DSci529: Security and Privacy In Informatics

Technical Means of Protection: a Primer on Computer Security

Prof. Clifford Neuman

Lecture 3
29 Jan 2021
Online

Course Identification

- DSci 529
  - Information Privacy
  - 4.0 units
  - Website http://ccss.usc.edu/529
- Class meeting schedule
  - Noon to 3:20PM Friday's
  - Online (and if we do transition to Hybrid, OHE100C)
- Class communication
  - bcn@isi.edu (for now)

Course Outline

- Overview of informatics privacy
- What data is out there and how is it used
- **Technical means of protection**
- Identification, Authentication, Audit
- The right of or expectation of privacy
- Social Networks and the social contract
- Measuring Privacy
- Big data – Privacy Considerations
- Criminal law, National Security, and Privacy
- Civil law and privacy
- International law and conflict across jurisdictions
- The Internet of Things
- The future – What can we do

Current Event Discussion

- http://csclass.info/USC/INF529/s21-lec3-ce.html

A primer in-security

- Much of todays lecture will be review for students in the security informatics program.
- The objectives of today's lecture are to provide an overview of security for the non-security specialist.
  - Useful for those in data informatics
  - Useful for those outside of engineering
- What you need to know about the security of the information you manage

Next Weeks Lecture

- A second lecture on security techniques focused on Identification, Authentication and audit.

The Three Aspects of Security

- Confidentiality
  - Keep data out of the wrong hands
- Integrity
  - Keep data from being modified
- Availability
  - Keep the system running and reachable
  - Keeping the data available.

Policy v. Mechanism

- Security policy defines what is and is not allowed
  - What confidentiality, integrity, and availability actually mean
- Security mechanisms are tools we use to protect our systems.
  - Mechanisms enforce policy.
  - Mechanisms may solve intermediate problems.
    - Authentication, Audit
    - Containment

Important Considerations

- Risk analysis and Risk Management
  - Impact of loss of data.
  - Impact of disclosure.
  - Legislation may play a role.
- The Role of Trust
  - Assumptions are necessary
- Human factors
  - The weakest link

In The Shoes of an Attacker

- Motivation
  - Bragging Rights
  - Revenge / to inflict damage
  - Terrorism and Extortion
  - Financial / Criminal enterprises
  - Nation State motivations
- Risk to the attacker
  - Can play a defensive role.

Security and Society

- Does society set incentives for security.
  - OK for criminal aspects of security.
  - Not good in assessing responsibility for allowing attacks.
  - Privacy rules are a mess.
  - Incentives do not capture gray area
    - Spam and spyware
    - Tragedy of the commons

Why we aren't secure

- Buggy code
- Protocols design failures
- Weak crypto
- Social engineering
- Insider threats
- Poor configuration
- Incorrect policy specification
- Stolen keys or identities
- Denial of service

Security Mechanisms

- Encryption
- Checksums/Hashes
- Key management
- Authentication
- Authorization
- Audit
- Firewalls
- Virtual Private Nets
- Intrusion detection
- Intrusion response
- Development tools
- Virus Scanners
- Policy managers
- Trusted hardware

Loosely Managed Systems

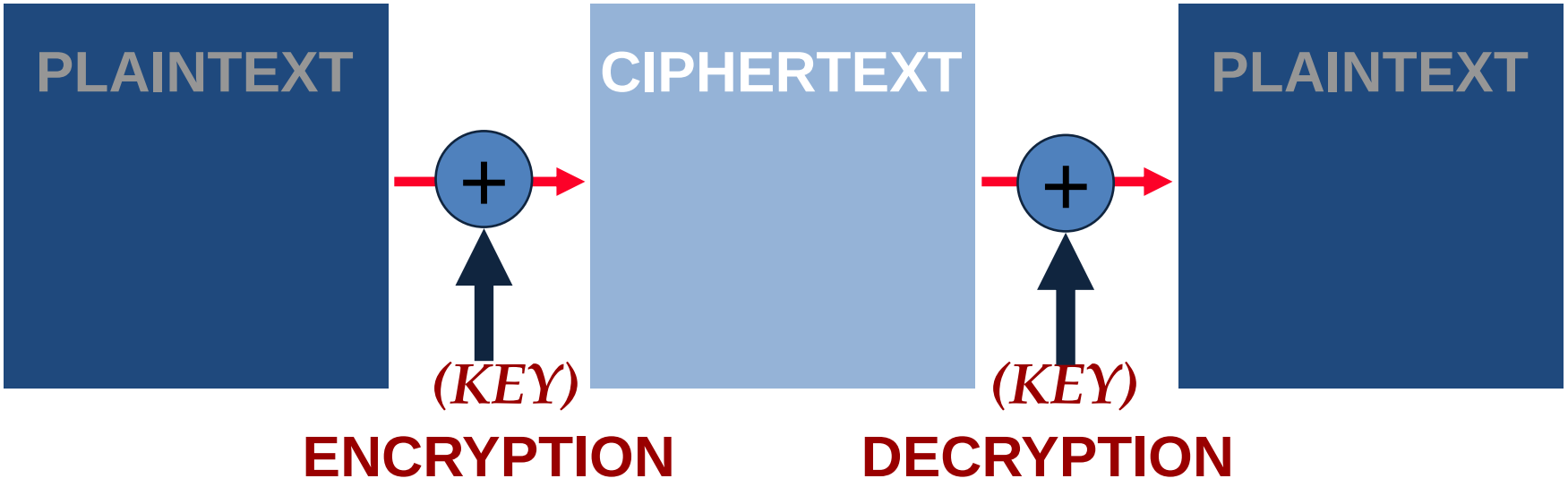
- Security is made even more difficult to implement since today's systems lack a central point of control.
  - Home machines unmanaged
  - Networks managed by different organizations.
  - A single function touches machines managed by different parties.
    - Clouds
  - Who is in control?

