隐私数据？这些数据被用来干嘛？ 2. 如果发现隐私泄露甚至知道泄露方后该怎么泄露是否有相应的应急措施？ 补偿机制？ 归根到底，谁要为实名制带来的风险买单？

亮了

masteruser 2018-01-22

5楼 回

法律和政策还需要制定，制定后的执行也要不知道多久的时间，数据泄漏啥时间发现，怎么罚，不知道中间会发生多少变化

如果大家自己都能提高隐私意识，还是最有效的

划重点，，，

- 1.避免连接公共WiFi，及时在所有设备上安全更新
2. 开启账户两步验证功能，安装密码管理器
3. 注意所有应用程序的权限设置
4. 斟酌填写真实信息，适当选择备用方式
5. 开启浏览器拒绝跟踪功能，及时清理 cookie
6. 适当使用匿名互联网服务

亮了

AI 2018-01-22

6楼 回

AI发展依赖这些数据，如果为了保护公民权利管控很严格，在和他国的竞争上又会落后。。。。。

目前能够在国际上获得各种奖项，也是因为我们的获取到了足够的数据，公共交通，医疗健康，电商消费。有数据才能继续研究

亮了

大数据时代，也许根本就没有隐私

亮了

轻狂书生2018-01-23

8楼回

[doge

亮了

选择

呢

请输入昵称
关注我们 分享每日精选文章

必须 您当前尚未登录。[登陆?](#) [注册](#)

邮箱

请输入邮箱地址

必须（保密）

表情

插图

提交评论(Ctrl+Enter)

[取消](#)



有人回复时邮件通知我



[Elaine_zz](#)

这家伙太懒，还未填写个人描述！

119

文章数

25

评论数

最近文章

- 从刷支付宝乘地铁谈起，浅议大数据时代的隐私安全

2018.01.22

- 机器学习对抗案例：愚弄Google图像识别算法

2018.01.16

相

支



浏览更多

[淘玉姐支付宝乘地铁让支付宝漏洞...](#)

关注我们 分享每日精选文章

[注意！Win10的11月更新重启数据收集...](#)

[浅谈互联网隐私安全](#)

[趋势科技发现支付宝安卓版漏洞](#)

特别推荐



[OpenVAS开源风险评估系统部署方案](#)

[双刃剑与灰色地带：“泄露数据收藏家”的素描](#)

[2015最酷的Hack方式有哪些？](#)

[【限时优惠】FreeBuf精品公开课 | 36W漏洞奖金先生CplusHua：](#)

[简单](#)

2016-01-05

[FB客服](#)

2017-09-16



关注我们 分享每日精选文章



Copyright © 2018 WWW.FREEBUF.COM All Rights Reserved [沪ICP备13033796号](#)

 阿里云 提供计算与安全服务