Information Security Engineering

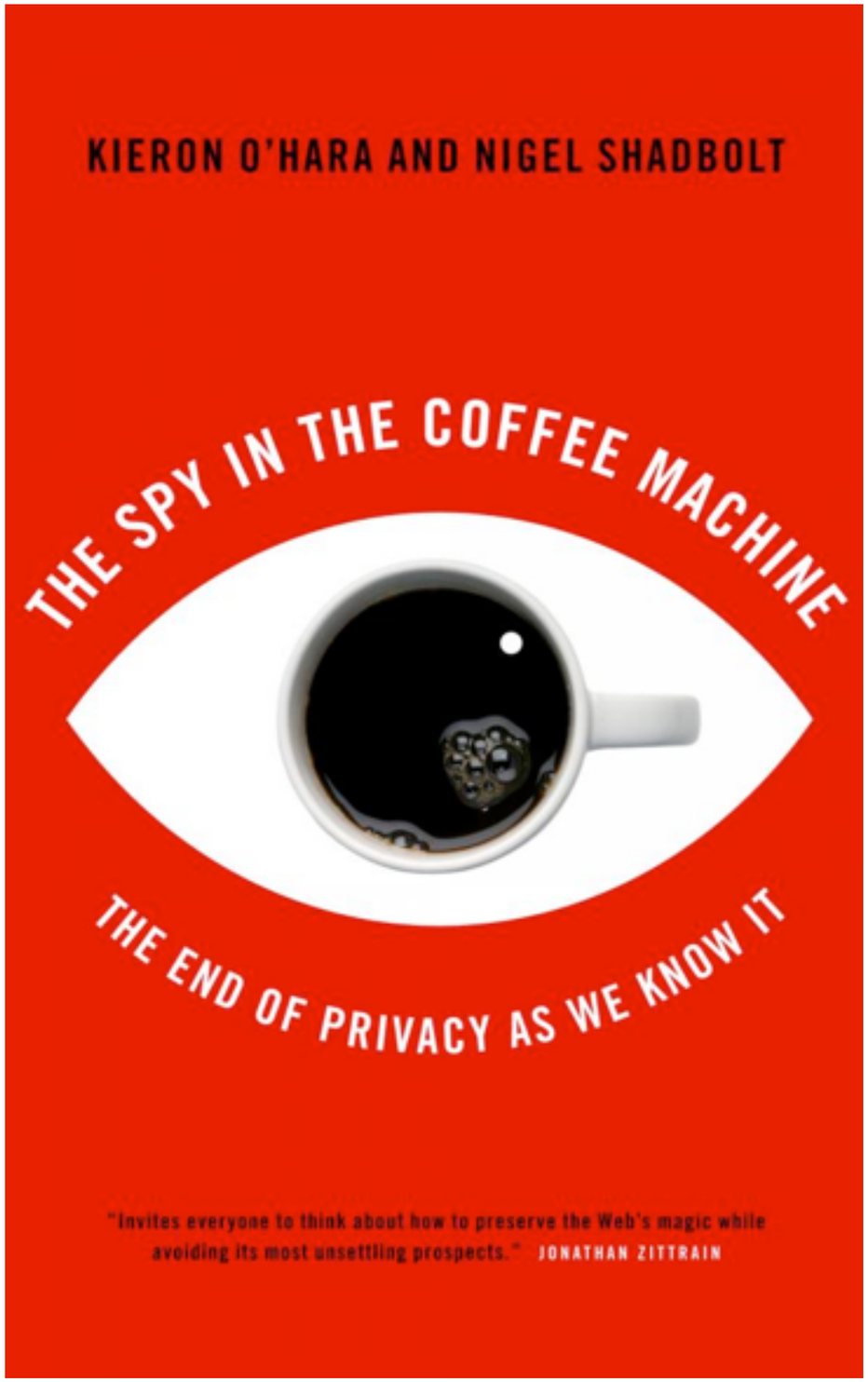
2017.12.06

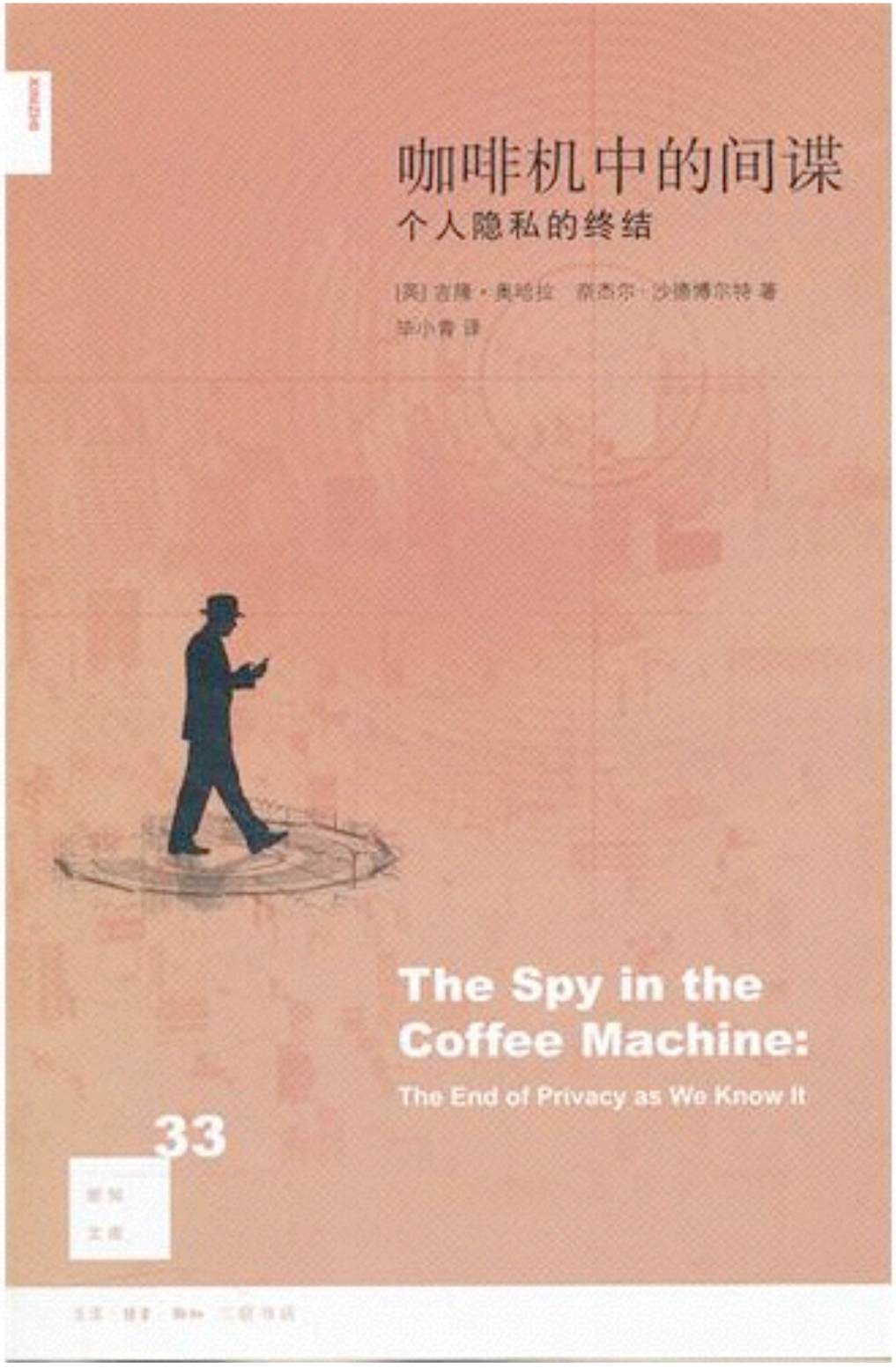
Privacy

北京大学 软件与微电子学院
School of Software and Microelectronics, Peking University

Huiping Sun(孙惠平)
sunhp@ss.pku.edu.cn

推荐一本书


KIERON O'HARA AND NIGEL SHADBOLT
THE SPY IN THE COFFEE MACHINE
THE END OF PRIVACY AS WE KNOW IT
"Invites everyone to think about how to preserve the Web's magic while avoiding its most unsettling prospects." JONATHAN ZITTRAIN