Cryptography and Security

- Cryptography underlies many fundamental security services
  - Confidentiality
  - Data integrity
  - Authentication

- It is a basic foundation of much of security.

Encryption used to scramble data



Digital Signatures

- Provides data integrity
  - Can it be done with symmetric systems?
    - Verification requires shared key
    - Doesn't provide non-repudiation
- Need proof of provenance
  - Hash the data, encrypt with *private* key
  - Verification uses public key to decrypt hash
  - Provides "non-repudiation"
    - But what does non-repudiation really mean?

Policy: The Access Matrix

- Policy represented by an Access Matrix
  - Also called Access Control Matrix
  - One row per object
  - One column per subject
  - Tabulates permissions
  - But implemented by:
    - Row – Access Control List
    - Column – Capability List

Malicious Code / Subversion

- Modification of data
  - Deletion, changes to balances
- Exfiltration
  - Obtain sensitive information
- Advertising
  - Targeting or generating
- Propagation
  - Extend ones reach
- Self Preservation
  - The Subversion issue

Zombies/Bots

- Machines controlled remotely
  - Infected by virus, worm, or trojan
  - Can be contacted by master
  - May make calls out so control is possible even through firewall.
  - On order of 10-30 percent
  - Other malicious code
    probably 60%

Spyware

- Infected machine collect data
  - Keystroke monitoring
  - Screen scraping
  - History of URL's visited
  - Scans disk for credit cards and password.
  - Allows remote access to data.
  - Sends data to third party.

Economics of Malicious Code

- Controlled machines for sale
- “Protection” for sale
- Attack software for sale
- Stolen data for sale
- Intermediaries used to convert online balances to cash.
  - These are the pawns and the ones that are most easily caught

Economics of Adware and Spam

- Might not ship data, but just uses it
  - To pop up targeted ads
  - Spyware writer gets revenue for referring victim to merchant.
  - Might rewrite URL's to steal commissions.

Architecture: A first step

- Understand your applications Information Flow:
  - What is to be protected
  - Against which threats
  - Who needs to access which apps
  - From where must they access it
- Do all this before you invest in the latest products that salespeople will say will solve your problems.

What is to be protected

- Is it the service or the data?
  - Data is protected by making it less available
  - Services are protected by making them more available (redundancy)
  - The hardest cases are when one needs both.

Classes of Data

- Decide on multiple data classes
  - Public data
  - Customer data
  - Corporate data
  - Highly sensitive data

  (not total ordering)
- These will appear in different parts of the network

Classes of Users

- Decide on classes of users
  - Based on the access needed to the different classes of data.
- You will architect your system and network to enforce policies at the boundaries of these classes.
  - You will place data to make the mapping as clean as possible.
- You will manage the flow of data

How to think of Firewalled Network

Crunchy on the outside.
Soft and chewy on the inside.

– Bellovin and Merrit

Firewalls

- Packet filters
  - Stateful packet filters
    - Common configuration
- Application level gateways or Proxies
  - Common for corporate intranets
- Host based software firewalls
  - Manage connection policy
- Virtual Private Networks
  - Tunnels between networks
  - Relationship to IPsec

Protecting the Inside

- Firewalls are better at protecting inward threats.
  - But they can prevent connections to restricted outside locations.
  - Application proxies can do filtering for allowed outside destinations.
  - Still need to protect against malicious code.
- Standalone (i.e. not host based) firewalls provide stronger self protection.

