INF529: Security and Privacy In Informatics

The Truth is Out There

Prof. Clifford Neuman

Lecture 2
18 Jan 2019
OHE 100C

Course Identification

- INF 529
  - Information Privacy
  - 4.0 units
  - Website http://ccss.usc.edu/529
- Class meeting schedule
  - Noon to 3:20PM Friday's
  - Room OHE 100C
- Class communication
  - inf529@csclass.info

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

Semester Project

All students are expected to prepare and present a 30 minute lesson on a topic related to privacy that is of interest to them.

- If on a topic that is already in the syllabus, your presentation will be made in the week that the topic is covered in class. The next slide shows some possible topics that align with lectures (your title should be more specific).
- If on a topic that is not already in the syllabus, I will assign a week from your presentation, based on available time in lecture, and based on relevance.
- Please send me proposed topics for your class presentation by Thursday the 25th. You can suggest multiple topics if you like... if so let me know your order of preference. All that you need is a short title and a one sentence description. Topics may be chosen from among the topics listed in the syllabus for the class, or you may propose topics around any particular problem domain (e.g. type of system, type of business, type of activity) for which you will provide a thorough discussion of privacy (or privacy invading) technology and policy.

Possible Presentations

- Week3: Technical means of protection
- Week4: Identification, Authentication, Audit
- Week5: The right of or expectation of privacy
- Week6: Social Networks and the social contract
- Week7: Measuring Privacy
- Week8: Big data – Privacy Considerations
- Week9: Criminal law, National Security, and Privacy
- Week10: Civil law and privacy
- Week11: International law and conflict across jurisdictions
- Week12: The Internet of Things
- Week13: The future – What can we do

Weekly Current Event Assignment

Beginning today students should find a current event regarding security or privacy (preferably privacy) in the news and and me a URL, title, and three sentence write-up which we will discuss in class. The write-up should be sent to me before 7AM the morning of our lecture. These will be due from all students for each class and will count toward you class participation grade.
We will discuss all submissions in class and you will be called on to provide additional information about your submission.

Example:

Google's Art Selfie App Offers A Lesson In Biometric Privacy Laws – NPR 1/18/18
Story explains why the Art Selfie app is not available to users in Illinois or Texas.
Explains the unintended consequences of privacy legislation.

INF529: Security and Privacy In Informatics

The Truth is Out There

Prof. Clifford Neuman

Lecture 2
18 Jan 2019
OHE 100C

What is Cloud Computing

- In the 1990's Sun Microsystems (now owned by Oracle) used the phrase the network is the computer.
  - This afternoon we will talk about the kinds of data you can obtain from these parts of computers.
  - With cloud computing, many functions are no longer performed on a physical device, but are performed elsewhere on the network.

Examples of Cloud Services

From an Investigators Perspective

- What clouds mean is that:
  (+) You may be able to find evidence without gaining possession of a subject computer.
  (-) This evidence may be located elsewhere on the internet, sometimes out of your jurisdictional reach.
  (-) The issue of jurisdictions is not known with any clarity.
  (-) Most users don't understand where their data is stored, or that copies of old data remain present in the cloud for extended periods.
  (+) This might provide you with the ability to seize such data with appropriate warrants or orders.
  (-) You will have to contact many organizations to obtain relevant data.
  (-) Data is harder to protect in the cloud.

Cloud Discussion

- What is stored in these and other services.
- How is this data protected.
- What kind of access can you get to this data.
- What you need to know before using.
- What are the implications.

Initial Homework Assignment
(due before todays class)

- What sensitive information is available about you?
  - Apple
  - Google
  - Amazon
  - Enumerate the data
  - Where is it stored?
  - To whom is this data available?
    - Under what conditions?
  - How long is the data retained?
    - How (or) can you remove it?
  - What can be done with this data?
- Let's review your discussion

You Are Being Tracked

- **Location**
  - **From IP address**
  - **From Cell Phones**
  - **From RFID**
- **Interests, Purchase History, Political/Religious Affiliations**
  - **From RFID**
  - **From Transaction Details**
  - **From network and server traces**
- **Associates**
  - **From network, phone, email records**
  - **From location based information**
- **Health Information**
  - **From Purchases**
  - **From Location based information**
  - **From web history**

Why Should you Care?