咖啡机中的间谍
个人隐私的终结
The Spy in the Coffee Machine:
The End of Privacy as We Know It
33

PETS

Proceedings on
Privacy Enhancing Technologies
Symposium

PoPETs Home | 2018 CFP | Info for Authors | About Reviews | FAQ | Proceedings | Acceptance Rates | Advisory Board | Code of Conduct

Caspar Bowden PET Award | Andreas Pfitzmann Best Student Paper Award | Best HotPETs Talks

2018 Symposium | Accepted Papers | Past PETs

PETS 2018

The 18th Privacy Enhancing Technologies Symposium
July 24–27, 2018
Barcelona, Spain

@PET_Symposium

https://petsymposium.org/

Privacy definitions are not one-size-fits-all

定义

- Wiki

  The right to be let alone

  ✴个人人格上的利益不受非法侵害，个人与大众无关的私事不能被发布公开，私人生活不能被非法侵入。

- 世界人权宣言

  ✴任何人的私生活、家庭、住宅和通信不得任意干涉，他的荣誉和名誉不得加以攻击，人人有权享受法律保护，以免受这种干涉和攻击

- 涉及一个人对有关自己信息的控制权

  ✴私人空间、私人决定、思想自由、私人财产

  ✴隐私意味着选择、自主、尊重、权利

来源

私人空间

公共空间

家庭
城邦

private：私
privare：剥夺

照相机
摄像机
摄像头
闭路电视监控

我让别人看我手机中的照片

我把手机中照片传到互联网上去

私人空间

社交空间

公共空间

隐私系统

信息收集

信息处理

聚合
鉴别
不安全
第二次使用

数据产生者

监控

数据拥有者

干扰决策

侵扰

入侵者

信息传播

侵害保密性
泄露
曝光
敲诈
扭曲

隐私泄漏渠道

- 数据经常被偷偷的收集

✴容易、廉价、自动化

- 数据被从多个数据源合并

✴可信与否

- 从商业目的收集的数据可能被用于其余目的

隐私面临挑战

- 公众使用信息的权利 vs 个人隐私
- 一种平衡
- 数据：公有 vs 私有
- 限制
- 神话：提供法律分类记录和数据
- 现实：需要一个框架、策略贺工具

  ✴是否需要

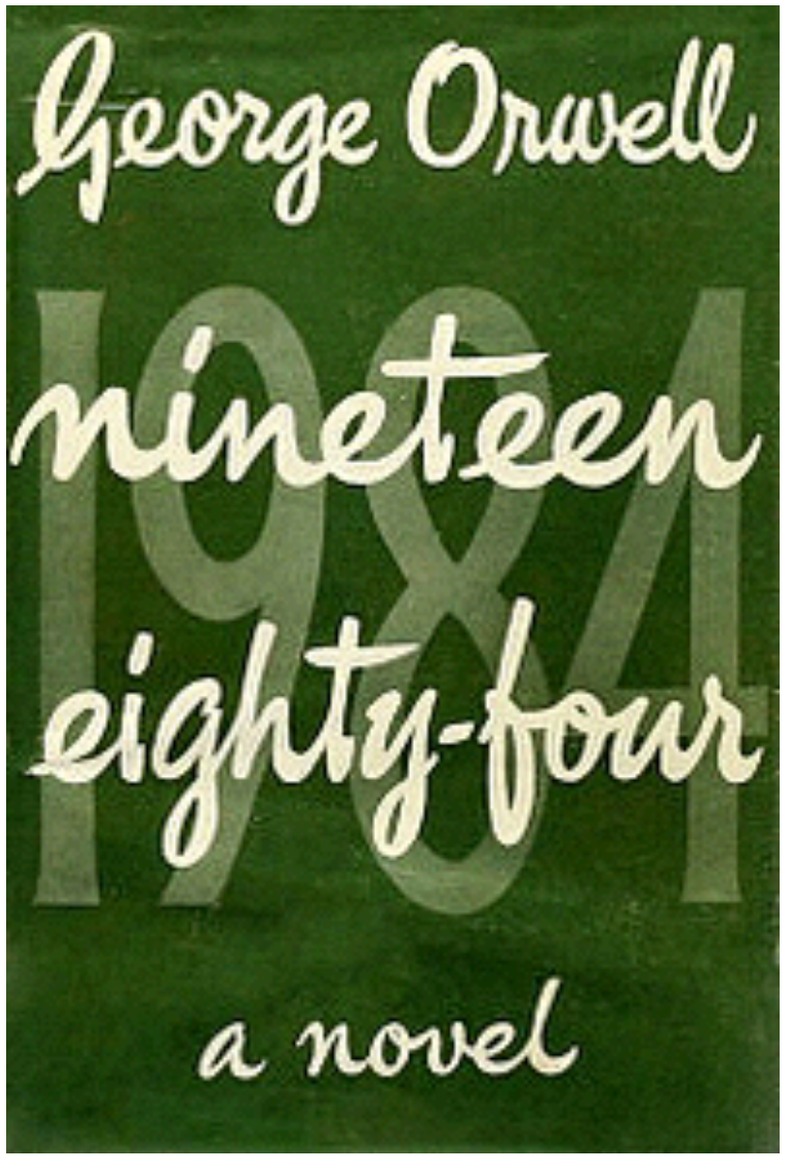
  ✴如何做

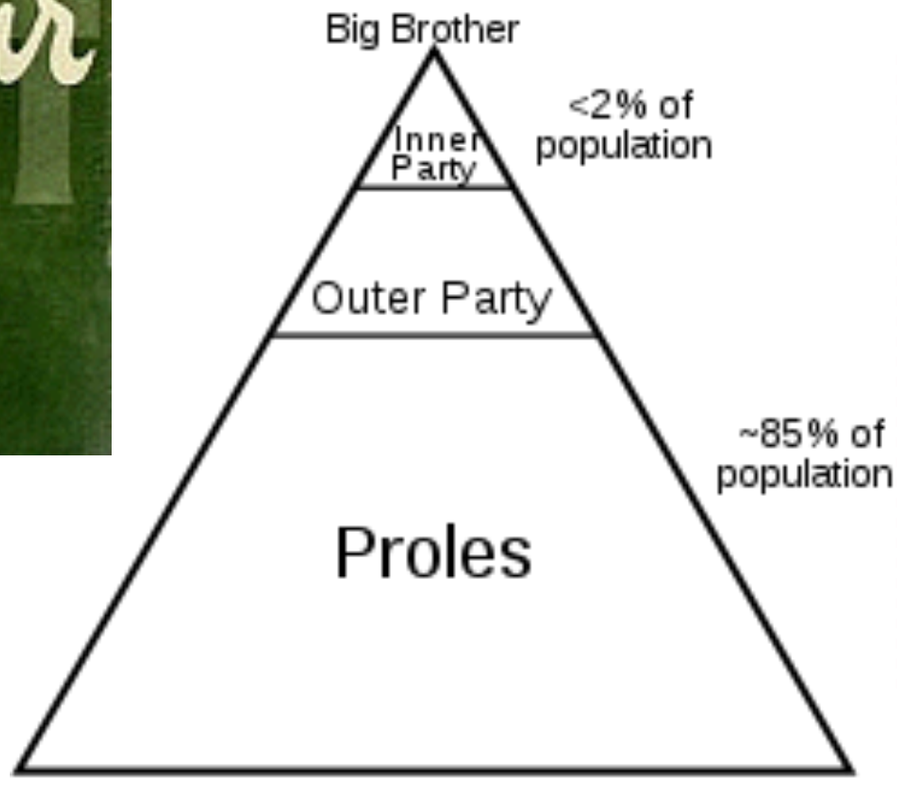
相关概念

- 匿名
- 伪名
- 不可连接
- 不可发现
- 位置隐私
- 监管受限

Why

Why Big Brother





斯诺登

http://en.wikipedia.org/wiki/Edward_Snowden

http://en.wikipedia.org/wiki/PRISM_(surveillance_program)