Intrusion Types

- External attacks
  - Password cracks, port scans, packet spoofing, DOS attacks
- Internal attacks
  - Masqueraders, Misuse of privileges

Attack Stages

- Intelligence gathering
  - attacker observes the system to determine vulnerabilities (e.g, port scans)
- Planning
  - decide what resource to attack and how
- Attack execution
  - carry out the plan
- Hiding
  - cover traces of attack
- Preparation for future attacks
  - install backdoors for future entry points

The Human is the Weak Point

- Humans make mistakes
  - Configure system incorrectly
- Humans can be compromised
  - Bribes
  - Social Engineering
- Programmers often don't consider the limitations of users when designing systems.

Some Attacks

- Social Engineering
  - Phishing – in many forms
- Mis-configuration
- Carelessness
- Malicious insiders
- Bugs in software

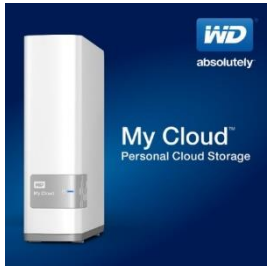
Trusted vs. Trustworthy

- We trust our computers
  - We depend upon them.
  - We are vulnerable to breaches of security.
- Our computer systems today are not worthy of trust.
  - We have buggy software
  - We configure the systems incorrectly
  - Our user interfaces are ambiguous regarding the parts of the system with which we communicate.

Defining The Cloud

- The cloud is many things to many people
  - Software as a service and hosted applications
  - Processing as a utility
  - Storage as a utility
  - Remotely hosted servers
  - Anything beyond the network card
- Clouds are hosted in different ways
  - Private Clouds
  - Public Clouds
  - Hosted Private Clouds
  - Hybrid Clouds
  - Clouds for federated enterprises

Examples of Cloud Services



Risks of Cloud Computing

- Reliability
  - Must ensure provider's ability to meet demand and to run reliably
- Confidentiality and Integrity
  - Service provider must have their own mechanisms in place to protect data.
  - The physical machines are not under your control.
- Back channel into own systems
  - Hybrid clouds provide a channel into ones own enterprise
- Less control over software stack
  - Software on cloud may not be under your enterprise control
- Harder to enforce policy
  - Once data leaves your hands

Defining Policy

- Characterize Risk
  - What are the consequences of failure for different functions
- Characterize Data
  - What are the consequences of integrity and confidentiality breaches
- Mitigate Risks
  - Can the problem be recast so that some data is less critical.
    - Redundancy
    - De-identification
  - Control data migration within the cloud

Controlling Migration

- Characterize Node Capabilities
  - Security Characteristics
    - Accreditation of the software for managing nodes and data
  - Legal and Geographic Characteristics
    - Includes data on managing organizations and contractors
  - Need language to characterize
  - Need endorsers to certify
- Define Migration Policies
  - Who is authorized to handle data
  - Any geographic constraints
  - Necessary accreditation for servers and software
    - Each node that accepts data must be capable for enforcing policy before data can be redistributed.
  - Languages needed to describe

Enforcing Constraints

- With accredited participants
  - Tag data and service requests with constraints
  - Each component must apply constraints when selecting partners
    - Sort of inverting the typical access control model
- When not all participants are accredited
  - Callbacks for tracking compliance
  - Trusted computing to create safe containers within unaccredited systems.

Cloud Security Summary

- Great potential for cloud computing
  - Economies of scale for managing servers
  - Computation and storage can be distributed along lines of a virtual enterprise.
  - Ability to pay for normal capacity, with short term capacity purchases to handle peak needs.
- What needs to be addressed
  - Forces better assessment of security requirements for process and data.
  - Accreditation of providers and systems is a must.
  - Our models of the above must support automated resolution of the two.

INF529: Security and Privacy In Informatics

Technical Means of Protection (part 2)

Prof. Clifford Neuman

Lecture 4
5 February 2021
12:00 Noon
Online

Why Identity is So Important

Most policy specifications are identity based
- CIA policies last week, depend on knowing who is trying to read or change data.

Most security breaches include some form of impersonation
- Malicious code runs as an authorized user
- Passwords stolen by phishing

Identifiers link data and make it findable/searchable.
- Whether right or wrong, this identification has significant impact on users.

Identification vs. Authentication

Identification

- Associating an identity with an individual, process, or request

Authentication

- Verifying a claimed identity

Basis for Authentication

Ideally
- Who you are

Practically
- Something you know
- Something you have
- Something about you
  - (Sometimes mistakenly called things you are)

Something you know

Password or Algorithm

- e.g. encryption key derived from password

Issues

- Someone else may learn it
  - Find it, sniff it, trick you into providing it
- Other party must know how to check
- You must remember it
- How stored and checked by verifier

Something you Have

Cards

- Mag stripe (= password)
- Smart card, USB key
- Time varying password

Issues

- How to validate
- How to read (i.e. infrastructure)

Case Study – RSA SecureID

Claimed - Something You Have

Reduced to something they know

How it works:

Seed

Synchronization

Compromises:

RSA Break-in

Or man in the middle

Implication of Authentication Failures

Access to data (confidentiality or integrity) as if attacker were the authorized user.

