Cybersecurity for Future Presidents

Lecture 8:
How can individuals be associated with
actions in a computer (and when should they be)?
What would conducting public elections by computer require?

Any Questions?

My office hours:
Wed. afternoon,
12-3pm, 442 RH

- About previous lecture?
- About homework?
- About reading? D is for Digital, Part III, Communications, introduction and Chapter 8, pp. 117-134 (Networking).

- Homework for next week:
- Debate readings on Canvas
- Supplementary for today's lecture: Chapter 2, "Authentication in the Abstract," pp. 33-54, in *Authentication through the Lens of Privacy*, NRC report. – in Supplementary readings file on Canvas

Cybersecurity events from the past week of interest to future (or current) Presidents:

- Utah Republican caucus used online voting yesterday
  - "$80,000 contact to London-based SmartMatic, which has set up online voting in the small country of Estonia."
  - http://www.smartmatic.com/
- FBI backs off legal confrontation with Apple (for now)
  - Maybe found another way in to the iPhone in question
- Android "Stagefright" exploit announced, can defeat ASLR protections on devices with patch level prior to Oct 15, 2015; hardware-specific attack required
  - Attack details here:
  - https://www.exploit-db.com/docs/39527.pdf
- Lithuanian "elves" counter apparently mercenary pro-Russian trolls on social websites

The lecture on one slide

How can individuals be associated with actions in a computer?

1. Accountability: being able to hold someone responsible for an action.

   Why it's important:

   - can provide incentives for corrective actions that otherwise won't exist

   When you may not require it

2. Fundamental technical issues, trusted path
3. Identification: a claim of who you are: userID, token?

   Identity for a context

4. Authentication: verification of the claim.
5. Authorization: decision to allow entity to perform some restricted function
6. Forensics: providing accountability after the fact
7. What are the requirements for voting systems for public elections with secret ballots?

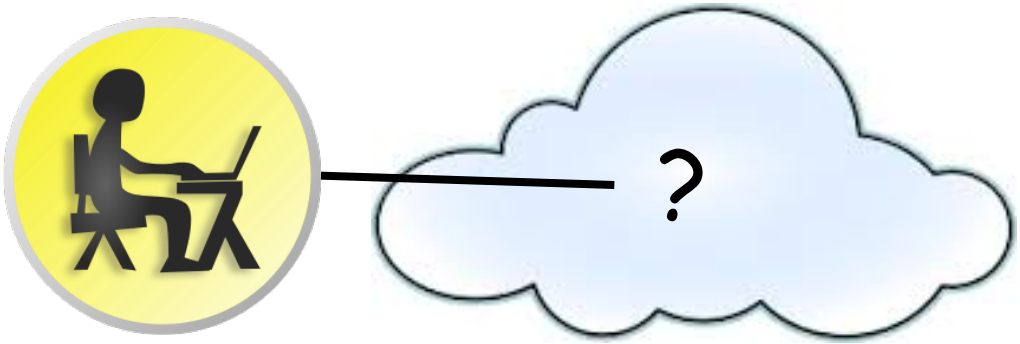
Why is accountability important for cybersecurity?

- Accountability provides a basis for accepting risk, for example in business transaction:
  - Amazon will do business with you if you can be held accountable for things you order (i.e., your credit card is valid)
  - You will do business with Amazon if you can hold them accountable for delivering what you order and standing behind it
- Security-critical operations need to be performed on behalf of an authorized individual
  - Software installation / update
  - Enrolling / removing users
  - Installing certificates
  - Etc.
- Note that not all operations need to be individually accountable
  - Web browsing from a public library: need to have a library card but need not be individually identified
  - Barbed wire and trespassing

Context for Establishing Accountability in Human-Computer Interaction

How does the system know whose fingers are on the keyboard?

When we type in a password, how do we know where it goes?

