Information processing

- manual processing (10^2)
- mechanical processing (10^5)
- PCs and LANs (10^8)
- mainframe (10^13)
- the Internet of things, ubiquitous computing, pervasive computing, ambient intelligence (10^12)
- Internet and mobile (10^9)

Information storage and transmission

- 2010: digital universe is 1.2 Zettabyte; this corresponds to 600 million hard drives with a capacity of 2 Terabyte (2020: 80 Zettabyte)
- 2014: global internet traffic will grow to 64 Exabyte/month (2009: 15 Exabyte)

Exponential growth

- Ray Kurzweil, KurzweilAI.net
- human brain: 10^14 ... 10^18 ops and 10^13 bits memory
- 2025: 1 computer can perform 10^16 ops (2^23)
- 2013: 10^13 RAM bits (1 Terabyte) cost 1000$

Uniqueness

- physics and electronics (accidental)
  - process variation in deep submicron processes
  - radio fingerprinting: unique pattern of each wireless antenna, modulator, filter, oscillator
  - fibers in paper
  - magnetic behavior of certain materials
- human: biometry
  - fingerprint
  - iris
  - DNA
  - face
  - gait
  - ...
Uniqueness

physics and electronics (deliberate)
- MAC address, IMEI
- Pentium III Processor Serial Number 1999
- yellow dots produced by laser printers
- PUF Physical Unclonable Function

"Chattering" devices

RFID
Bluetooth/Zigbee
WLAN
WiMAX
2G/GSM
3GSM
GPS/Glonass/Galileo

Location Based Services

location-based traffic monitoring and emergency services
- e-Call, traffic congestion control
location finder:
- where is the nearest restaurant, gas station,...
- variable pricing applications
- congestion pricing
- pay-as-you-drive
- social applications
- Geotagged Twitter
- Google Latitude

Why is this a problem?

- do you want to be seen at certain locations?
  - abortion clinic, AIDS clinic, business competitor, or political headquarters (Google Street View)

- what can be automatically inferred about a person based on location?
  - any important location...
  - desk in a building
  - home location
  - future locations
  - and even identification!
  - http://www.batchgeocode.com/lookup/

Intelligent processing

uniqueness + connectivity + processing power

- create "big brother" or "Kafka" for specific purposes
  - protecting children
  - road pricing and congestion control
  - public transport
  - car insurance
  - car pool
  - social networking
  - anti-counterfeit
  - copyright infringements
  - ...

- individual applications are legitimate
- cost effective
  - limited need for tamper resistance: cost reduction
  - allows for effective pricing (and price discrimination)
  - long term incentive for integrating solutions and function creep
Uniqueness, Identity and Privacy
Bart Preneel

Scott McNealy, co-founder, Sun Microsystems (1999)
- “You have zero privacy anyway. Get over it”

Eric Schmidt, CEO, Google (2009)
- 10/12/09: When asked during an interview about whether users should be sharing information with Google as if it were a "trusted friend": "If you have something that you don’t want anyone to know, maybe you shouldn’t be doing it in the first place.”

Mark Zuckerberg, CEO, Facebook (2010)
- 09/01/10: The age of privacy is over
  - “People have really gotten comfortable not only sharing more information and different kinds, but more openly and with more people. That social norm is just something that has evolved over time.”
- Back in 2008: “Privacy control is the vector around which Facebook operates.”

US Department of Commerce
- 16/12/10: “Commercial Data privacy and Innovation in the Internet Economy: A Dynamic Policy Framework”
  - the adoption of Fair Information Practices (FIPs)
  - the development of privacy codes of conduct
  - the creation of a privacy office in the Department of Commerce

The privacy debate
- “if you care so much about your privacy it’s because you have something to hide”
- “surveillance is good and privacy is bad for national security. We need a tradeoff between privacy and security”
- “people don’t care about privacy”

Solove:
- “the problem with the ‘nothing to hide’ argument is its underlying assumption that privacy is about hiding bad things.”
The privacy debate

"surveillance is good and privacy is bad for national security. We need a tradeoff between privacy and security"

"we need more surveillance" is a powerful argument
- if attacks increase, you can argue that you need even more
- if attacks decrease, you take credit

The privacy debate

“people don’t care about privacy”

people want to control information:
- impression management /self-presentation
- what do we tell to whom
- concerns over information taken out of context
- personal safety
- we value friends who are discreet

The privacy debate

[Solove] “Part of what makes a society a good place in which to live is the extent to which it allows people freedom from the intrusiveness of others. A society without privacy protection would be suffocation.”

[Diffie and Landau] “Communication is fundamental to our species; private communication is fundamental to both our national security and our democracy.”

[Diffie] “In the long run privacy and individual autonomy have no chance against increase in communications.”