Google facebook skype AOL msn

YAHOO! Microsoft YouTube Hotmail

审查和监视

- 不可能对每一个人进行监视

  ✴确定性 vs 不确定性

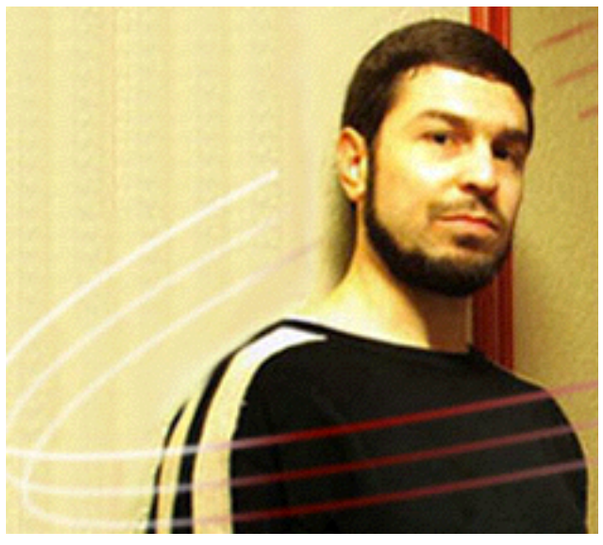
- 有目的的、常规的、系统的、有针对性的

  - 不管目的多么善良

    ✴信息可能被滥用

    ✴可能会犯错误

    ✴马希尔.阿拉尔



http://maherarar.net/

- 功能蠕变（Function Creep）

英国身份证

- 英国是普通法国家
  - ✴公民可以做任何没有被法律所禁止的事情
  - ✴没有法律禁止个人使用不同的身份和别名
  - ✴澳大利亚、新西兰、美国、加拿大
- 1915－1918和1939－1952两次建立了零时战时身份证制度
- 民众支持，2006年，50%支持，反对39%
- 身份登记数据库将不可避免的出现数据不准确和管理错误的问题
- 身份盗用问题

大數据

隐私的灾难

主要隐私威胁

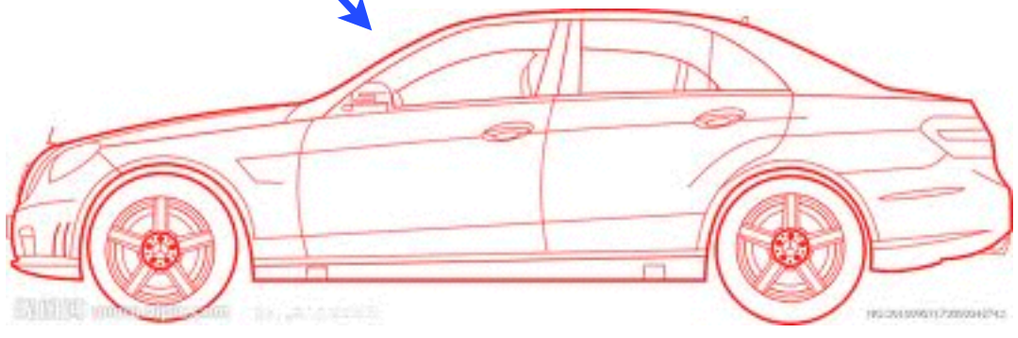
- Big Brother
  - ✴海量数据＋智能搜索
  - ✴上网记录、通话记录、位置信息、摄像、卫星
- Little Sister
  - ✴企业、银行、ISP、IDM
  - ✴隐私信息恶意泄露
- Big Mess
  - ✴垃圾短信、垃圾邮件、电话营销

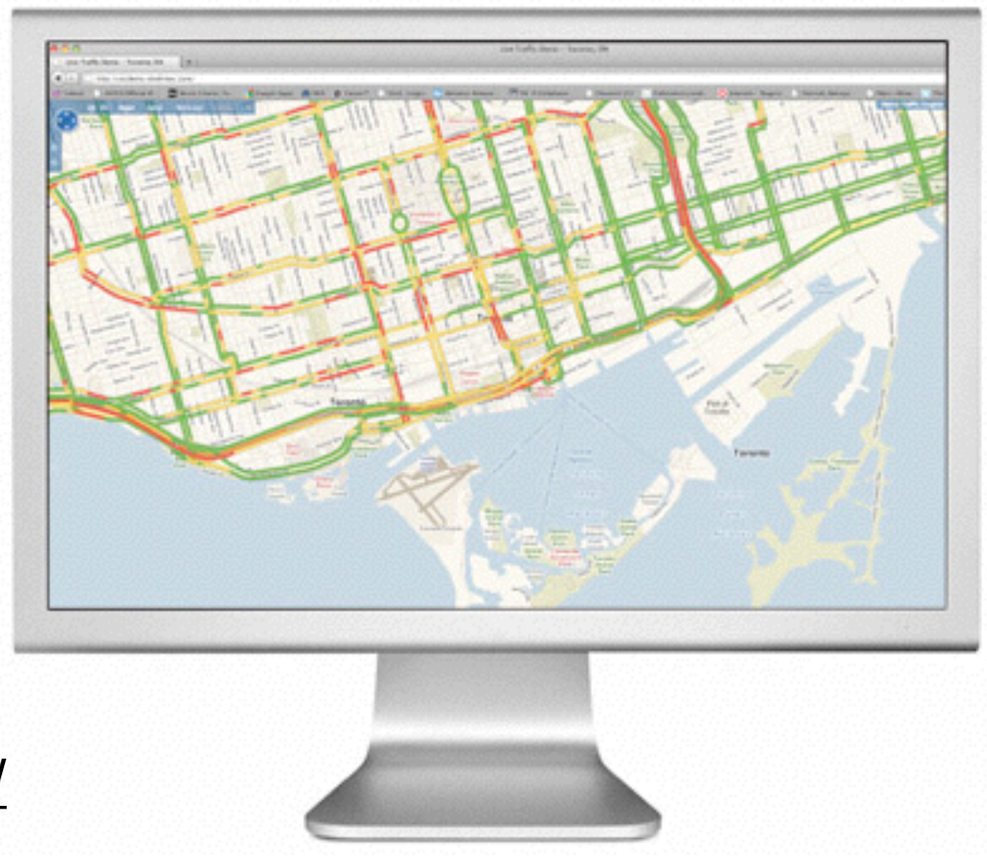
Why

智能家居

Why

智能交通



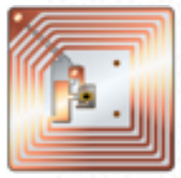


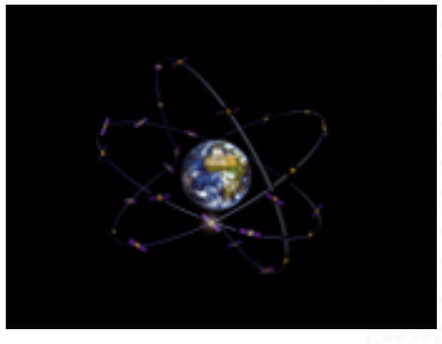
http://www.intellimec.com/m2m-telematics/traffic/

智能海关

- 速度快、安全、方便、有效
- 隐私保护?
  - ✴ 两个职员
  - ✴ 没有面部
  - ✴ 不存储数据
  - ✴ 不允许记录设备进入
  - ✴ 自动删除

其余





- 手机
- 笔记本
- 信用卡
- 银行卡
- 车辆
- 摄像头
- 工卡
- 蓝牙
- ……

Why

SNS隐私

- 2000年

  ✴87%的美国居民可以通过如下几个简单的信息来唯一确定：性别、ZIP编码、生日

- 2005年，Facebook profile

  ✴90.8%包括图片、87.8%包括生日、39.9%包括电话号码、50.8%包括居住地点

  ✴政治观点、约会喜好、婚姻状态、兴趣等

- 2005年

  ✴生日＋家乡＋居住地，可以用来估计用户的SSN

- 可能存在恶意的第三方应用

  ✴2008年，BBC News开发了一个

- 应用开发者可能违背开发策略

  ✴2010年，The Wall Street Journal发现有应用发送用户信息给广告和跟踪公司

- 第三方应用可能查询更多的应用根本不需要的数据

  ✴2008年，前150个应用有91%超需要收集数据

去匿名化攻击

- 大量的SNS数据被收集
  - ✴2010年，Twitter，4亿profile、14亿关系、10亿tweets
- 匿名也不安全
  - ✴使用外部背景知识可以去匿名
  - ✴标识、属性、关系、图特征
  - ✴2006年，AOL公布65万用户的2千万搜索关键词，纽约时报使用电话本标示了一些人
  - ✴主动攻击 vs 被动攻击
  - ✴边调整机制、基于聚类的泛化
- 推理攻击：使用用户公开的信息来推理其余隐私信息