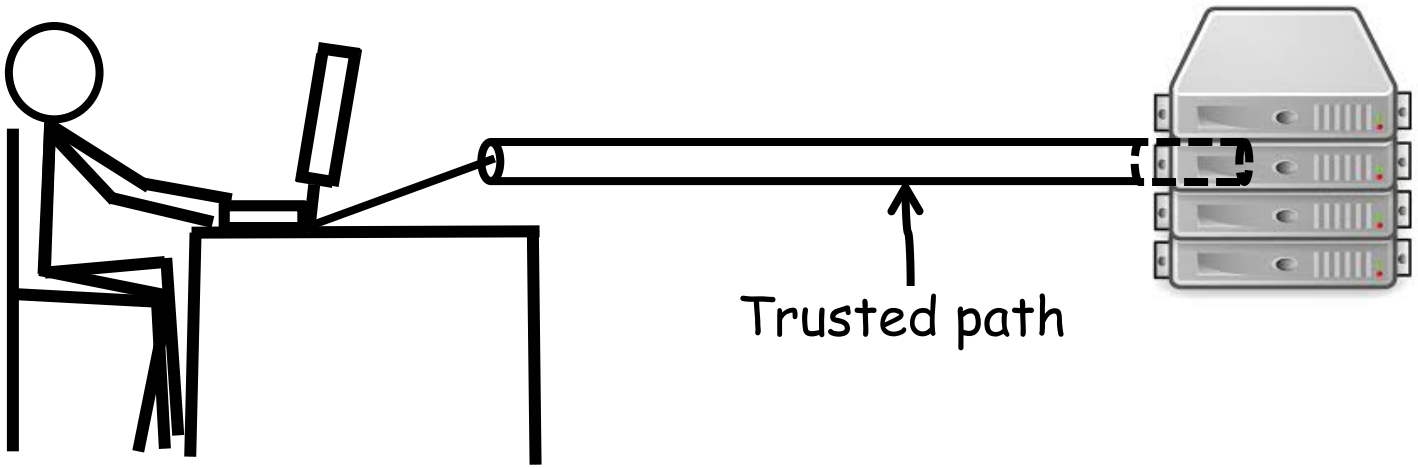


Fundamental Technical Issues

Smartphone observing its owner

- Identifying the user
  - Self declaration
  - Observation
- Trusted Path
  - How to be sure you are not being spoofed by the computer
  - How the computer can be sure you are not spoofing it
- Degrees of authentication
  - Authentication for "normal use"
  - Authentication for critical acts (installing software, adding/removing users, changing permissions
- Authentication over time

Trusted* Path mechanisms



- Trusted Path: mechanism that provides confidence that the user is communicating with what the user intended to communicate with, ensuring that attackers can't intercept or modify whatever information is being communicated <Wikipedia>
- Original intent: prevent malware from spoofing security labels
  - "secure attention key" allows user to cause a hardware interrupt, assuring "Trusted Computing Base (TCB)" gets control
- Modern equivalent: Windows: Ctl-Alt-Delete, MacOS: Apple-Opt-Esc, iPad, iPhone: the button at the bottom of the screen
- Virtual Private Networks (VPNs) use encryption to provide trusted path through network
- Problem: today's "TCB" is often the whole operating system

What is Identification?

How do you identify yourself?

- Talking to a human:

  "Hi, I'm <insert name>"

  - It's an assertion that you are (who you say you are)
  - If you are meeting in person, normally you can see each other
  - If you are speaking over the phone, your voice may be recognizable
- If you are "talking" to a computer:
  - For a typical laptop, identity = user ID
  - Also for a website (bank, store) also a user ID (which is often also an e-mail address
- If identity = user ID then nearly all of us maintain multiple identities
- What if your are talking to a smartphone or tablet? A "smart" appliance?
- An identified individual may have attributes (age, height, etc.); sometimes only these attributes and not full identity are needed for authorization

What is Authentication?

- Authentication is the process of establishing confidence that you are the person (or identity) you claim to be
  - Three parties to authentication: presenter, issuer, verifier
  - Presenter provides credential from issuer to verifier
- At a hotel desk or airport: providing a driver's license, passport, etc. (OK, these documents are referred to as "ID's" – but we will consider them authenticators)
  - For a computer, typically it's a password, could be a fingerprint

    Windows 10 to accept fingerprint, iris, face biometric (March 2015)
  - For a smartphone/tablet: usually a PIN, could be a fingerprint
- Without authentication, is confidentiality possible?