Taking privacy to create security

Privacy = Security Property

- individuals
  - freedom from intrusion, profiling and manipulation, protection against crime / identity theft, flexibility to access and use content and services, control over one’s information
- companies
  - protection of trade secrets, business strategy, internal operations, access to patents
- governments / military
  - protection of national secrets, confidentiality of law enforcement investigations, diplomatic activities, political negotiations
- shared infrastructure
  - despite varying capabilities infrastructure is shared
  - telecommunications, operating systems, search engines, on-line shops, software, ...
- denying security to some, means denying it to all: crypto wars redux?
What is privacy?
- abstract and subjective concept, hard to define
  - fuzziness can be seen as an advantage
- dependent on cultural issues, study discipline, stakeholder, context
- privacy as confidentiality
  - “The right to be let alone”; focus on freedom from intrusion
- privacy as control: informational self-determination
- privacy as a practice
  - focus on user experience

Recent definition of privacy
(US) National Strategy for Trusted Identities in Cyberspace - Creating Options for Enhanced Online Security and Privacy
http://www.dhs.gov/xlibrary/assets/ns_tic.pdf
The appropriate use of personal information under the circumstances. What is appropriate will depend on context, law, and the individual’s expectations; also, the right of an individual to control the collection, use, and disclosure of personal information.

Data protection: legal basis
- 1950: European Convention on Human Rights (ECHR)
  - Art. 8 provides a right to respect for citizen’s “private and family life, his home and his correspondence,” subject to certain restrictions.
  - very broad interpretation by the European Court of Human Rights (Strasbourg)
  - part of Lisbon treaty (2009)
- 1981: Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Council of Europe)

EU Data Protection Directive
- data collected for specific and legitimate purpose
- proportional: adequate, relevant and not excessive (data minimization)
- with the subject’s awareness and consent
  - unless data is necessary for...
- data subject’s right to access, correct, delete her data
- data security: integrity, confidentiality of the data
  - unfortunately, millions of records with personal data are breached every year
  - weak enforcement, low penalties
  - creates database of databases
  - USA: fair information practices
    - many individual laws (HIPAA, California disclosure laws)

Soft privacy
- controller: main security “user”
- policies, access control, audits (liability)
- goal (data protection): purpose, consent, data security

Soft privacy
- data subject has already lost control of her data
  - in practice, very difficult for data subject to verify how her data is collected and processed
  - needs to trust data controller
Hard privacy

- system model: subject provides as little data as possible
- reduce as much as possible the need to "trust" other entities
- threat model
  - adversarial environment: communication provider, data holder
  - strategic adversary with certain resources motivated to breach privacy (similar to security systems)

Hard privacy

- subject is an active security "user"
- data minimization
- goal: protect against surveillance, interrogation, aggregation, identification [Solove]
- hard privacy solutions: technology (PETs)

Outline

- Context: information processing and uniqueness
- Do we need privacy?
- What is privacy anyway?
- Identity management
  - What is identity management?
  - ID management 1.0
  - ID management 1.5
  - Principles of identity and ID management 2.0
- Privacy by design
- Conclusions

What is Identity Management (IDM)?

- secure management of the identity life cycle and the exchange of identity information (e.g., identifiers, attributes and assertions) based on applicable policy of entities such as:
  - users/groups
  - organizations/federations/enterprise/service providers
  - devices/network elements/systems
  - objects (application process, content, data)

Pseudonymous identity management

- one-time pseudonyms: anonymity
- persistent pseudonyms: they become an identity
- solutions in between: partial identities
Identity: definitions (1)

- **attributes**: distinct & measurable properties belonging to a particular entity
- **identity**: dynamic collection of all of the entity’s attributes (1 entity: 1 identity)
- **partial identities**: specific subset of relevant attributes
- **identifier**: attribute or set of attributes of an entity which uniquely identifies the entity in a given context
- **credential**: piece of information attached to an entity and attesting to the integrity of certain stated facts

If these definitions reflect a specific vision on identity and identity management.

Identity: definitions (2)

- **entity authentication or identification**: using claimed or observed attributes of an entity to distinguish the entity in a given context from other entities it interacts with
  - Note: in computer security, often identification is providing one’s username and authentication is proving who an entity is
- **authorization**: the permission of an authenticated entity to perform a defined action
- **registration**: process in which a **partial identity** is assigned to an entity and the entity is granted a means by which it can be authenticated in the future

If these definitions reflect a specific vision on identity and identity management.

Identity management

- **physical world**
- **consumer space**
- **business environment**
- **e-government**
- **services and objects**

Identity management has many dimensions

Real life: growing number of applications

- financial, e-commerce, e-government, e-health, social networks, airlines, car rental, …
### Changing IT landscape

#### Cloud Computing
- RIA’s, AJAX, Flash, Silverlight, SaaS, IaaS, PaaS, Virtualization, RSS, Social Media, Wikis, ...

#### Web Services & SOA
- XML, SOAP, WS-*; REST, ESB, WSM, Java

#### Web Applications
- HTTP, HTML,.Net, Java, J2EE, TCP/IP

#### Client/Server & Distributed Computing
- VB, C++, SmallTalk, ERP, Tuxedo, MQ, DCE, COM, DCOM, Corba

#### Web Applications
- HTTP, HTML,.Net, Java, J2EE, TCP/IP

#### Mainframe/mi
- MVS, Top Secret, RACF, ACF

---

### Step 1: centralize (identity 1.0)

- **Integrate** entity authentication
  - but move authorization decision to application and services