- Aren't the only ones that need to be concerned about privacy the ones that are doing things that they shouldn't?
- Consider the following:
  - Use of information outside original context
    - Certain information may be omitted
  - Implications may be mis-represented.
  - Inference of data that is sensitive.
  - Such data is often not protected.
  - Data can be used for manipulation.

Traffic Analysis

- **Even when specifics of communication are hidden, the mere knowledge of communication between parties provides useful information to an adversary.**
  - **E.g. pending mergers or acquisitions**
  - **Relationships between entities**
  - **Created visibility of the structure of an organizations.**
  - **Allows some inference about your interests.**

Information Useful for TA

- **Lists of the web sites you visit**
- **Email logs**
- **Phone records**
- **Perhaps you expose the linkages through web sites like linked in.**
- **Consider what information remains in the clear when you design security protocols.**

Linkages – The Trail We Leave

- **Identifiers**
  - **IP Address**
  - **Cookies**
  - **Login IDs**
  - **MAC Address and other unique IDs**
  - **Document meta-data**
  - **Printer microdots**
- **Where saved**
  - **Log files**
  - **Email headers**
- **Persistence**
  - **How often does Ip address change**
  - **How can it be mapped to user identification**

Unlinking the Trail

- **Blind Signatures**
  - **Enable proof of some attribute without identifying the prover.**
  - **Application in anonymous currency.**
  - **Useful in voting.**
- **What about BitCoin**
  - **Contrary to popular belief, the flow of funds in bitcoin are completely public (public blockchain)**

Unlinking the Trail

- **Anonymizers**
  - **A remote web proxy.**
  - **Hides originators IP address from sites that are visited.**
  - **Usually strips off cookies and other identifying information.**
- **Limitations**
  - **You are dependent on the privacy protections of the anonymizer itself.**
  - **All you activities are now visible at this single point of compromise.**
  - **Use of the anonymizer may highlight exactly those activities that you want to go unnoticed.**

Onion Routing

- **Layers of peer-to-peer anonymization.**
  - **You contact some node in the onion routing network**
  - **Your traffic is forward to other nodes in the network**
  - **Random delays and reordering is applied.**
  - **With fixed probability, it is forwarded on to its destination.**
- **TA requires linking packets through the full chain of participants.**
  - **And may be different for each**

Protecting Data in Place (at rest)

- **Many compromises of privacy are due to security compromised on the machines holding private data.**
  - **Your personal computer or PDAs**
  - **Due to malware or physical device theft**
- **Countermeasures**
  - **For device theft, encryption is helpful**
  - **For malware, all the techniques for defending against malicious code are important.**
  - **Live malware has the same access to data as you do when running processes, so encryption might not be sufficient.**

Hardware Key Loggers

- Key loggers could be located inside the computer case or hidden inside a keyboard. Some loggers store the keystrokes in onboard memory while others could transmit the keystrokes wirelessly via 802.11/Bluetooth etc.

Forensics

- **Tools are available to recover supposedly deleted data from disks.**
  - **Similar tools can reconstruct network sessions.**
  - **Old computers must be disposed of properly to protect any data that was previously stored.**
    - **Many levels of destruction**
  - **Tools like whole disk encryption are useful if applied properly and if the keys are suitably destroyed.**

Privacy – Retention Policies

- PII (personally identifiable information)
  - Is like toxic waste
  - Don’t keep it if you can avoid it
- Regulations
  - Vary by Jurisdiction
  - GDPR
  - But if you keep it, it is “discoverable”

Visibility of Addresses

- MAC or physical addresses seen only on the local network.
- IP Addresses visible to the endpoints and intermediate nodes.
- Private IP addresses behind NAT boxes (Network Address Translators) may not be visible
- IP addresses are often transient, assigned as needed through DHCP.
  - Attributing an action based on IP address requires knowing the IP address assignment at a particular point in the past.

Volatile Information

- Information that is lost when a system is powered down or lose power that can be found by a “live” forensics analysis
  - System time
  - Logged-on user(s)
  - **Open files**
  - Network & connection information
  - Process information
  - Process-to-port mapping
  - **Process memory** (possibly including passwords/crypto keys)
  - Network status
  - **Clipboard contents**
  - **Command history**
  - Mapped drives
  - Shares (resources made available by system over a network)

Nonvolatile Information

- Information kept on secondary storage that persists after power down, that can be collected in a forensic analysis
  - **Hidden files**
  - Swap files
  - Index and metadata files
  - Hidden alternate data streams
  - Search indices
  - Unallocated clusters
  - Unused partitions
  - **Hidden partitions**
  - **Registry settings** (including some kinds of encryption keys)
  - Connected devices
  - **Event logs**

Deleted Files

- Usually remain on the storage system until overwritten with new data later
  - Until completely overwritten, these “partially removed” files can be partially or completely recovered using special forensic program tools that can read every sector and piece the old information together.
  - Different storage devices use different methods for removing these deleted files from the directory structure and sooner or later overwriting them.