Sidebar: State Dept. Has Dept. of Authentications!

(for documents, not people) Provides assurance to foreign countries that U.S. documents (e.g., diploma, birth certificate, business incorporation) are legally correct (i.e., that the notarizations on them are valid)

Aspects of Authentication

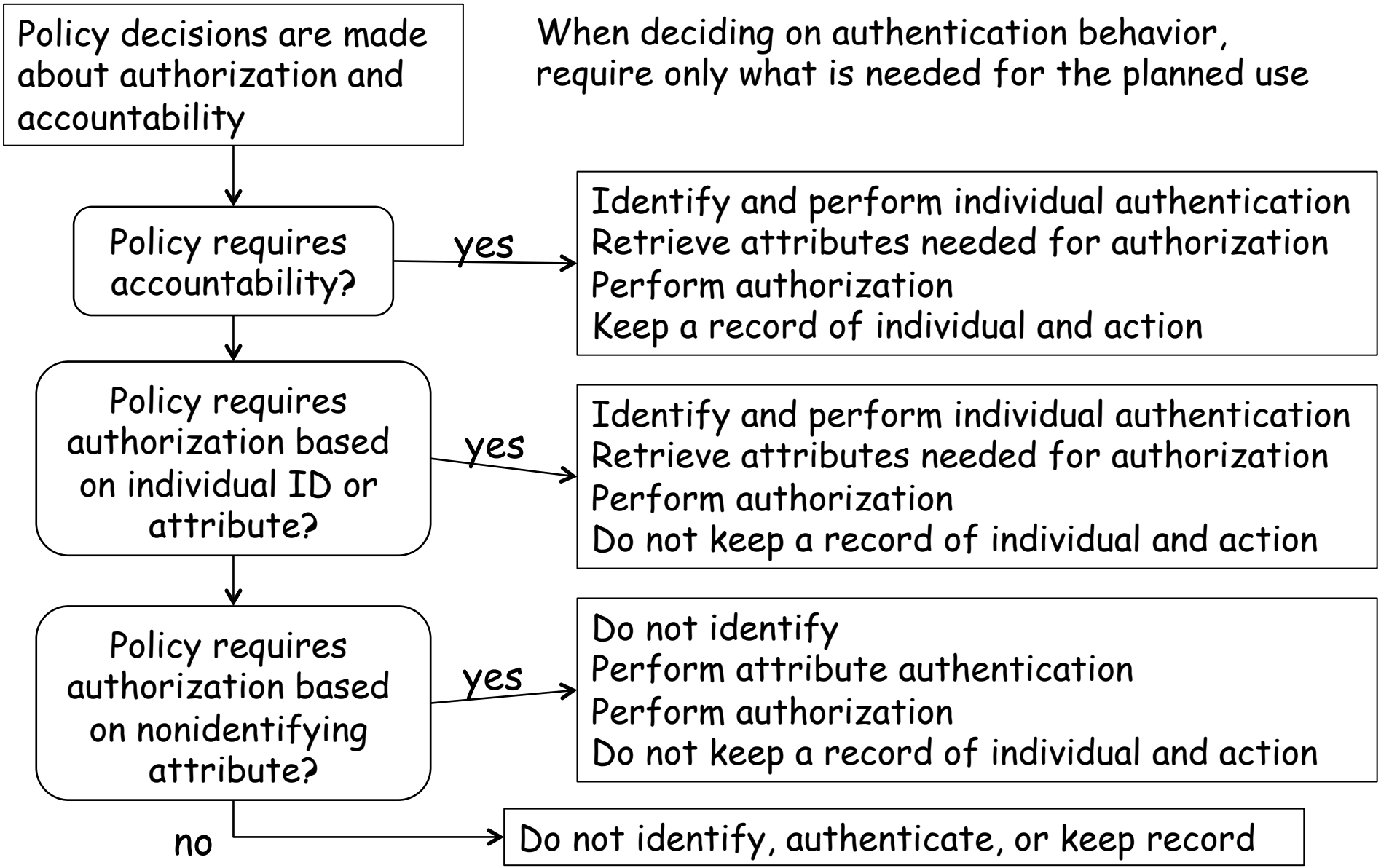
- Degrees of authentication
  - You might provide more or less evidence, for example
    work ID < driver's license < passport < birth certificate
- Multi-factor authentication
  - Something you know: password
  - Something you have: token
  - Something you are: biometric
- Discrete vs Continuous authentication
- Mutual authentication: assuring this is the device (website) you think it is, and the website assuring you are the person you claim to be
  - CAPTCHA*s: to prove to a machine the claim you are human

  - *Completely Automated Public Turing test to tell Computers and Humans Apart

What kind of authentication is appropriate?