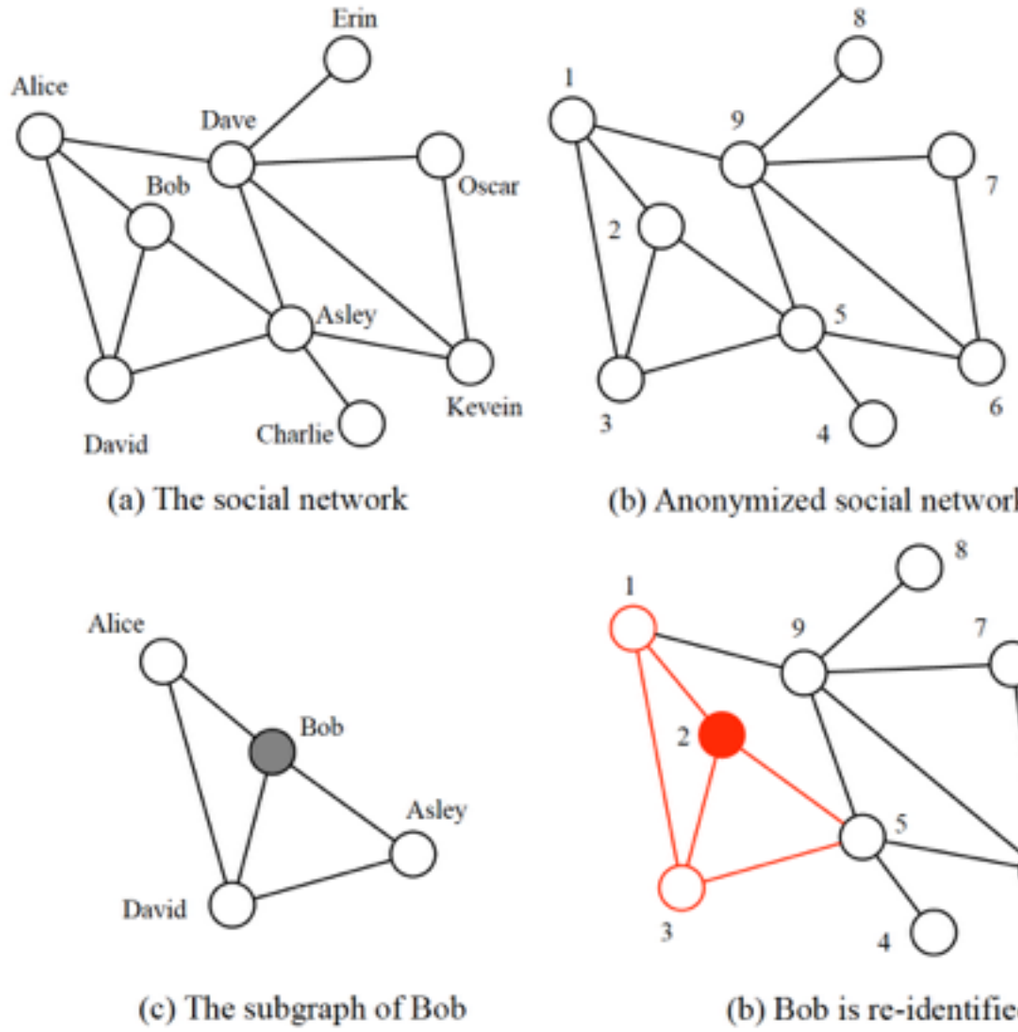


Fig. 5: Anonymization and de-anonmymization attacks.

How

欧盟隐私数据保护法律

- Reporting the processing
  - ✳Data Protection Authority
  - ✳收集目的
- Transparent Processing
  - ✳用户要知道收集目的，要知道谁来处理，进行哪些处理
- “As required” Processing
  - ✳不能进行和目的不一致的处理
- Lawful basis for data processing
  - ✳要依照法律

欧盟隐私数据保护法律

- Right of the parties involved

  ✴修改更新

- Processing personal data by a processor

  ✴要符合限制，有义务保证

- protection again loss and unlawful processing of personal data
- Data traffic with countries outside the EU

  ✴要依照法律

- Data quality

  ✴准确、足够、合时、最小化

个人信息泄漏披露

- 个人信息
  - ✴ 名字、地址、Email、电话、SSN、信用卡、银行账户、...
- 所有服务均需要填写大量信息
  - ✴ 哪些信息是必须的
- 如何保证这些信息不被滥用