Gathering Email Header Information

- Is different for each email client application
- See the following for a good reference:
  - https://support.google.com/mail/answer/22454?hl=en

Analyzing the header information

- Use tools like:
  - https://toolbox.googleapps.com/apps/messageheader/
  - (paste header information into web application for ease of analysis)

Discussion of Data Sources

- De-Privacy by Sophia Catsambi from Yale News 18 January 2018

  "By clicking or navigating this site, you agree to allow our collection of information on and off Facebook through cookies." This was the message that greeted me when I opened my page on Monday evening. I promptly clicked "Agree" and rushed to meet a friend for dinner. The notification didn't catch my attention for more than the minimum amount of time required to dismiss it. Not for a second did I pause to think: "Wait, do I actually want Facebook to have access to my personal information, even when I'm engaging in activities that aren't even happening on its platform?"

Current Events

Is the " 10 Year Challenge" on Facebook a privacy scheme discussed as a meme? – CBS News Story explains about how facebook denies about the initiation of the ten year challenge. The article explains about the unintended consequences of privacy on usage of this data for advanced facial recognition Algorithms. - Deepti Rajashekharaiah Siddagangappa

Cyber Security Experts Weigh In on Marriott/Starwood Data Breach – HospitalityTech 12/03/2018 - This article states a bit details about technical aspects about Marriott's data breach, it shows that one of the most serious reasons of the breach was a wrong management of encryption keys. -- Fumiko Uehara

LA Sues Over Weather Channel App's 'Misleading' Data Collection US News 01/04/19 This article explains about how huge tech companies such as IBM (which owns The Weather Channel) do ask for permissions to access personal data but do not make it very clear how the data will be used. In this case, the users' location data was sold to advertisers and interested third parties. -- Chloe Choe

Unprotected Government Server Exposes Years of FBI Investigations - The Hacker News 01/17/2019 This article talks about exposed data belonging to the Oklahoma Department of Securities (ODS) since at least November 30, 2018. Talks about the kind of data that was exposed. -- Sevanti Nag

Current Events

[Tim Cook wants FTC to let consumers both track and delete their personal data]

The Verge 1/17/19 - The article talks about how Tim Cook wants stricter privacy legislation in the US. Also, the article highlights a new bill in Congress, the American Data Dissemination Act. Explains that the act while useful for regulating tech companies has the consequence of making it difficult for states to have stricter regulations. -- Ahmed Qureshi

[We Need New Privacy Laws,' urges Apple CEO Tim Cook]

Apple CEO Time Cook has stated this week that in 2019 that it's important to stand up for ones' right to privacy- citing multiple data breaches and non-transparent data practices as a reason. Cook stated that there are 9 data brokers in the U.S. and there aren't many regulations on what data they have and what they do with it. He urges that those data broker firms need to register with the FTC and provide transparent information on how the data is used, as well as creating new laws allowing consumers to see what data is held and collected about them and why- as well as having the right to correct or delete the data. - Kate Glazko

[Senator Rubio Unveils Privacy Proposal] Digital News Daily 1/17/18

Story explains a new Act that overrides many state privacy laws. The Act addresses issues about consumers not having the ability to elect what is and is not collected on them. This Act is the starting point for that topic. -- Andrew Carmer

Current Events

Apple Is Part of the Privacy "Shadow Economy" The Mac maker has little choice but to indirectly contribute. TMFNewCow 1/17/19

Apple tried protecting consumer's privacy. CEO Tim Cook addressed that consumers should not tolerate about data breaches. He compared with GDPR (General Data Protection Regulation by European Union) that Apple is a small part of the "shadow economy" by using Google search in their safari. -- Sophia Choi

Apple CEO Tim Cook pushes for increased privacy oversight by Congress 1/17/19 Fox Business News

Explains Tim Cook's idea for an increase in privacy regulations in the US, allowing users to track and delete their data and see exactly how their data is being bundled and sold to different companies -- Anupama Sakhalkar