Policy decisions are made about authorization and accountability

When deciding on authentication behavior, require only what is needed for the planned use



Taken from Fig. 2.1, p. 35, *Who Goes There? Authentication Through the Lens of Privacy*, NRC CSTB http://www.nap.edu/catalog/10656.html

Knowledge-based authentication ("Something you know") -- 1

Things not widely known ("security questions" – not secrets):

- Could be one or more questions for you to answer (e.g., mother's maiden name, your high school, pet's name, etc.). More questions might boost confidence of authentication
- Convenient – you don't have to remember anything special (except the answers you gave the system when you enrolled)
- Vulnerable – if attackers can discover this information about you (e.g. on Facebook, LinkedIn, etc.)
- Also, each time you reveal one some system now knows it
- OK for not-too-important services; not suitable for high assurance situation (e.g. bank withdrawal)

Knowledge-based authentication ("Something you know") -- 2

Secrets

- A PIN, password, passphrase
- Terrible: password should be hard to guess, easy to remember, not written down, regularly changed, … (??!!)
- But useful: can key it into any system, can share it (then change it!)
- Calculating the size of the password space
- Imposing constraints on passwords: characters vs words
  - Oxford English Dictionary (OED): $6.15*10^5$ words.
  - # random 10 character strings: $26^{10} = 1.4*10^{14}$
  - But $(6.15*10^5)^4 = 1.4*10^{23}$ -- so four words is much bigger space than 10 characters, and probably easier to remember
- Storing passwords so they can't easily be stolen (one-way functions again)
- Passwords for website logins

Token-based authentication ("Something you have")

- Car keys: active or passive
- One-time password tokens (or lists)
- RSA SecurID
- Mobile phone as a token (w/texts for codes)
- IP address of your computer? (not really useful, because it changes)
- MAC (Media Access Control) address of your computer's network interface (this is built in to the network interface card (NIC) on your computer and is unique to that card)
- Cookie on your computer: once you are authenticated, server may store a cookie on your machine that it can request on the next visit. This may be used to authenticate you and your machine. Not so good on a public computer (e.g. library, lab).

Biometrics ("Something you are")

- Note: biometrics are NOT secrets!
  - Even though you may think of them that way, you leave them everywhere: fingerprints, facial image, iris, DNA, etc.
- How are biometrics stored in a computer?
  - In general, features are extracted during enrollment
  - Template constructed from extracted features is stored
  - During authentication, biometric is re-sensed, features extracted, and features compared with those in the stored template
    - Note the importance of Trusted Path between user/sensor and authentication software
    - "Close enough" → match. False positive, false negative is possible
    - Difference between confirming identity (this input matches the claimed identity) and determining identity: here's an input, who is it?
- What if a biometric database (e.g. template store) is stolen?
  - Do you need to change your fingerprints/iris/face? [no]
  - Can someone spoof your identity with the stolen information?[maybe a little easier, but Trusted Path is an important protection
  - Do your biometrics become useless for identification?

Authorization

- Authorization: assuring that rules concerning how some (computing) resource may be used are obeyed
  - Who can read this message?
  - Who can post to this website?
  - Who can install software on this machine?
- Why we have I&A: you want to perform some actions that are authorized for your claimed identity
- Sometimes, authorization may not require authentication, if full accountability is not needed, as noted in the flow chart earlier
  - "Must be at least 4' tall to ride the roller-coaster"
- Sometimes, accountability can be provided even though identification is not: deposit for billiard balls, for example

Forensics: accountability after the fact

- "Forensic" having to do with argument (debates are forensics) but particularly legal argument, arguments in court
- In the context of this lecture, trying to establish accountability for some criminal act or act of war, after the fact
- Computer forensics
  - Trace evidence in computers
- Network forensics
  - Evidence from network traffic or network infrastructure
- Relationship to side channels: often forensic evidence is provided by analysis of digital breadcrumbs the perpetrator may be unaware of
  - May lead investigators deeply into implementation details
- How much evidence do you need for criminal attribution?
- How do we know who was behind the Sony attacks?
- How is all this affected by the fact that bots are so easy to obtain?

