Identity, Security, and Privacy

Steven M. Bellovin

smb@cs.columbia.edu

http://www.cs.columbia.edu/~smb

This work by Steven M. Bellovin is licensed under a Creative Commons Attribution-Noncommercial 3.0 United States License.
Our Goals

- Protect our systems
- Protect our networks
- Protect our data
Our Strategies

- Build better walls
  - Operating systems
  - Firewalls
  - Applications
  - (Can this work?)

- Encrypt
  - Sometimes, encryption even makes sense, though not always...

- Authenticate
Why Authenticate?

- Restrict access to some resources
- Encrypt to the right party
- Accountability?
- End anonymity?
- Solve the cybersecurity problem?
- Because we can?
Accountability

- A primary stated purpose
- “On the Internet, nobody knows if you’re a dog” - but what if the dog bites?
- Some governments just want to restrict freedom of speech and access - but even in democratic societies, there are abuses of anonymity
Is Anonymity Good?

- Anonymity can be a powerful force for good
- It permits “whistleblowers” to disclose government or corporate wrong-doing
- In the U.S., there is a long tradition of anonymous political speech; it is strongly protected by law
The Cybersecurity Threat

- We all know there are serious security problems on the Internet
- If there is authentication, will bad guys be deterred?
- There is strong pressure from some quarters to mandate authentication, purportedly for that reason
The U.S. View

- Craig Mundie, head of Microsoft Research:
  - An “Internet Driver’s License”
  - You can lose your license for misbehavior

- The White House:
  - “Strong, interoperable” authentication schemes
  - Use online and offline
  - Changes - as yet unspecified - to (already weak) U.S. privacy laws

- (And what about the EU Data Retention Directive?)
The Real Cybersecurity Threats

- Hackers - these days, mostly motivated by profit
- Industrial espionage - quite possibly sponsored by governments
- Foreign government espionage
- Cyberwarfare (if there is such a thing)?

Will strong authentication help against any of these?
Hackers

• Hackers don’t use their own machines for most of their work

• Instead, they create *botnets* - armies of “bots”

• They are demonstrably capable of running arbitrary code on many computers belonging to many innocent people

• They steal all sorts of authentication credentials *today* - why should a new authentication scheme be stronger?

• Can it be stronger?
Thought Experiment: What Identity Should Be Used?

• Suppose I send virus-infected mail to my ISP’s mail server. It forwards the mail to my target. What identity is asserted for that hop?

• If it uses its own, it will be blamed for the virus

• If it uses my identity, it means identities are forgeable. Besides, it doesn’t have my private key

• Second thought experiment: what if I hack into a mail server and tamper with outbound mail? (Perhaps I insert a buffer overflow into the digital signature section of the mail.)
Governments

• Governments effectively control all CAs within their jurisdiction

• If a government wishes to issue fake credentials to spies - or to industrial spies benefitting its own country’s businesses - it will do so
  • There are many reports of fake passports issued by intelligence agencies today...

• No government will trust credentials issued by another government. How do such credentials protect against cyberespionage or cyberwarfare?
“Strong” Authentication

- A strong authentication scheme can’t use passwords - they’re too easily guessed or captured, and then replayed.

- Some sort of cryptographic solution is needed, most likely based on public key technology

- If the private key is stored in a file system, it will be compromised

- Some sort of trusted hardware is needed
Trusted Hardware

• Suppose the private key is stored in a smart card or TPM chip. Will this help?

• The smart card or TPM chip can’t talk directly to the outside. They can’t even talk to the web browser directly. Instead, they speak via the operating system.

• But we know that our operating systems are very vulnerable to attackers - which means that our trusted hardware can be controlled by the attackers.

• You think you’re logging in to your bank - but in reality, it’s the hacker who’s logging in... *This is already happening*. It’s a man-in-the-browser attack...
An Obvious (Over-Simplified) Authentication Protocol

A (Alice) wishes to authenticate to B (Bob)

A \rightarrow B: Certificate Authority, Certificate

B \rightarrow A: N

A \rightarrow B: \sigma_A(f(N))

What are the (non-cryptographic) problems?

(Note: analogous solutions with a KDC present a serious security risk in event of KDC compromise.)
Problems...

**Trustworthiness**

- Can we trust the signer?
- Can we trust the CA?
  - What if the CA is corrupt?
  - “A CA will protect you against anyone from whom it won’t take money” (Matt Blaze)
- But if these are the major threats, what is the point of strong authentication?

**Privacy**

- Bob learns A’s identity
- Exactly what is learned depends on what’s in the certificate - at the least, Bob can track uses of Alice’s public key
- The issue isn’t just governments; it’s also private corporations (especially in the U.S.)
Cybersecurity Through Authentication?

- It seems like it doesn’t work
  - The hackers can steal weak credentials or abuse strong ones
  - They don’t use their own machines in any event
  - The CAs can’t be trusted if governments are involved

- So why do it?

- Because - in its simpler forms - authentication is a solved problem

- We can’t secure our systems, and we can’t stop nasty governments, but we can authenticate...
“Something must be done. This is something. Therefore, it must be done.”
Real-World Issues

- How do we authenticate *people*?
- What about lost credentials?
- What about compromised credentials?
- What about accountability?
Identity Management

- Use secret-sharing to recover lost private key
- Give shares to people trusted by the individual - family, close friends, etc.
- Rotate share-holders as time passes: add a new spouse, remove an old one, etc.
- Properly identifying an individual is hard - but no harder (and no easier) than is done for passports, driver’s licenses, etc.