- 公司主要为了减少风险
  - ✴ 支付欺诈
  - ✴ 交易是否合法
- 过量收集个人信息
- 但不保护个人信息

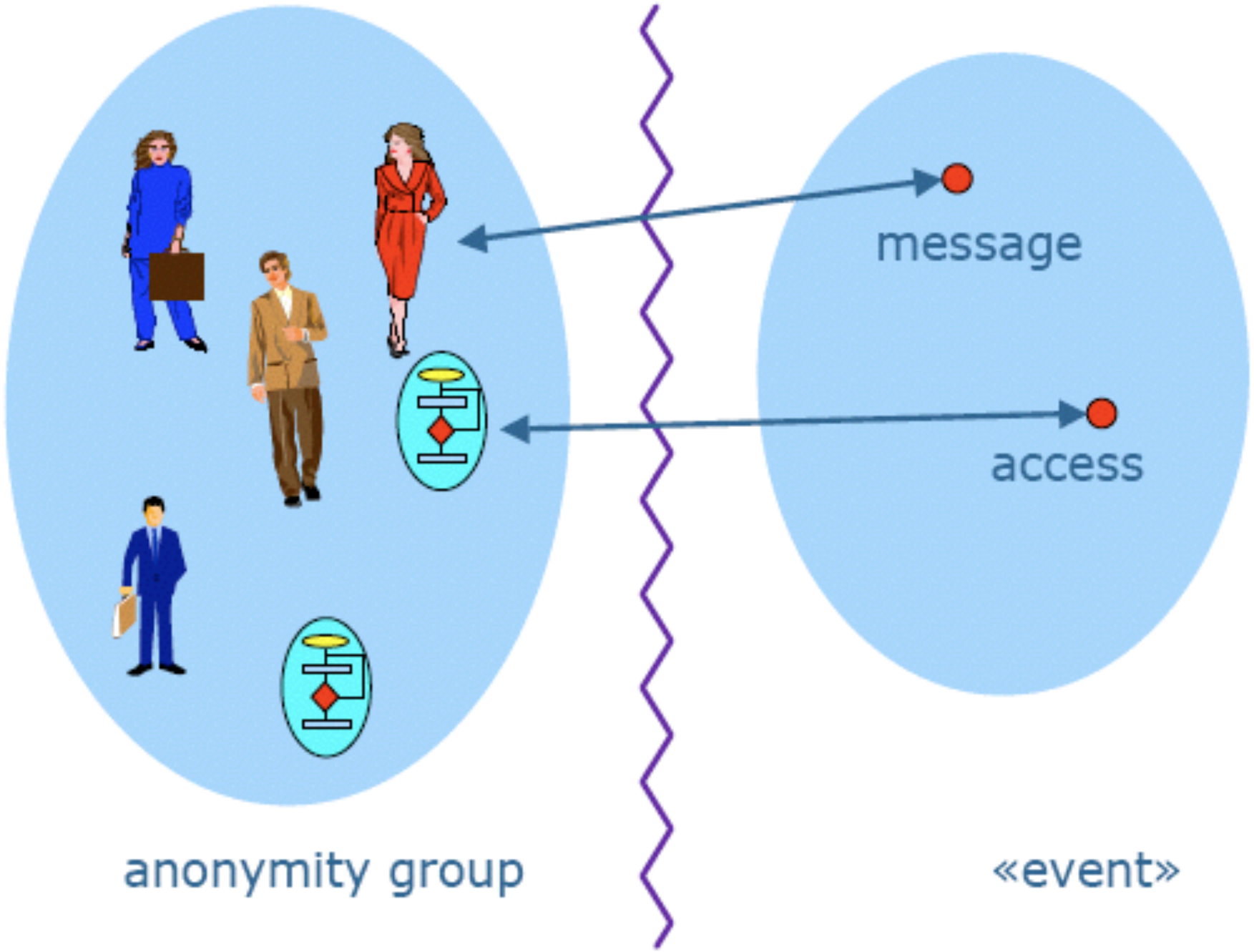
How

个人信息泄漏披露

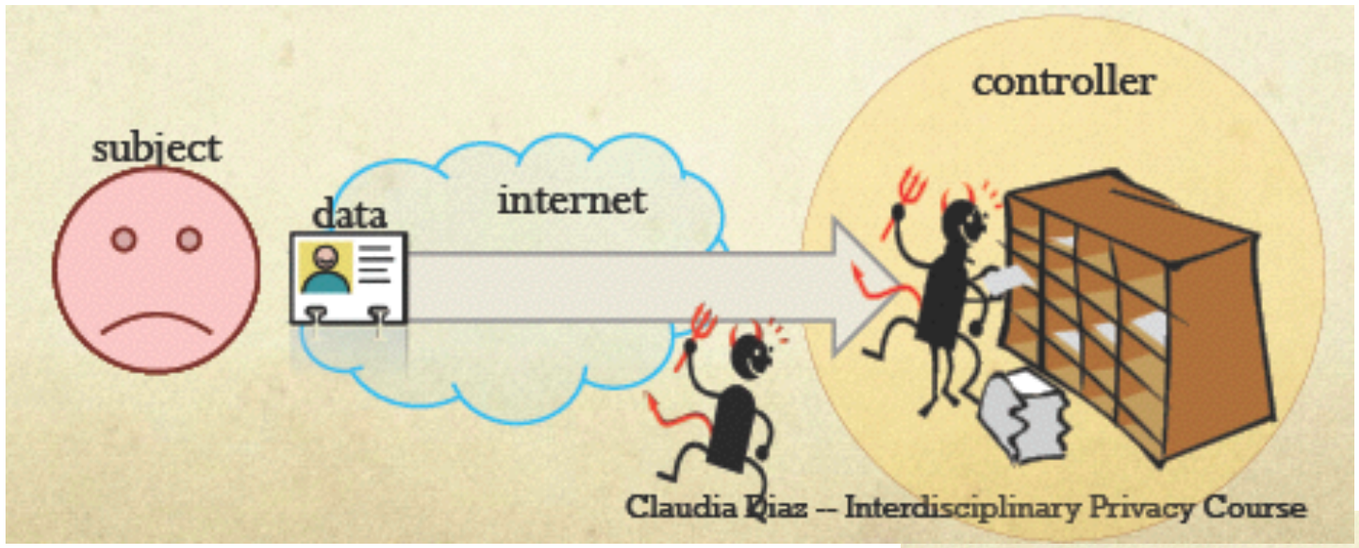
- 美国的44个州已经建立法律，要求涉及个人信息泄露的时间必须披露
- 要求掌握个人信息的公司必须保证安全
- 目的
  - ✴ sunlight is the best disinfectant
  - ✴ right to know
- Toxic Release Inventory (1986)
- 高的公司花费 vs 低的社会花费

- 个人信息保护
  - ✴ 事先防范
  - ✴ 事后惩罚
  - ✴ 信息披露
- 分析
  - ✴ 激励公司投资于个人信息保护，减少类似事件发生
  - ✴ 帮助消费者自我保护
  - ✴ 公司并不补偿消费者，也不需要强制要求事先的安全机制

隐私保护的目的

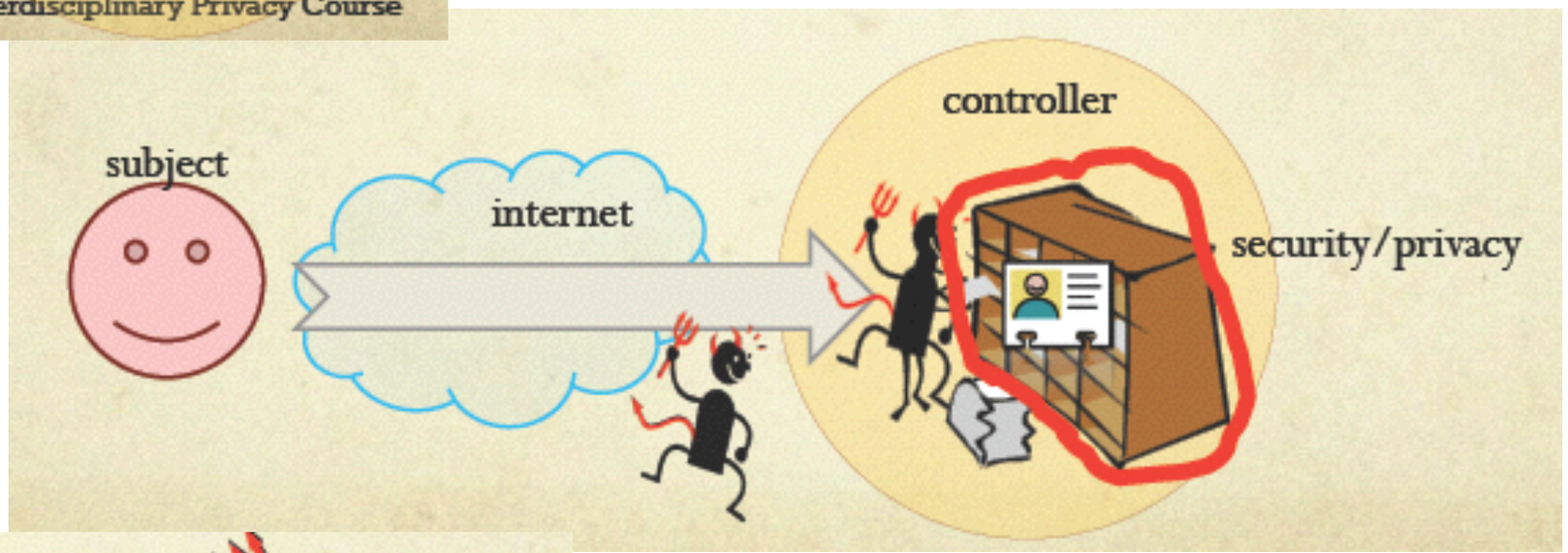


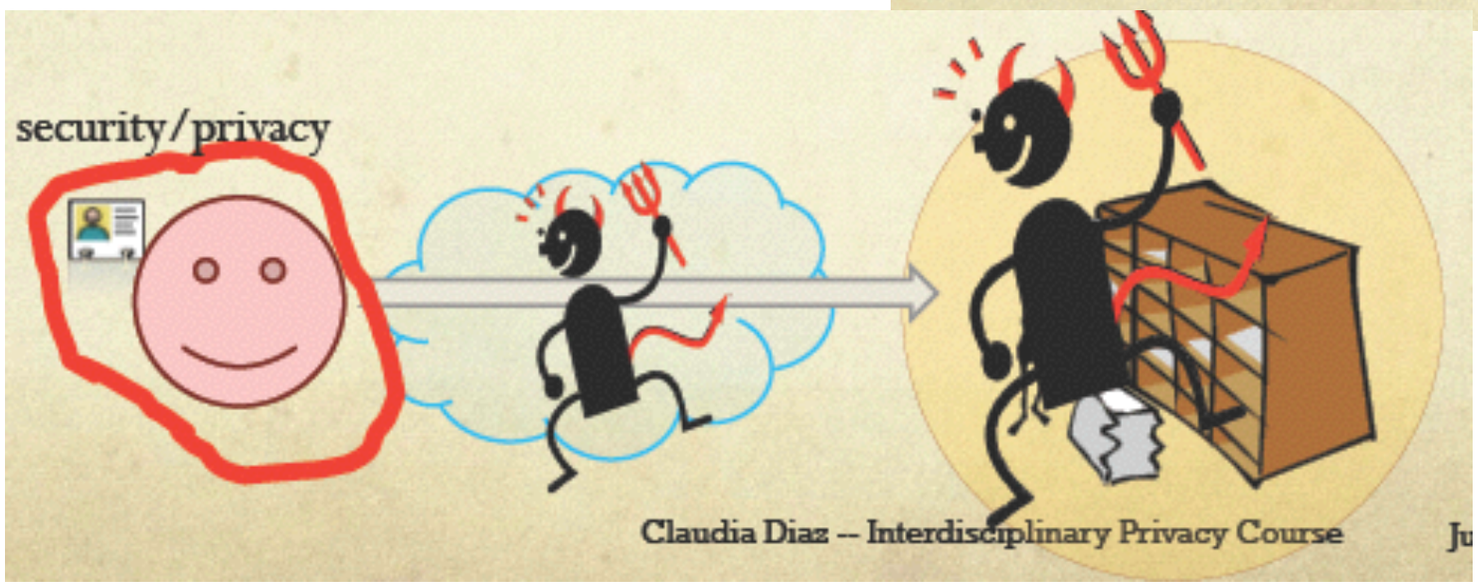
三种假设



信任数据
拥有者

信任数据拥有者
的安全体系





信任自己的
安全体系

如何做

- 身份认证
- 访问控制
- 加密
- 隐写
- 混淆
- 填充

PET

- 隐私丰富技术（PET： Privacy Enhancing Technology）
- 是一种用于隐私保护的信息技术和系统
  - ✳ 减少和最小化对个人数据的访问
  - ✳ 防止对于个人数据不必要和不希望的处理