Some types of forensic evidence

- Files found on a hard drive
  - How did the file get there? → whose fingers were on the keyboard?
- Files that the user may have deleted, but weren't overwritten
- Data remaining in bits of disk drive that have been discarded by the drive
- Data in backup files (local or remote)
- Data in temporary files or RAM
- Network data
  - Network traffic logs on client
  - Server records of traffic
  - Router logs

What if some or all of the data are encrypted?

- Can the accused be legally coerced to supply the encryption key?
  - It depends: at a border crossing, probably so
  - In a criminal investigation, if the prosecution can show probable cause

Attribution

Attribution: regard something as being caused by someone (or something); ascribing a work to a particular author, artist, speaker

Attack attribution is often a difficult issue to resolve in cybersecurity, because of a general lack of accountability: it's hard to know

- Whether a particular machine with a particular IP address was used wittingly or unwittingly as a source of attack or storage
- Whose fingers were on the keyboard when an attack occurred
- Who might have provided incentives to the individual whose fingers were on the keyboard?

Evidence sometimes used to support attribution

- System logs showing evidence of malware infestation (or not)
- External observation of actual computer use (e.g. video surveillance)
- Internal observation of computer use (via keylogger, e.g.)

National Identity Systems – policy and practice

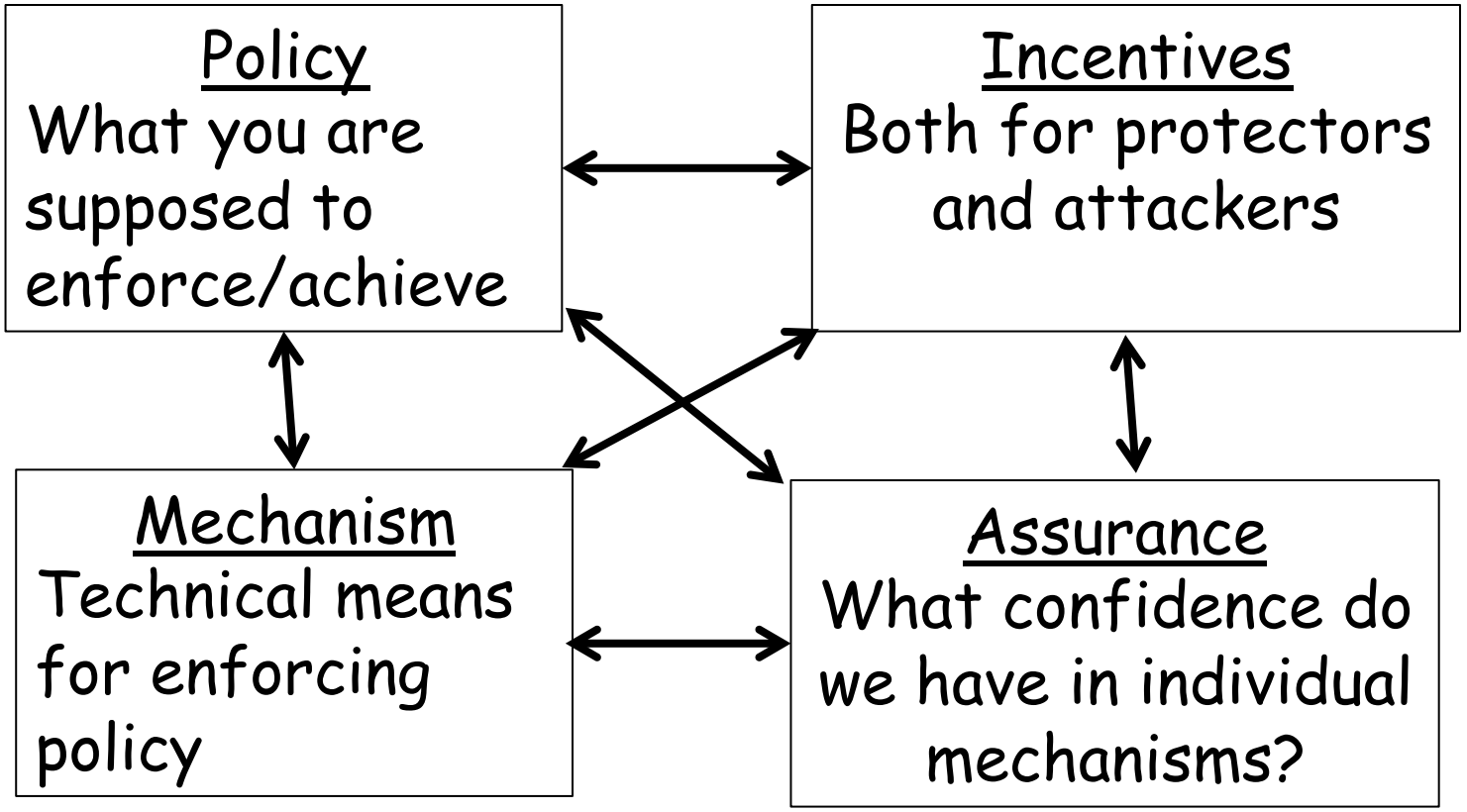
- Other countries
- US
  - History
  - Government record systems
  - Citizen concerns
  - Medical records
  - National Strategy for Trusted Identities in Cyberspace
  - RealID (state drivers license) program
  - Passports