For one system, or for many systems.

Failure can propagate through system.

Don’t depend on a less critical system.

How Authentication Fails

Stolen Credentials

- Passwords
- Cards / devices
- Copied biometrics

- The role of malicious code
  - GP devices can not protect credentials

Problems of e-mail authentication

And password recovery

- General email security is weak
  - Emails can be intercepted
  - Or are sent to a compromised account

  - http://abcnews.go.com/Business/online-security-time-upgrade-passwords/story?id=36223462

Implications of password reuse

If users use same password on multiple systems.

- The security of the users account on any system becomes dependent on the security of the weakest system used with that password.

- https://thestack.com/security/2016/02/03/taobao-hack-20-59-million/

Implications of Data Compromise

The biggest reason most people are concerned with data breach is:

- The data is used for authentication
  - Social Security Numbers
  - Credit Card Numbers
  - PINs

The future of second factors

What do we have

Who takes responsibility

- This is a major stumbling block
- Responsibility means liability

Back to Identification

Identification is important for attribution

- Audit trails and logs
- Identifying wrongdoers

- Identification can be wrong
  - Attacks facilitated through compromised machines
  - IP Addresses that change

Points of Identification

Biometric Data
Surveillance Data
Internet Addresses
MAC Addresses
Payment details

Audit and Detection

Identification data is recorded in audit logs routinely together with observed actions
- Accesses, authentication attempts, failures, etc.

Systems use tools to process this audit data and alert on suspicious actions.

Attack Detection

- External attacks
  - Password cracks, port scans, packet spoofing, DOS attacks
- Internal attacks
  - Masqueraders, Misuse of privileges

Attack Stages

- Intelligence gathering
  - attacker observes the system to determine vulnerabilities (e.g, port scans)
- Planning
  - decide what resource to attack and how
- Attack execution
  - carry out the plan
- Hiding
  - cover traces of attack
- Preparation for future attacks
  - install backdoors for future entry points

Intrusion Detection

- Intrusion detection is the problem of identifying unauthorized use, misuse, and abuse of computer systems by both system insiders and external penetrators
- Why Is IDS Necessary?

IDS types

- Detection Method
  - Knowledge-based (signature-based) vs behavior-based (anomaly-based)
- Behavior on detection
  - passive vs. reactive
- Deployment
  - network-based, host-based and application - based

The Anonymity Debate

- Should we be required to identify ourselves when using the internet?
  - What about other situations

- Authentication of Attributes vs Identity
  - Over 21, but without showing your DL

Are we less secure today

No: It is the environment that has changed

- Users today demand instant and universal access to everything they can get.
- In the past, data was better protected because it wasn't accessible
- Some data was better protected because no-one collected it to begin with.

Understanding this can help you prepare

- Develop a containment architecture
- Different data can have different accessibility
- Collect and distributed data to mitigate the impact of the inevitable breach

Containment Architecture Action PLAN

Conduct an Inventory – of data

- What Kinds of Data do you have in your business
- How is it handled and where is it handled
- Who needs access to this data
- Which systems need access to this data
- How is it protected in transit, and in situ

Containment Architecture Action PLAN

Conduct an Inventory – of physical assets

- What Kinds of systems do you have
  - E.g. POS terminals, servers, network hardware
- Understand the access to each system
  - Employees, customers, etc
- How are the different classes of systems protected from one another
  - Network zones, etc
- How do you contain breaches to particular zones.

COLLECT BASLINE INFO ON ALL ASSETS

Software and system checksums
- Used to detect changes to the system
- To identify which assets are affected
- To enable recovery – reinstall those affected systems

Baseline data communication from all assets
- In you network infrastructure, use this to identify anomalous flows
- As they happen to block exfiltration
- From Logs to identify where data went and how much, and over what time periods

This much PREPARATION CHANGES THE STORY

From:
XYZ corporation is the latest company to report that the personal information of 70M customers may have been compromised.
TO:
XYZ corporation reports that users of its beach city store between October 1st and 3rd may have been affected by …

What Technology should I deploy

- Audit and intrusion detection
- Encryption throughout the systems
  - Data in transit and data on disk
  - As close to the source as possible
- System mapping/configuration tools
  - Align with your containment architecture
- Effective identity and policy management
- Configuration management systems

Addressing Data Compromise

Don’t collect the data
- If you don’t need it
- Design systems so you don’t need it

Don’t use the data for authentication
- Why do we use public information for authentication:
  - Mothers maiden name
  - Password reset information
  - SSN

Why such poor practices

Internet services require scalability to be viable.
Automation provides that scalability.
Effective Customer service does not.

It is all about avoiding personal contact with the customer, which would require more staff.