- 隐私保护
  - ✳ 伪名工具
  - ✳ 匿名服务（Email、Web、Key、…）
  - ✳ 加密工具（Email、Document、…）
  - ✳ 过滤工具（Email、Web、Pop-up、…）
  - ✳ 痕迹擦除工具（Data、Activity、Browserß）

- 隐私管理
  - ✳ 策略创建和核对
  - ✳ 管理工具
    - ➡ 身份管理
    - ➡ 取证共举办
    - ➡ 审计工具

P3P

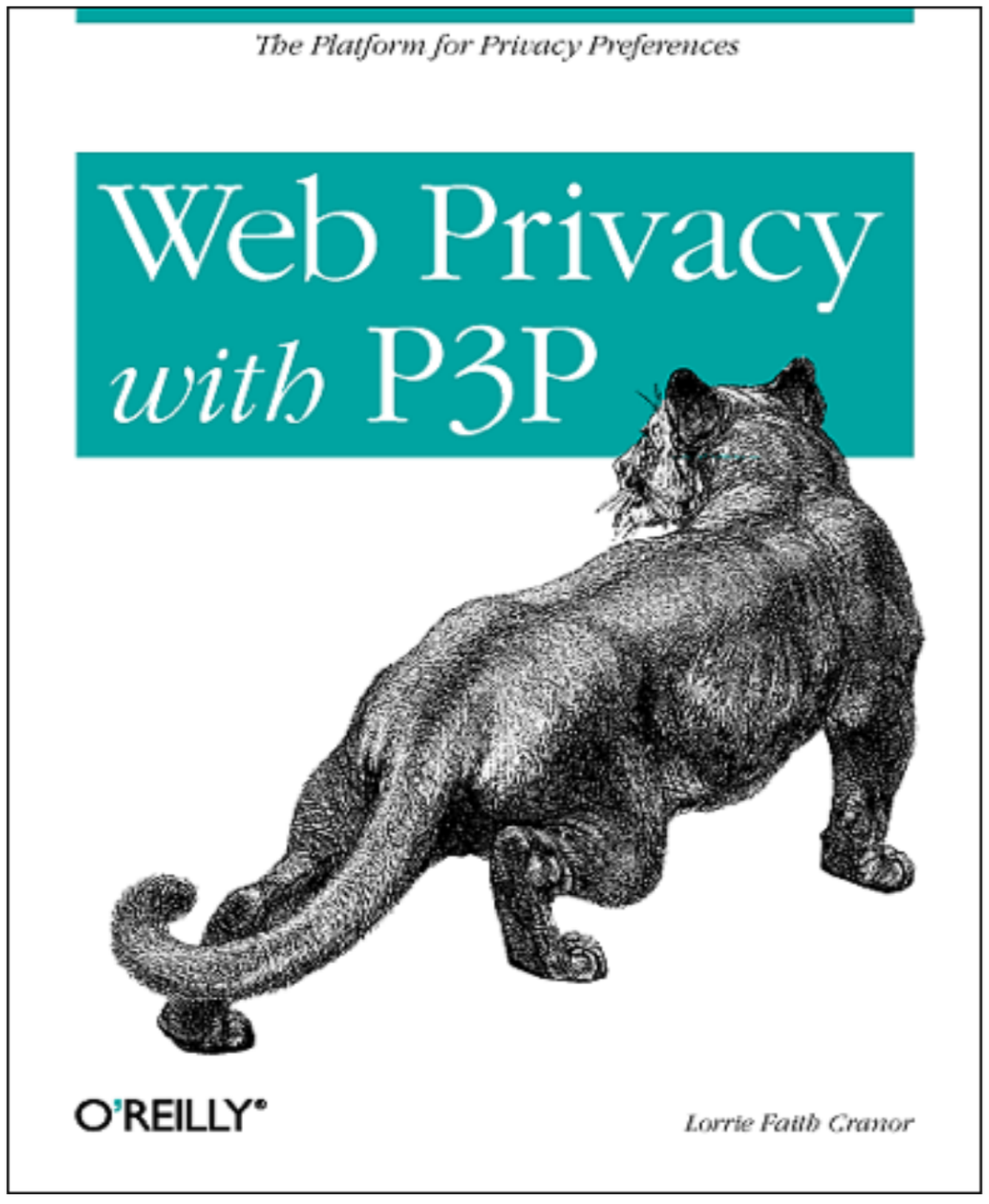
Platform for
Privacy
Preferences
Projects