Public Policy

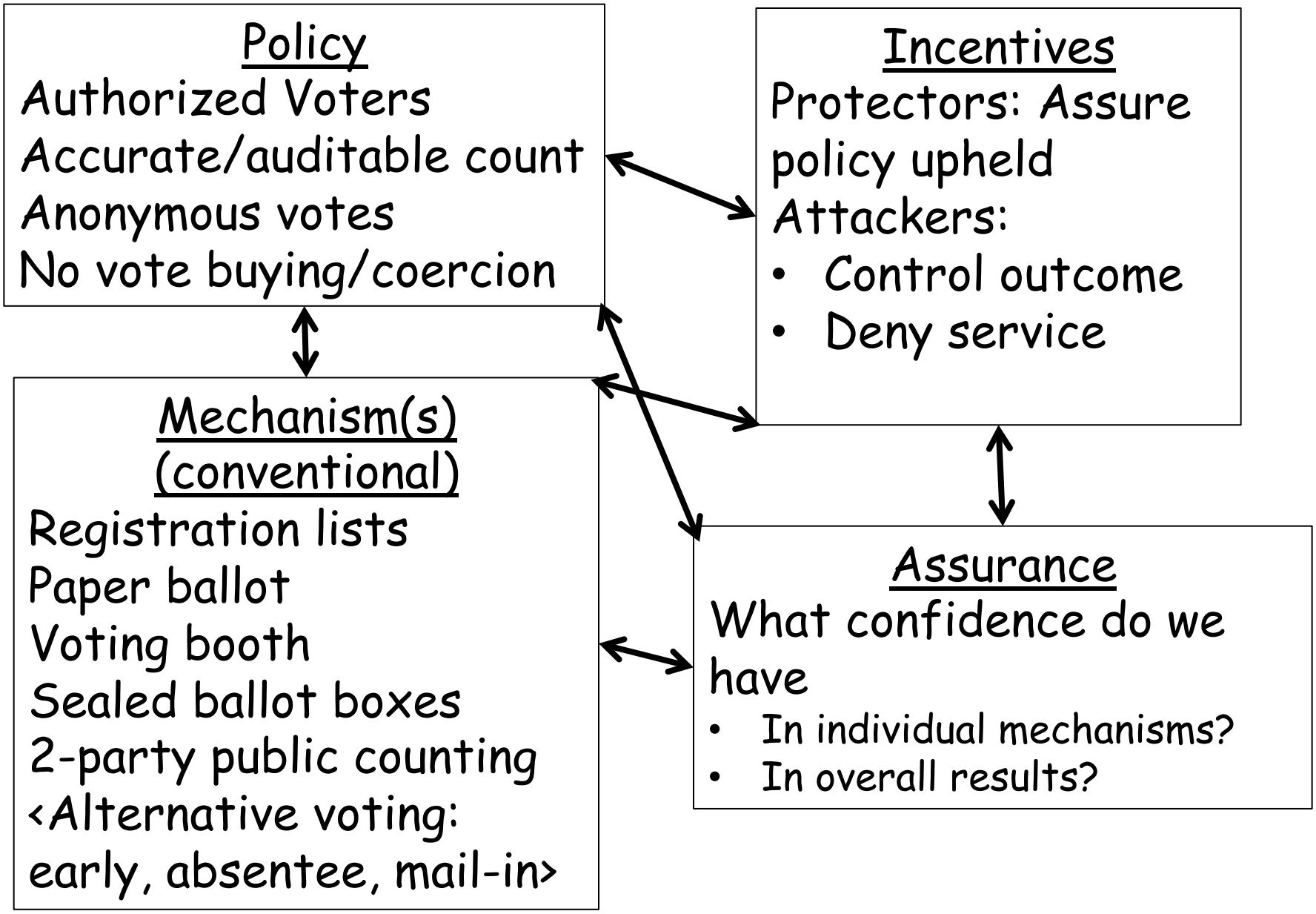
Requirements for Voting in Public Elections