(Androulaki, Vo, and Bellovin, Engaging Data 2009)
Real World Credentials

- A credential to authenticate you to the government must be valid cradle-to-grave

- There may be a stretch of years when it isn’t used

- How is it issued? To whom? How are lost credentials handled?

- N.B.: the best way to acquire a fake passport is to steal someone’s identity when talking to the passport office; that way, the passport will be 100% genuine - and owned by the wrong person
Privacy Issues

- When the same pseudonymous identity is used in different contexts, a profile of the user can be built up.
- One link to a real person can tie a real person’s activities to that person.
- Such tracking can be and is being done by many parties.
- (Anonymization is very hard.)
Authorization Credentials

- To protect privacy, do not use identity-linked credentials
- Rather, use authorization credentials: the bearer has certain rights, regardless of identity
- Each use has its own credential
- Example: the person who deposited money to a bank account is the one who can withdraw it - but the credential that authorizes this doesn’t have any relationship to any other credential, even for the same bank
Authorization Certificates

- Not the conventional way of doing things - X.509 certificates are generally identity-based
- Still - well-understood mechanisms (e.g., SDSI/SPKI) for authorization certificates
- Some acceptance in the X.509 world (RPKI certificates for IP address blocks)
The Attribute is What Matters...
Unlinkable Credentials

- Work by Brands and by Camenisch and Lysyanskaya provide us with *unlinkable* credentials

- Each user has a master key pair

- The master private key can be used to generate *subcredentials* - a key pair that is verifiably derived from a given CA-issued certificate

- Subcredentials cannot be linked to each other or to the master credential

- Knowledge of a private subkey reveals the master private key
What Do We Have?

- Strong authentication
- Pseudonymity - as many (or as few) pseudonyms as you want
- Privacy
- No accountability
- No revocability in event of private key compromise
Accountability

- Revoke pseudonymity?
  - (By whom? Can you trust them?)
  - Focus of much prior work

- Reputation?

- Blacklisting?
Reputation in a Pseudonymous World

- Reputation should adhere to the real identity
  - A bad guy should not be able to discard a bad reputation by issuing a new pseudonym
- Positive and negative reputation
- Protocol non-adherence should not prevent assignment of negative reputation points
Pseudonymous Reputation

- After a transaction, Alice uses a digital cash “coin” to give Bob positive or negative repcoins
  - Complex mechanisms to ensure that Bob deposits negative coins...
  - Blind signatures used during deposit to hide Bob’s pseudonym from the bank

- The reputation bank uses blind group signatures to issue “certified balance” statements

- Unsolved (and probably unsolvable within the system): collusion to run up Bob’s score - but that’s a problem in non-anonymous reputation systems, too

(Androulaki, Choi, Bellovin, and Malkin, PETS 2008)
Blacklisting

- Sometimes, you never want to deal with a given individual again

- It is possible to blacklist a master credential: based on seeing a single subcredential, all future subcredentials derived from the same master credential can be rejected

- Unlinkability is still maintained - you cannot link the rejected subcredential to previously-accepted subcredentials

(Androulaki, Vo, and Bellovin, 2009)
Paying Taxes

- Suppose you open many bank accounts using anonymous, unlinkable credentials
- How can the government ensure that you pay taxes on your accounts
Simplified version

• When opening an account, people pay the bank a digital cash “account coin”
  • People can get as many account coins as they want, but the government knows how many they start with
  • When paying taxes, people also turn over their remaining account coins, so the government knows how many have been spent, and hence how many accounts exist

• The bank sends each (anonymous) account holder a signed account statement; both parties pass that information to the tax authority

(Androulaki, Vo, and Bellovin, ESORICS 2010)
More Privacy

• Instead of turning over each account balance, the blinded tax reports are created with a homomorphic commitment scheme

• As a result, the tax authority sees only the total balance, rather than the balances of each anonymous account
Disclaimers

- At this point, the protocols I’ve described are theoretical constructs
- The real world is far more complex
- We assume that certain underlying mechanisms - cryptographic primitives, digital cash schemes, anonymous networking technology, etc. - are available, adequately efficient, and secure
- Usability is a major challenge
Back to the Real World

- The White House scheme purports to be privacy-enhancing
  - Attribute certificates
  - Anti-linkage policies
  - Some anti-linkage technology - mechanisms are as-yet unspecified

- But - it calls for the “ability to support robust forensic capabilities”. Who can engage in such forensics, and under what conditions?
Where Does That Leave Us?

- Many people in high places want strong authentication when using the Internet.
- Such technology *cannot* solve the problems it is nominally aimed at.
- It may or may not use available privacy technologies, but the mention of forensics makes me skeptical.
What are the Policy Questions?

- There is (often) a societal interest in accountability
- There is also a societal interest in privacy
- What is the right tradeoff?
- What is the proper cost - temporal, financial, and procedural - for revoking anonymity?
- (Computer scientists have no more right to speak on policy issues than anyone else, but they have no less right. They’re also more qualified to discuss technical tradeoffs.)
What are the CS Questions?

- Given some set of answers to the policy questions, can we devise suitable technical mechanisms?
  - What are the assurance arguments for these mechanisms?
- If there is a revokability feature, how is it protected?
- How do we prevent leakage via lower-level (i.e., network layer) or higher-level (login name, writing style, interests) channels?
References 1


References 2

