
Identity: the final frontier

Dr Geraint Price

Information Security Group

Royal Holloway, University of London

Outline

- Why is identity important?
- What is your identity?
- How do you identify yourself?
- Technology
- Who has information about you?
- When things go wrong...
- Conclusions

Why is Identity Important?

- First, we need to understand what we mean when we use the term *Information Security*...
- Some features include:
 - **Confidentiality**
 - Protecting information from **unauthorised** disclosure.
 - **Integrity**
 - Protecting information from **unauthorised** modification, and ensuring that information can be relied upon and is accurate and complete.
 - **Availability**
 - Ensuring information is available to **authorised** users when they need it.

Who am I?



- While joined by umbilical cord, there is undeniable proof you are your mothers offspring.
- Once cord is cut, “proof” of identity relies on (removable) tags plus procedures.

The Jackal

- In Frederick Forsyth's novel *The Day of the Jackal* the would-be assassin of General de Gaulle escapes detection through the generation of a false identity:
 - He assumes the identity of a dead child by obtaining the child's *birth certificate* and using it apply for a *passport*.
- The procedure - often referred to as the "Day of the Jackal technique" has been exploited in many real crimes.

Lord Buckingham – I

- Christopher Edward Buckingham had used the name of a dead baby for 23 years.
- He stole his name from the birth certificate of a baby who died 20 years previously.
- The real Christopher Buckingham was born in 1962 and died a year later
- Arrested in Dover in January 2005.
- During interviews he admitted assuming a false identity but refused to reveal his true identity.

Lord Buckingham – II

- Jailed in Nov 2005 for 22 months.
- His former wife and two children also did not know his true identity.
- His daughter decided to try and track down his true identity.
- In May 2006, it was discovered that he was missing American ex-serviceman called Charles Albert Stopford III.
- He was then due to be deported to the US.

What is an Identity? – I

- Attributed Identity:
 - Given name
 - Date of birth
 - Place of birth
 - Parents name
 - Mother's maiden name
 - ... i.e. birth certificate

What is an Identity? – II

■ Biographical Identity:

- What you've done – education, qualifications.
- Where you live(d).
- What you do – employment.
- Previous interaction with structured society.

■ Biographical Data:

- *Informal Defⁿ*: Any information that can be combined to identify an individual
- ... which in the online world becomes anything we do that leaves a *digital footprint*.

What is an Identity? – III

- Biometric Identity:
 - Different types of biometrics
 - Accurate (?)
 - Intrusive
 - Requires technology
 - Useful for verifying post-enrolment

How do you Identify Yourself?

- The Basic Principle:
 - A user identifies themselves to a trusted body.
 - The user is given an “*identifier*”.
 - The *identifier* is then given to authenticate the user.

 - **Note:** The authentication process merely confirms that the person producing the identifier is the person to whom it was issued.

The Main Uses for an Identity

■ Identification

- Its use in registration.
- Verification of the credentials held by an individual.
- Secure creation of the digital entity bound to that identity.

■ User authentication

- The cornerstone of most computer security.
- Using some process or procedure to verify the validity of a claimed identity.
- Binding the claimed identity to the one which it was originally linked.

■ Identity verification

- passports, ID Cards, etc.

Requirements to Prove Identity

- *Validity:*

- Is there sufficient supporting evidence to confirm that a person of that name exists.

- *Verification:*

- Can you establish whether the applicant is the “data subject” or “owner” of the valid identity references.

Requirements for authentication

- Test “validity” by:
 - Accessing a wide range of data
 - Examining the history of the data
 - Evaluating the quality of the data
- How do we verify that “this is John Smith”?
 - Test “verification” by:
 - Verifying that only the genuine “data subject” would know e.g. pervious address – marital status – time in employment – time at current address – ...

Electronic v Documentary Evidence

- Some credit reference agencies prefer electronic to documentary evidence.
- They consider documentary evidence to not be robust because...
 - How do you “reconstruct” the visual check of the document?
 - Data on documents is relatively static.
 - Logistics of a centralised checking process.
 - Documents easily forged/bought.
 - Genuine documents easily obtained falsely.
 - Documents used to breed other documents.
- Setting up multiple, corroborative, long-term electronic data sets *should* be more difficult...