- Establish eligibility to vote in election
- Assure vote is cast as voter intends
- Assure the cast vote can't be mapped back to the voter (beyond what vote totals may tell)
- Provide assurance to voter that her/his vote was counted correctly
- Prevent vote buying, vote fraud
- Auditable process in case of trouble

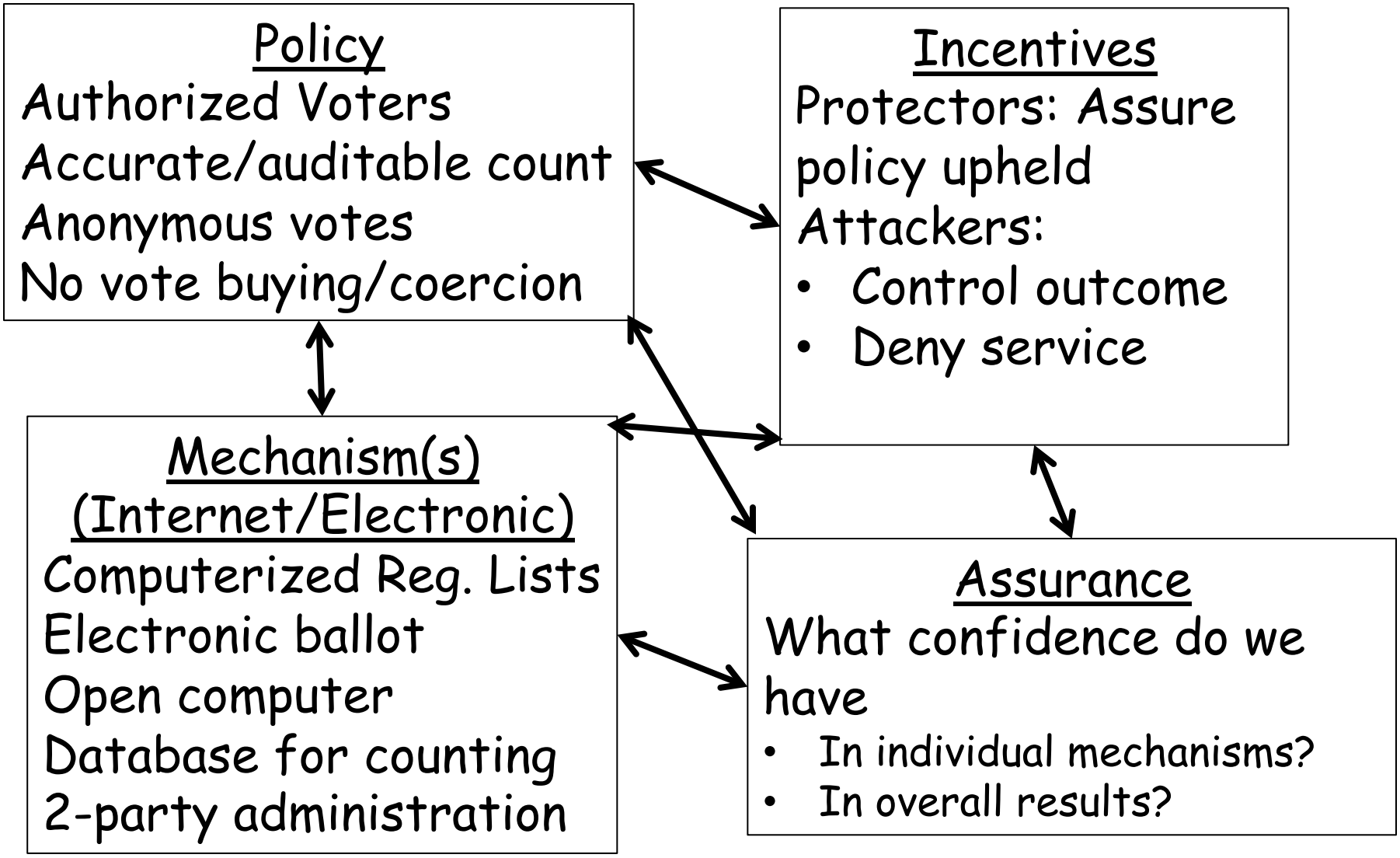
Security Engineering view of a system



Security Engineering view of public elections



Security Engineering view of public elections



Cryptographic Methods Useful in Elections

- Cryptographic Checksums
  - To permit detection of changes to data (e.g., ballots, but also software components)
- Secret sharing
  - To permit decryption keys to be split into pieces (like the pieces of a treasure map)
  - (k,n) threshold cryptography: n shares, k<n needed to reconstruct key
  - Think of a polynomial of degree n: need n points on the curve to reconstruct it
    - Give each party a subs
- Secure Multiparty Computation
  - Simple example (compute average weight)

Public Policy

Some U.S. Election History

- U.S. adoption of secret ("Australian") ballot after 1884 Presidential election, requiring
  1. Government printed ballot with all nominees, parties, issues
  2. Distributed only at polling place
  3. Marked in secret
- Note that mail-in ballots (and absentee ballots, depending) violate 2 & 3, yet Washington and Oregon conduct all elections this way.
- U.S. election administration is generally local; elections are run by states, counties, precincts
- 2000 election problems in Florida prompted Congress to pass Help America Vote Act (HAVA) of 2002 with aims