隐私
喜好
项目
平台

一种隐私策略语言
一种隐私信息描述语言

不能解决隐私问题

http://www.w3.org/p3p
http://p3pbook.com

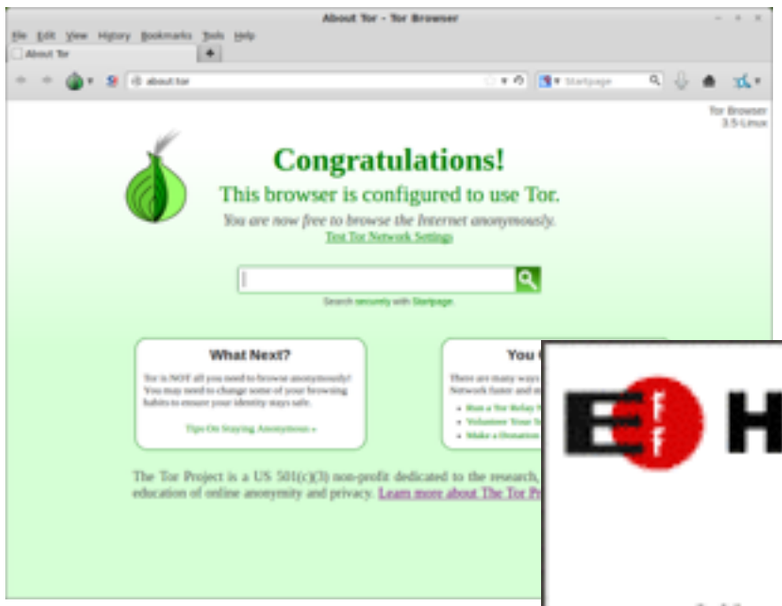


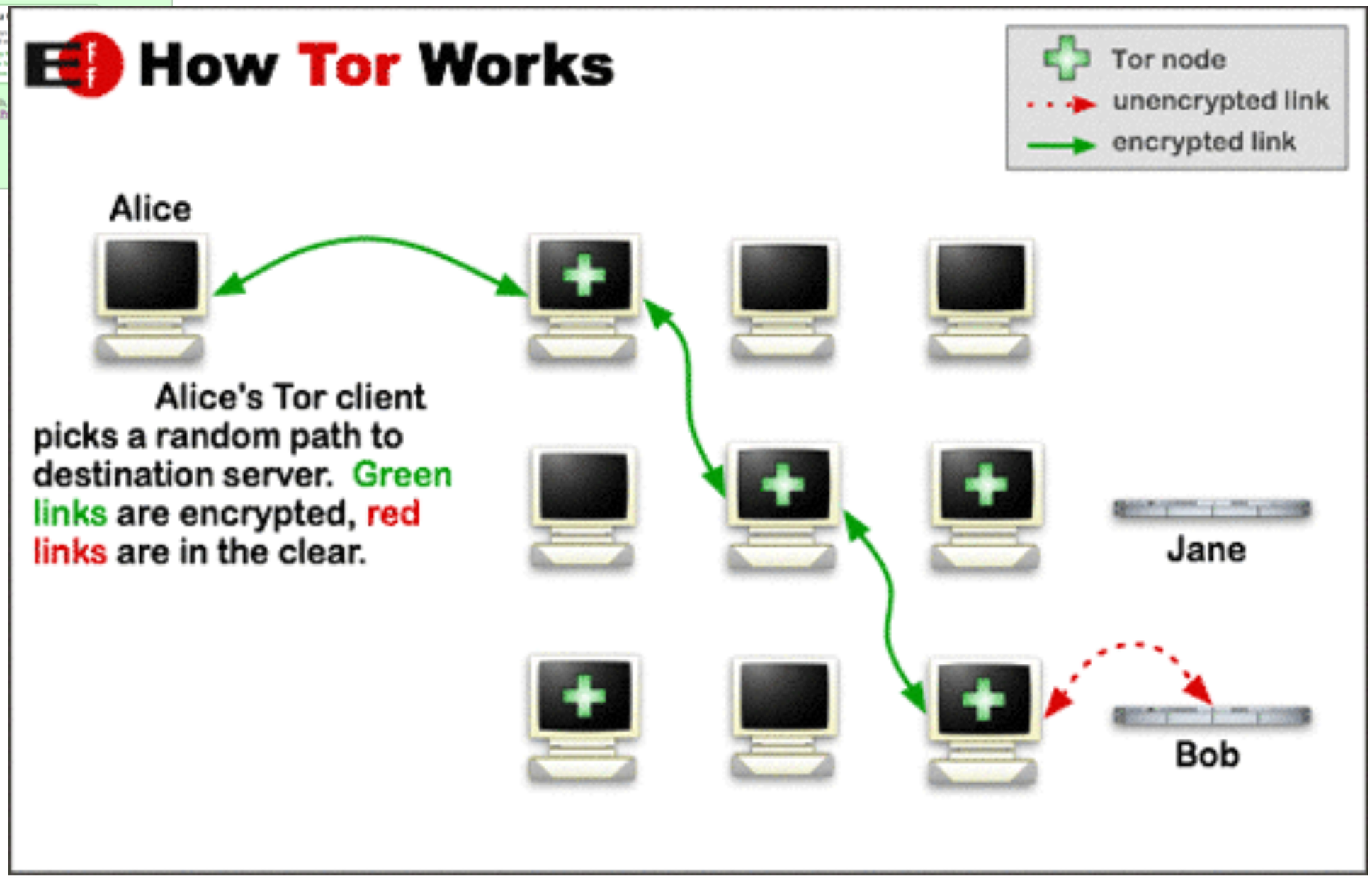
How

P3P

- 谁收集数据?
- 什么数据被收集?
- 为了什么目的使用这些数据?
- 谁是数据的接收者？是否给第三方?
- 数据保留策略？存储多久？如何存?
- 策略冲突如何解决?
- 人类可读的隐私策略在哪里?
- 用户是否可以查看这些数据?

How Tor


About Tor - Tor Browser
Congratulations!
This browser is configured to use Tor.
You are now free to browse the Internet anonymously.
Test Tor Network Settings
What Next?


How Tor Works
Tor node
unencrypted link
encrypted link
Alice
Alice's Tor client picks a random path to destination server. Green links are encrypted, red links are in the clear.
Jane
Bob

Tor使用情况

The anonymous Internet

Daily Tor users per 100,000 Internet users

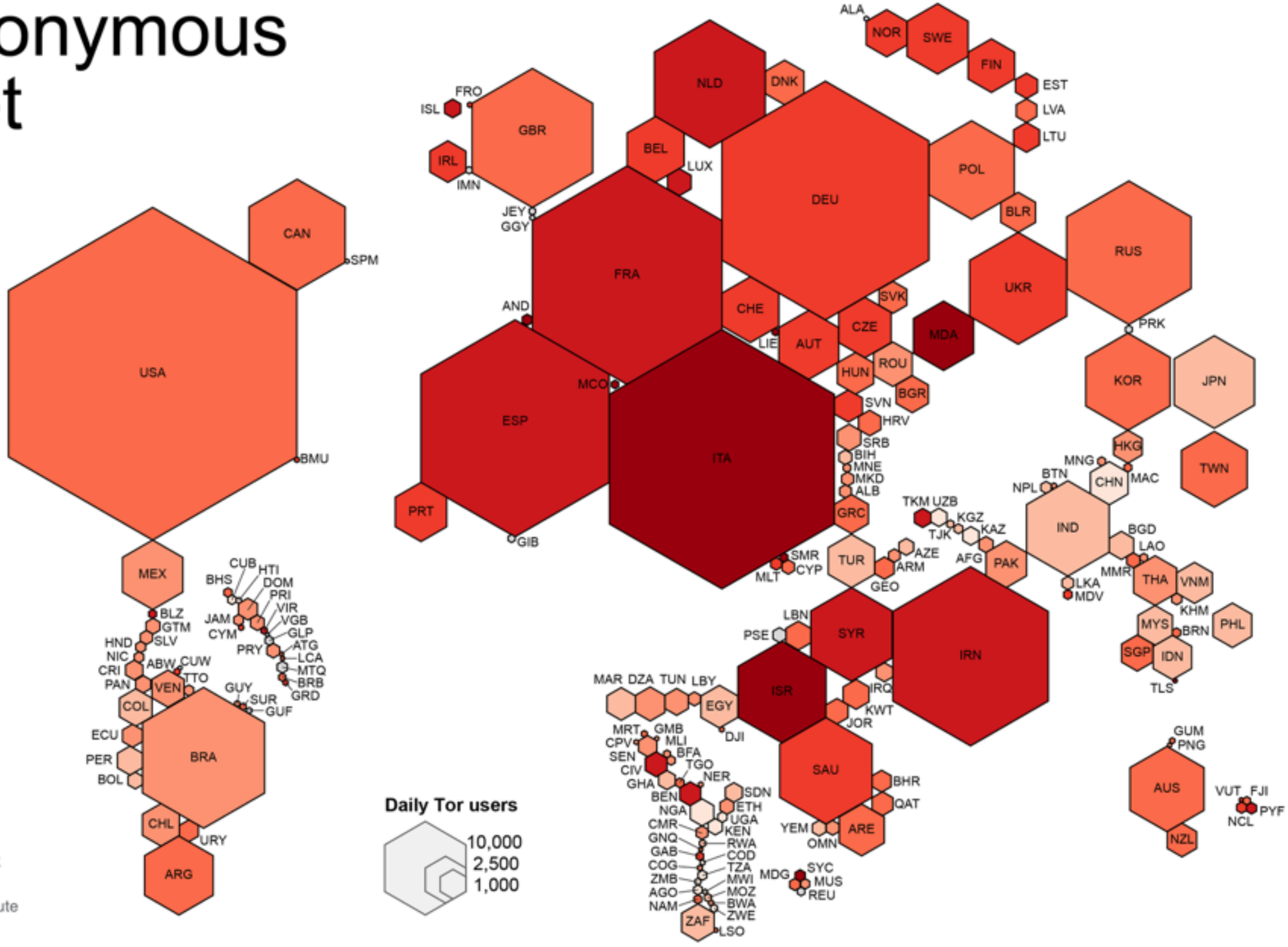
- > 200
- 100 - 200
- 50 - 100
- 25 - 50
- 10 - 25
- 5 - 10
- < 5
- no information

Average number of Tor users per day calculated between August 2012 and July 2013

data sources:
Tor Metrics Portal
metrics.torproject.org
World Bank
data.worldbank.org

by Mark Graham (@geoplace) and Stefano De Sabbata (@maps4thought)
Internet Geographies at the Oxford Internet Institute
2014 • geography.oii.ox.ac.uk