Registration and Enrolment

- Registration issues:
 - It is crucially important
 - The registration process should be fit for purpose – what are you trying to protect?
 - Registration = Identification + Enrolment + ...
 - Reliance (and any potential over-reliance) on the resulting system.
 - High-assurance registration and enrolment procedures do not scale well.
 - Do not forget the insider threat.

Types of Authentication Mechanism

- In general, automated authentication mechanisms can be broken down into three categories:
 1. Something the user *knows*:
 - e.g. passwords, PIN numbers, passphrases, ...
 2. Something the user *has*:
 - e.g. token, smart card, ...
 3. Something the user *is*:
 - *Biometrics*: these can be further subdivided into *physical* characteristics and *behavioural* characteristics.

Password Policy

- It is often recommended that:
 - Users should adopt a large alphabet (at least alphanumeric with upper and lower case letters)
 - Passwords should be long (at least 8 characters?)
 - Passwords should be randomly generated
 - Passwords should be different for each system
 - Passwords should be changed frequently
 - Passwords should not be written down

To quote a famous tennis player:

“You cannot be serious!”

Tokens

- Idea well-established:
 - keys for doors, cabinets, cars, ...
 - magnetic stripe cards – used for ATMs, access control to secure sites, ...
- What types of tokens/cards?
 - Physical tokens
 - One Time Password generators
 - Challenge-response tokens
 - Smart Cards with handheld readers
 - Signing tokens

Biometrics

- The term biometric is derived from the Greek words *bio* (life) and *metric* (to measure).
- Biometrics is the measurement and statistical analysis of biological data.
- Definition by Biometrics Consortium:
 - “*automatically recognising a person using distinguishing traits*”
- Biometrics can be separated into two categories:
 - *Physiological* characteristics: e.g. fingerprint, iris pattern, hand structure, ...
 - *Behavioural* characteristics: e.g. signature, speech, ...

Implementation Issues for Biometrics

■ Errors

- False Rejection
- False Acceptance

■ Liveness testing

- Although biometrics should not be transferable, they can be copied and forged.
- *Is the biometric data being captured from a legitimate, live user who is physically present at the point of acquisition?*
- This is particular importance in remote authentication applications.

How good are biometric products?

- How can we find out how good a biometric product is?
 - Empirical tests of the product.
- In 2002, there were two independent test series of biometric products.
 - in Japan – Prof Tsutomu Matsumoto
 - in Germany – c't magazine
 - Both were able to fool the sensors a significant proportion of the time.
- In 2005/06 Deloitte/NPL carried out a survey of biometric products.

Comparing the 3 Factors

- Something known:
 - May be learnt by attackers
 - May be forgotten by users
 - Can be changed if compromised
- Something owned:
 - Can be stolen by attacker
 - May be copied by attacker
 - May be lost by user
 - May be replaced or changed
- Biometrics
 - Difficult to “copy”
 - Difficult to recover from compromise
 - Enrolment may be difficult (impossible?) for some users.

One-Way Authentication...

- A perceived weakness in Chip and PIN is that the terminal does not authenticate itself to the card.
 - Danger of false readers or readers which have been tampered with.
- Changes were made between GSM and UMTS (3G) mobile security to account for *false base-station* attacks.
- The system's job is to protect its users/clients.
- What's good for the clients is good for the system as a whole.
- ... but at what cost?

Additional Safeguards

- The banking industry have various forms of *Pattern Detection Systems*. For example:
 - Transaction Anomaly Detection:
 - Monitors usage pattern behaviour.
 - Geo-location information.
 - Suspicious access patterns.
 - Customer Device Identification:
 - Fingerprint the client device, e.g. IP address.
- Acts as an additional line of defence. Remedial action taken could be:
 - Real time alerting of the customer via email, SMS, etc.
 - Outbound manual phone call to customer.
 - Offline reversal and suspension of suspicious transactions.

Who Has Information About You?

- According to think tank Demos:
 - *“The average economically active individual in the developed world is on about 700 databases”.*
- According to the 2006/07 annual report of the Information Commissioner there are more than 287,000 data controllers in the UK.
- According to Garlik, personal details of the average Briton can be found in more than 1,000 places on the web.

Managing Personal Information

- EU Data Protection Directive:
 - Implemented by the Data Protection Act of 1998, establishing 8 basic principles:
 - Fair and lawful
 - For limited purposes
 - Adequate, relevant and not excessive
 - Accurate
 - Not kept for longer than necessary
 - In accordance with the data subjects rights
 - Secure
 - Not transferred to countries without adequate protection
- Information Commissioner's Office is responsible for UK policy and enforcement for data protection.