  - Replace punch-card and lever voting machines
  - Create the Election Assistance Commission
  - Establish minimum election administration standards
- Funded with $3B appropriation for states to purchase new equipment satisfying
  - Permit the voter to verify ballot selections
  - Provide the voter the ability to change/correct ballot before casting
  - Notify voter of any over-votes and permit corrections

Public Policy

HAVA consequences

- Boom in electronic voting equipment marketplace
  - Funding bubble provoked one-shot products rather than long term stable development
  - Decentralized purchasing (many states, municipalities) meant many inexperienced groups making decisions
- Systems put in place without adequate security vetting
- Various studies, initiatives, exposed vulnerabilities in these systems
- Some states now turning to paper-based systems (for auditability) with scanners (for checking valid ballots and counting)

(Potential) Roles for Cryptography in Voting Systems

- Encryption
  - Secrecy of ballot
- Digital signatures
  - Integrity of ballot
- Secret sharing
  - Enforce "2-man rules"
- Vote verification
  - Schemes for voter to verify that vote is counted

Public Policy

Alternative Voting in the US

- Early Voting
- Absentee Voting
- Mail Voting

Role of Access Control in Voting Systems

- Roles:
  - Voter
  - Identification/authentication
  - Registration checker
  - Election judge/administrator
  - Vote tallying

How would you manipulate / defraud such a system?

What does the security/integrity of the system depend upon?

- Software
- Hardware
- Procedures