Apple keeping up the privacy promise despite market concerns

Apple is losing out to market pressure and eventually announced support for iTunes and airplay in competitors like Samsung, Sony etc. But apple still keeps up to its promise to uphold privacy. "Apple tells The Verge that Samsung's smart TV ad-tracking features cannot track viewing usage within the iTunes Movies and TV Shows app, in another example of Apple's focus on privacy. -- Kavya Sethuraman

Current Events

Apple Maps gooses DuckDuckGo in search privacy partnership CNET 1/15/2019

Search engine DuckDuckGo which was known for its promises in protecting users' privacy is partnering with Apple to share users' location history of DDG's users. Even though DDG promised to delete users' data immediately after using it; however, DDG will share information such as location history, IP address, along with other information with Apple. Both companies declined to provide information to the public on their terms of partnership. -- Faris Almathami

Could HIPAA Be Repealed, Replaced with a Unified Federal Privacy Law?

The Information Technology and Innovation Foundation (ITIF) has called to repeal privacy regulations, particularly HIPAA, to be replaced with a standardized framework. The think tank highlights that a unified approach would help bring about greater consumer privacy protections for the Digital Age by providing protections based on sensitivity of data. The article seeks to address different perspectives of regulating health-related data in the digital world that would maximize consumer welfare. -- Jacqueline Dobbas

Privacy Concerns Largely Ignored At Annual Consumer Electronics Show

NPR 1/12/19 - With the latest CES (Consumer Electronics Show) causing a lot of buzz about new technology and devices, these same companies lack a strategy to offering more privacy and security of consumer data. Rather than providing ideas on how they, for example Apple, are protecting data, they are simply stating that there is security through advertisements. -- Brianna Tu

Current Events

Facial recognition scanning unwitting tourists and shoppers in London's West End

The Independent 1/17/19 - An article that discusses the controversy of testing a new facial recognition technology. London's West End Police claims that refusing to be scanned won't raise suspicion. In addition to the privacy breach, it is believed that this technology has a high rate of false positives.
-- Abdulla Alshabanah

Ring Camera's Possible Privacy Issue The Intercept 1/10/2019

Story discusses how according to inside sources, Ring has been providing their teams with access to a folder that contains almost any video created by Ring cameras. Article describes the use of human operators to train AI systems using this sensitive data. Discusses privacy issue of this practice.
-- Lance Aaron See

Google Home and Alexa Hack to Improve Privacy FastCompany 1/14/2019

This article discusses a necessary functionality for the smart home devices that many view as a possible invasion of privacy. The smart homes require access to microphones to listen out for a keyword (i.e. Alexa), and when used accesses the buffer of recording to catch what was requested. It then goes on to describe newly created tool, that when attached to the smart home devices that essentially plays white noise until activated. -- Joseph Mehltretter

Current Events

hackers-attack-hundreds-of-high-profile-german-politicians-journalists

The article describes a recent breach of privacy involving several German politicians. Large amounts of their personal information was stolen and released on the internet, and it appears only one particular party was spared. This is an example illustrating that malicious privacy violation is prevalent, even if those responsible for the sensitive information do not distribute it themselves. -- Ann Bailleul

Cops told: No, you can't have a warrant to force a big bunch of people to unlock their phones by fingerprint, face scans

The Register 1/14/19 - Recently (Jan 10, 2019), US Magistrate Judge Kandis Westmore in the Northern District of California has ruled that compelling use of biometric features to unlock electronic devices in a warrant application is in violation of Fifth Amendment rights. Her ruling was based on Carpenter vs. US where the Supreme Court ruled that the government had violated the Fourth Amendment rights by seizing and collecting cell phone location records without a warrant. Previous rulings on pass codes for digital devices had been grey in the area of biometric features, but have found numeric or alpha-numeric pass codes to be testimonial, meaning that they cannot be compelled under the Fifth Admendment right to not incriminate oneself. -- Charlene Chen

Collection #1 Breach Exposes a Record 773 Million Email Addresses

PCMagazine – 1/17/19 - -- Collection #1 is a set of email addresses and passwords totaling 2,692,818,238 rows," Hunt explained in a Thursday blog post. "It's made up of many different individual data breaches -- suspicious collection of files on the cloud service Mega.
-- Gene Zakrzewski

Current Events

How The Government Shutdown is Affecting National Cyber Security - CNBC 1/14/2019