De-identification is hard

- Commonly used to remove data from a medical record that could be used to identify a patient.
- Techniques involve:
 - Personal identifiers removed
 - De-localisation
 - Record order scrambling
 - Numeric items banded, extremes truncated
 - Dates reduced
- Not restricted to medical data:
 - In August 2006 AoL released (supposedly de-identified search queries carried out by ½ M customers over a 3 month period.
 - AoL pulled the dataset when people started looking at the data to see if it might be identifiable... but there are copies out there.

The CD Scandal

- The missing disks
 - In October 2007, two CDs containing the personal details of 25 million people went missing.
- Jeremy Clarkson
 - Published an article via his column in The Sun expressing how this was not a problem.
 - To prove a point, he published his bank account details via his column.
 - Someone then set up a Direct Debit from his account to Diabetes UK for £500.
 - Was forced to admit he was wrong...

Identity Theft – I

- **Identity theft (or identity fraud)** occurs when someone wrongfully acquires or uses another person's personal data, typically for their own financial gain.
- Techniques:
 - Stealing mail or rummaging through rubbish.
 - Eavesdropping on public transactions to obtain personal data.
 - Stealing personal information in computer databases.
 - Infiltration of organisations that store large amounts of personal information.
 - Impersonating a trusted organisation in an electronic communication.
 - Spam

Identity Theft – II

- Surveys in the USA from 2003 to 2006 showed a decrease in the total number of victims but an increase in the total value of identity fraud to US\$56.6 billion.
- The 2003 survey from the Identity Theft Resource Centre found that:
 - Only 15% of victims find out about the theft due to a proactive action taken by a business.
 - The average time spent by victims resolving the problem is around 600 hours.
 - 73% of respondents indicated the crime involved the thief acquiring a credit card.

Identity Theft – III

- In Australia identity theft was estimated to be worth between AUS\$1 billion and AUS\$4 billion per annum in 2001.
- In the UK, the Home Office reported that identity fraud costs the UK economy £1.7 billion.
- Confusion over exactly what constitutes identity theft has lead to claims that statistics may be exaggerated.

Identity Theft – IV

- From Experian website:
 - Cases they had dealt with (up to June 2006):
 - It takes an average of 463 days for a victim to discover their details have been misused.
 - Average amount of credit gained fraudulently is £1921.80
 - Some of the ways in which they recommend you look after personal information:
 - Do not use “auto-complete” options in Internet browsers.
 - Taking care with WiFi and Bluetooth connections.
 - Taking care of portable storage devices (e.g. USB keyrings)

ID Theft/Fraud/Impersonation

- For a good example of how to get this type of credit fraud, the following YouTube link from *The Real Hustle* has an excellent exposition of such a scam:
 - <http://www.youtube.com/watch?v=A-3ShEroG40>
- Also, by *The Real Hustle* is use of a fake ID to get a locksmith to break into someone's house and steal £4,000 of goods.
 - <http://www.youtube.com/watch?v=t96SdVsVzUQ>

Phishing Attacks

- Social Engineering
 - Attacker discovers secret “information known”
 - Banking customers have been “prime” targets via email messages or fake websites.
- Countermeasures:
 - User education/awareness
 - Use of 2 or 3 factor systems so that compromise of 1 factor has limited impact.

Technology – a Benefit or Drawback?

- Technology definitely allows us to do things that wouldn't be otherwise possible.
- There are a few caveats though:
 - Function creep.
 - Management of large data stores is hard.
 - You need to have a clear understanding when you start out about what you want to be able to achieve in order to then manage that data securely.
 - Making it more “open” will make compromise (whether accidental or intentional) more likely.

The Future?

- Is this the future we are heading towards?
- <http://www.aclu.org/pizza>

Security is Never Perfect...



Conclusions

- In the modern world, our identities are formed by virtually everything we do.
- Protecting that identity is very much an end-to-end process:
 - Technology
 - Processes and procedures
 - People
- Managing large and well-connected data stores is non-trivial.
- There is now an easier flow between the formal and informal aspects of identity.
- More research needs to be carried regarding the protection of privacy in the modern world.