- Embrace multiple authoritative sources
  - authoritative for attributes, not people
  - account names should be ephemeral
  - users should be free to select and change

- Dynamic rules, not static roles

---

### Integrated identity management (inside one organization)

- **Identity Manager**
- **Active Directory**
- **Web Services**
- **Authoritative Repositories**
- **Domain Controllers**
- **Applications/Services**

---

### How to grow? Step 2: federate (identity 1.5)

- **Federated identity**: credential of an entity that links an entity’s partial identity in one **context or trust domain** to an entity’s partial identity in another **context or trust domain**

- **Note**: can also be used inside an organization for convenience

---

### Single sign on: login only once

- **Identity Provider (IDP)**
- **Relying Party (RP)**
- **Can use any mechanism to authenticate!**
- **Trustworthy end system**

---

### Single Sign-On Variants

- **Initiate contact with IDP or with RP**
- **Access token**: push or pull

- **Token**: symmetric versus public key
  - symmetric token: IDP and RP have to share a secret key (example: Kerberos)
  - asymmetric token (digital signature): IDP and RP have to trust a common CA (example: SAML)
Single Sign-On

- convenient
- more secure than multiple passwords
- can leverage a single but more secure authentication mechanism
- risk of breach of authentication mechanism is substantially larger

identity: principles [Kim Cameron, Microsoft, '05] also called “laws”

1. user control and consent
2. minimal disclosure of information for a constrained use
3. disclosure limited to justifiable parties
4. directed identities: omni-directional and uni-directional
5. open – operators and technologies
6. human integration
7. consistent experience across contexts

Identity meta-system

- identity/attribute provider
- relying party (service provider)
- relying party (service provider)
- identity/attribute provider
- identity selector
- identity/attribute provider
- relying party (service provider)
- relying party (service provider)

Main issues: “identity 2.0”

- need consistent view for user: identity selector
- move from enterprise centric to user-centric (user in control)
- increased privacy

Anonymous credentials [Chaum'85]
The great thing about standards is... there are so many to choose from!

Trends in identity management
- evolution towards further integration and open systems: Kantara Initiative, Identity Commons' Open Source Identity System working group
- integration with mobile phones (SIM/USIM) and eID?
- architecture:
  - more pull than push (since too many applications)
  - user control may be replaced by third party supervision or management
- reputation based mechanisms originating from social networks
- cultural differences very hard to overcome: role of government, banks, credit rating bureaus,...

Anonymous communications
- Applications assume that the communication channels are secured / maintain privacy properties
- previous protocols are useless if the adversary can link transactions based on traffic data (e.g., IP/MAC address, IMEI, GPS, browser: https://panopticlick.eff.org/)

Classical communications security model
- data confidentiality
- data authentication
- entity authentication
- non-repudiation: origin/receipt
- availability

Anonymity – Concept and Model

Privacy by design - PriPAYD: car insurance
- GPS + Black box (computation) + transmit billing
- Flexible: easy change
- Easy computation
- Low cost
- Privacy friendly
- Third parties do not carry personal data
Road pricing: straightforward implementation

Privacy-Friendly Electronic Toll Pricing
No personal data leaves the domain of the user

Cryptology versus privacy
- crypto is success story: 1975-2010
  - from engineering discipline to science (with heuristic assumptions)
  - massive deployment
  - essential building block in IT systems
- even if issues with
  - weak legacy systems
  - long term security (e.g., MD5 story)
  - insecure implementations
  - attacks that bypass cryptography
  - usability

Privacy challenges
- privacy requirements and privacy by design
- finding efficient and secure mechanisms
  - complex systems require privacy at every level: the chain is as strong as its weakest link
  - proposed techniques keep getting broken: lack of models and proofs
  - secure implementation is even harder
  - easy to defeat by “changing” abstraction layer
    - cameras, RFID tags, unique device properties, singulation protocols, traffic analysis, …

Privacy and identity management challenges
- usability issues
- economic incentives
- awareness and transparency
- PETs can be misused: conditional privacy
- identity management is closely intertwined with our social and economic interactions
- identity management technology is evolving quickly, yet the concepts in our society change only slowly
  - concept of identity will probably evolve
  - ease of use and increased profiling has higher importance than data minimization

New challenging scenarios
- location privacy
  - real time
  - space-time relation
  - dummy traffic?
- ubiquitous environments
  - constrained devices
  - securing the physical link
- social networks: tension with data sharing
- cloud computing (or is it swamp computing?): outsourcing of storage/computations
Conclusion (1)

- Privacy is not "opposed" to security, but rather a security property
- Compliance is a strong driver
  - Data Protection
  - US disclosure legislation
- Soft Privacy is the state of the art
  - hidden costs of securing the data silos
- Hard Privacy solutions:
  - active research
  - poor deployment: cost/security benefit

Conclusion (2)

- security for society will grow
- privacy of individual will erode
- security of individual:?
  - concept of identity will probably evolve
  - need for interdisciplinary research
  - impact on organization of society not understood

The end

Thank you for your attention

Further reading

- Privacy Enhancing Technologies proceedings, Lecture Notes in Computer Science

Further reading