Story explains about how our nation's risk may be at complete risk due to most of IT being furloughed. It begs the question that even if personnel were brought back on board now, how much damage has been done? It speaks of a few instances of when people with little to no motivation take information out of the government and how this can be replicated in a worse manner when people return to their work.
-- Jairo Hernandez

Massive 'Fortnite' security hole allowed hackers to take over accounts, eavesdrop on chats - NBCNews 01/16/19

Researchers in cybersecurity firm found flaws in Epic games' subdomains that allow hackers to have access to victims' accounts. Vulnerabilities in the login process, specifically in the single sign-on system, could be exploited to capture the victim authentication token provided by third party such as Facebook or Google. A hacker must first craft a phishing link and send it to the victim and the victim must click on the link for the attack to succeed. -- Abdullah Altokhais

Millions of Chinese CVs exposed on cloud server -- BBC 1/14/19

A security firm Hackenproof found a database containing two hundred million resumes and personal information, which did not have any security feature to protect the data. It was created by scraping a few Chinese job-seeking sites. The data had been copied twelve times before it was deleted from Amazon cloud where it was found. -- Yulie Felice

Current Events

Why your smartwatch and wearable devices are the next big privacy nightmare ZD Net | January 14, 2019.

With the rise of wearable technology, both device makers and relevant applications leveraging wearable data seek to sell or monetize data back to insurance companies - this is a multi-billion dollar market and is only expected to grow. This would also result more importantly in significant privacy and consent issues as many companies do not explain to users where and how their data is being used including anonymization, processes for in the event of data breach, and how relevant policies such as GDPR are met. In addition, in the healthcare industry, how patient data is being used to determine coverage (at the individual and aggregate) and receiving medical insurance could be in conflict with current regulatory coverage mandates and data protection laws (e.g., HIPAA). - Arjun Raman

Some Telco's continue selling consumer location data, Congress prepares to take action - ZDNet 1/12/19

Although many Telco's have promised to stop selling location data, many continue the practice including AT&T, T-Mobile and Sprint. Congress is now considering taking action to ensure private consumer location data is protected. -- Dewaine Reddish

There's a simple reason why your new smart TV was so affordable: It's collecting and selling your data – Business Insider

Companies like TCL and Vizio were among some of the many electronic manufactures to offer tremendous TV deals, such as a 65-inch 4K smart TV with HDR capability for less than $500. While this may seem like a great deal to most consumers, some of them may or may not be aware of that these companies are collecting your data and selling it to third-parties. - Louis Uuh

INF529: Security and Privacy In Informatics

Technical Means of Protection

Prof. Clifford Neuman

Lecture 3
25 Jan 2019
OHE 100C

Course Identification

- INF 529
  - Information Privacy
  - 4.0 units
  - Website http://ccss.usc.edu/529
- Class meeting schedule
  - Noon to 3:20PM Friday's
  - Room OHE 100C
- Class communication
  - inf529@csclass.info

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

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
- 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
- Key management
- Authentication
- Authorization
- Accounting
- Firewalls
- Virtual Private Nets
- Intrusion detection
- Intrusion response
- Development tools
- Virus Scanners
- Policy managers
- Trusted hardware