Oxford Internet Institute
University of Oxford



How

K匿名

Name	Age	Gender	State of domicile	Religion	Disease
Ramsha	29	Female	Tamil Nadu	Hindu	Cancer
Yadu	24	Female	Kerala	Hindu	Viral infection
Salima	28	Female	Tamil Nadu	Muslim	TB
sunny	27	Male	Karnataka	Parsi	No illness
Joan	24	Female	Kerala	Christian	Heart-related
Bahuksana	23	Male			
Rambha	19	Male			
Kishor	29	Male			
Johnson	17	Male			
John	19	Male			

Name	Age	Gender	State of domicile	Religion	Disease
*	20 < Age ≤ 30	Female	Tamil Nadu	*	Cancer
*	20 < Age ≤ 30	Female	Kerala	*	Viral infection
*	20 < Age ≤ 30	Female	Tamil Nadu	*	TB
*	20 < Age ≤ 30	Male	Karnataka	*	No illness
*	20 < Age ≤ 30	Female	Kerala	*	Heart-related
*	20 < Age ≤ 30	Male	Karnataka	*	TB
*	Age ≤ 20	Male	Kerala	*	Cancer
*	20 < Age ≤ 30	Male	Karnataka	*	Heart-related
*	Age ≤ 20	Male	Kerala	*	Heart-related
*	Age ≤ 20	Male	Kerala	*	Viral infection

RFID隐私

RFID应用

- 个人
  - ✴ 药品、钞票、行踪、衣物、...
- 公司
  - ✴ 资产
- 国家
  - ✴ 关键物品和行动

Killing、Sleeping和Renaming

- 使用PIN来保护Killing和Sleeping命令
- 32bit，EPC Class2 Gen2标准
- PIN管理是一个难题
- 物理触发

- RFID＝ProductTypeID＋UniqueID
- RFID＝RealID or TempID
- 两个Tags，RFID1、RFID2
- 随机ID
- 物理限制数据发送和接收

Re-encryption和Proxy

- 密钥

- Faraday Cage
- Watchdog Tag
- RFID Guardian
- RFID Enhanced Proxy

物理机制

RFID Inactive

RFID Active

- 基于能从实际环境和应用中得到的物理因素来设计RFID隐私保护算法
- 成功故事
  - ✴ 非接触卡－临近暗示可信
  - ✴ 物理阻塞
  - ✴ 物理触发的护照

为什么使用物理机制

基于标准密码学

非对称加密
对称加密

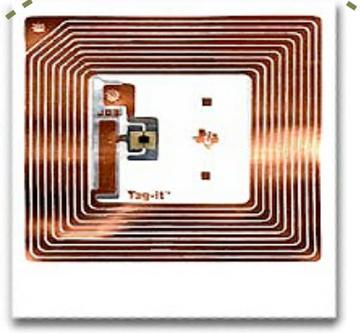
低成本Tag

有限存储
有限计算

基于轻量级密码学

Hash　　Pseudonyms　　HB

更容易被攻破

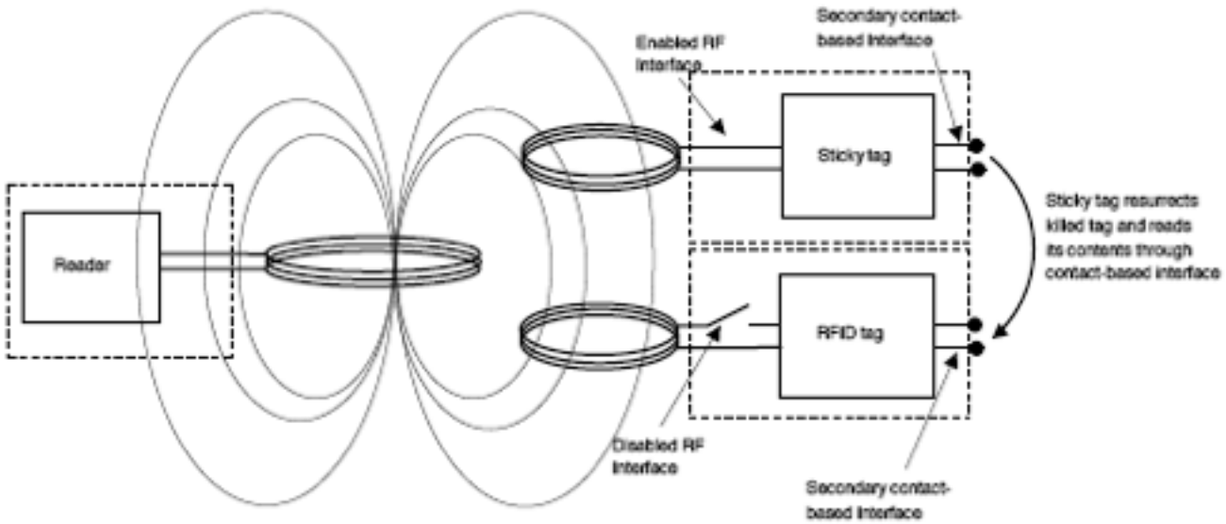


Tag

Reader

物理机制例子

- 增加Sensor
  - ✳ 湿度感应、衣服+洗衣机
  - ✳ 光感应、控制On/Off（直接、遮盖）
- 距离
  - ✳ 辅助Kill，近距可以唤起
  - ✳ 2个天线，2个接口



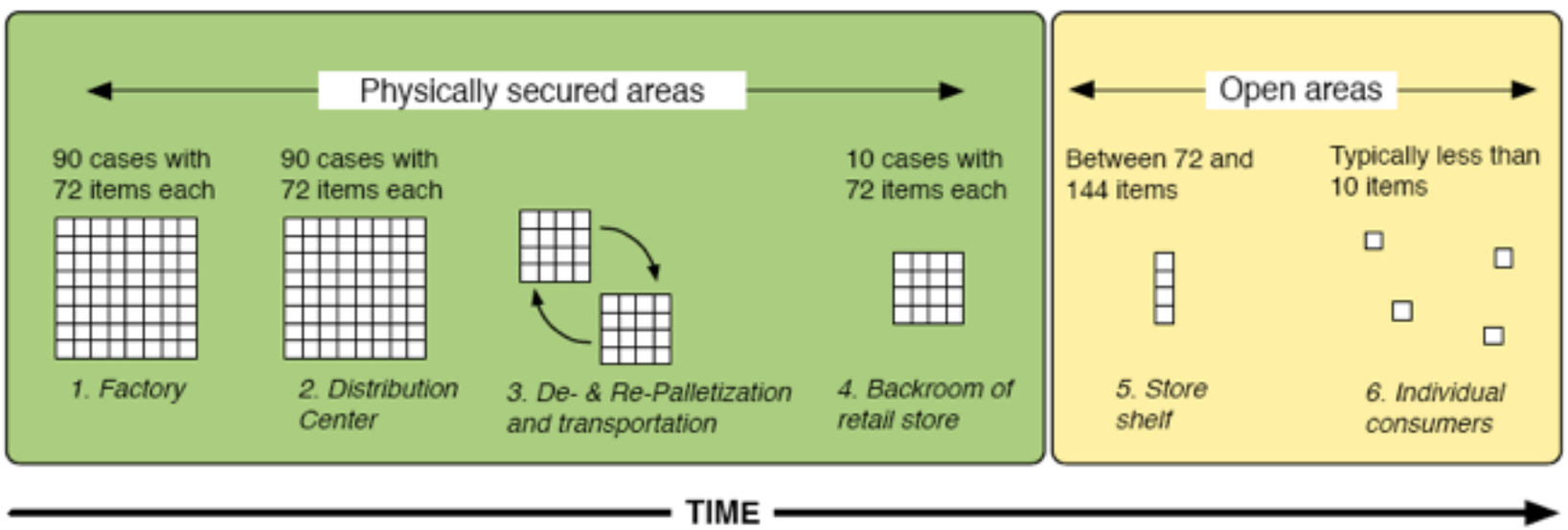


Figure 1: **Object hierarchies in RFID-enabled supply chains** *This schematic represents the path taken by an individual pack of razor blades from the factory to the consumer's home. Please refer to Section 2.2 for details.*

Unidirectional Key Distribution Across Time and Space with Applications to RFID Security，USENIX Security Symposium 2008

Homework 课后作业

- 要求阅读如下论文：

➡ *Designing Statistical Privacy for Your Data. In CACM 2015.*

下次上课测试！

谢谢！

孙惠平

sunhp@ss.pku.edu.cn