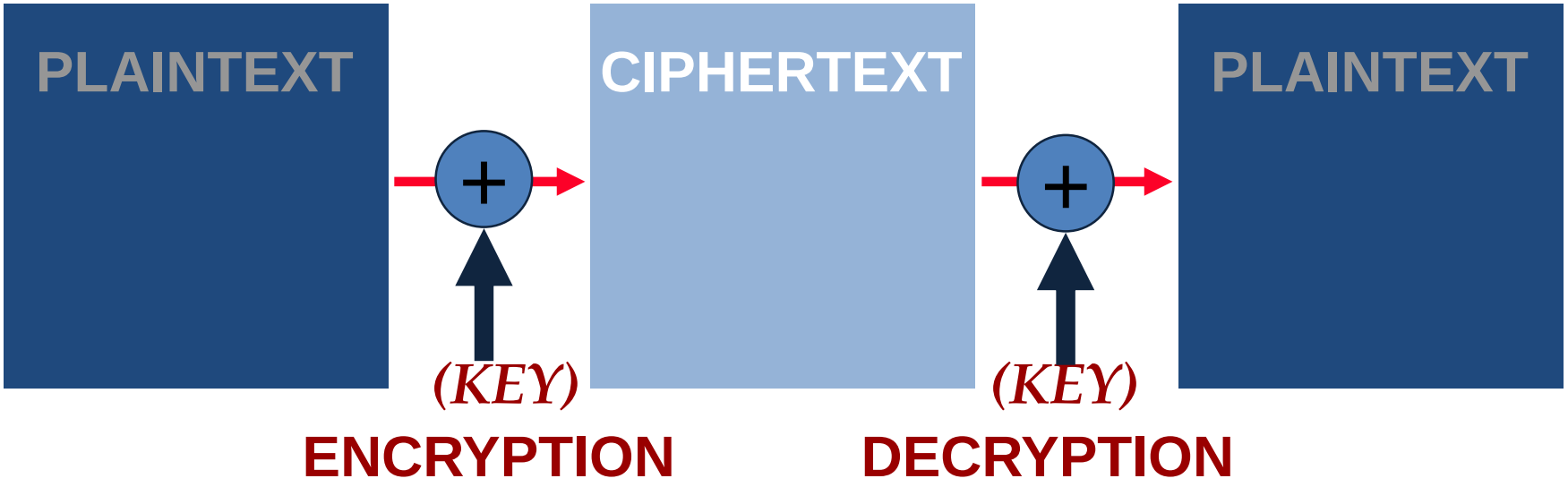
Loosely Managed Systems

- 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

Activities of Malicious Code

- 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

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

Prof. Clifford Neuman

Lecture 4
1 February 2019
12:00 Noon
OHE 100C

Course Identification

- INF 529
  - Information Privacy
  - 4.0 units
  - Website http://ccss.usc.edu/529
- Class meeting schedule
  - Noon to 3:20PM Friday's
  - Room OHE 100C
- Class communication
  - inf529@csclass.info

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

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 identification, authentication, and audit 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

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

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:

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

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.

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

Implications

http://www.networkworld.com/article/2683692/microsoft-subnet/is-swat-raid-on-wrong-house-based-on-open-wi-fi-ip-address-unconstitutional.html

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

Components of ID systems

- Collectors
  - Gather raw data
- Director
  - Reduces incoming traffic and finds relationships
- Notifier
  - Accepts data from director and takes appropriate action

Examples of Detection

The Anonymity Debate

- Should we be required to identify ourselves when using the internet?
  - What about other situations
  - Event:
  - https://www.bleepingcomputer.com/news/security/windows-drm-files-used-to-decloak-tor-browser-users/
- Authentication of Attributes vs Identity
  - Over 21, but without showing your DL
- Use of Internet Cafes

What shall you do if Compromised

What to do when your companies systems are attacked will depend on many factors, but most importantly:

What you have done to prepare.

- Emergency Response Plan
- Emergency Response Team
- How the system has been set up
  - Backups, Data Collection for Forensics, Baselines
  - Your containment architecture

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 …

After a breach: Containment

- You will need to shut down or take offline those systems affected by the breach (those you can no longer trust) to prevent further loss of data.
- Collect forensic data to assist in assessing impact, to identify attackers, and for prosection.
- Stronger containment (from preparation phase) means fewer critical services that you need to take offline.

After a breach: Notification

- California Breach Law SB-1386 November 2003 If personal information is stolen and not encrypted, the business has a fiduciary responsibility to advise every client or customer and employee their information has been compromised.
- SB-24 Addendum to SB-1386 January 2012 The business must also notify DOJ and provide particulars of how the

After a breach: analysis

- Check for changes to your systems and data.
- Use forensic data collected using technologies in place from your planning phase to identify sources of the attack, and the techniques used.
- Use this information to determine which customers data and which systems were affected.
- Use this information to fix the vulnerabilities in your existing systems.
- Subject to legal requirements, share this information as appropriate with the authorities and the security community.

After a breach: recovery

- Systems/data must be restored to a trusted state.
  - Using backups
  - Legitimate updates to data may need to be reapplied
- Vulnerably used in the attack must be patched before systems are brought back online.
- You need a plan for operating your business for some period without the impacted systems.
  - Part of your planning phase

What if I'm a small business

- Some parts of your operation are likely outsourced
- Carefully vet your providers, apply the same standards here, and ask how they address them.
- Consider using multiple sources so that you can switch from one to another in the event of a breach that it outside your control.
- Insist that they have containment architectures that better protect your information while it is in their hands.

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