

CSci 5271
Introduction to Computer Security
Day 26: Usability and security

Stephen McCamant
University of Minnesota, Computer Science & Engineering

Outline

Usability and security
Announcements intermission
Usable security example areas
AI/LLM safety and security, cont'd
DNSSEC

Users are not 'ideal components'

- Frustrates engineers: cannot give users instructions like a computer
 - Closest approximation: military
- Unrealistic expectations are bad for security

Most users are benign and sensible

- On the other hand, you can't just treat users as adversaries
 - Some level of trust is inevitable
 - Your institution is not a prison
- Also need to take advantage of user common sense and expertise
 - A resource you can't afford to pass up

Don't blame users

- "User error" can be the end of a discussion
- This is a poor excuse
- Almost any "user error" could be avoidable with better systems and procedures

Users as rational

- Economic perspective: users have goals and pursue them
 - They're just not necessarily aligned with security
- Ignoring a security practice can be rational if the rewards is greater than the risk

Perspectives from psychology

- Users become habituated to experiences and processes
 - Learn "skill" of clicking OK in dialog boxes
- Heuristic factors affect perception of risk
 - Level of control, salience of examples
- Social pressures can override security rules
 - "Social engineering" attacks

User attention is a resource

- Users have limited attention to devote to security
 - Exaggeration: treat as fixed
- If you waste attention on unimportant things, it won't be available when you need it
- Fable of the boy who cried wolf

Research: ecological validity

- User behavior with respect to security is hard to study
- Experimental settings are not like real situations
- Subjects often:
 - Have little really at stake
 - Expect experimenters will protect them
 - Do what seems socially acceptable
 - Do what they think the experimenters want

Research: deception and ethics

- Have to be very careful about ethics of experiments with human subjects
 - Enforced by institutional review systems
- When is it acceptable to deceive subjects?
 - Many security problems naturally include deception

Outline

Usability and security

Announcements intermission

Usable security example areas

AI/LLM safety and security, cont'd

DNSSEC

Presentation schedule posted

- Presentations will occur the next (last) three lectures
- Schedule posted today on Piazza
- 18 (13+5) minutes per project, with some administrative material interspersed

Outline

Usability and security

Announcements intermission

Usable security example areas

AI/LLM safety and security, cont'd

DNSSEC

Email encryption

- Technology became available with PGP in the early 90s
- Classic depressing study: "Why Johnny can't encrypt: a usability evaluation of PGP 5.0" (USENIX Security 1999)
- Still an open "challenge problem"
- Also some other non-UI difficulties: adoption, govt. policy

Phishing

- Attacker sends email appearing to come from an institution you trust
- Links to web site where you type your password, etc.
- *Spear phishing*: individually targeted, can be much more effective

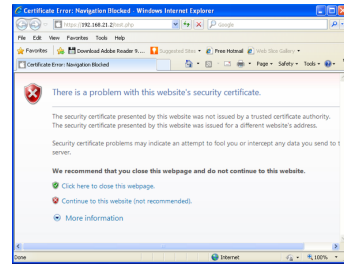
Phishing defenses

- Educate users to pay attention to X:
 - Spelling → copy from real emails
 - URL → homograph attacks
 - SSL "lock" icon → fake lock icon, or SSL-hosted attack
- Extended validation (green bar) certificates
- Phishing URL deny-lists

SSL warnings: prevalence

- Browsers will warn on SSL certificate problems
- In the wild, most are false positives
 - foo.com VS. www.foo.com
 - Recently expired
 - Technical problems with validation
 - Self-signed certificates (HA2)
- Classic warning-fatigue danger

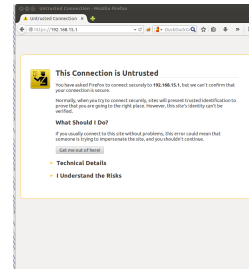
Older SSL warning



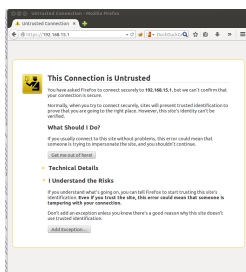
SSL warnings: effectiveness

- Early warnings fared very poorly in lab settings
- Recent browsers have a new generation of designs:
 - Harder to click through mindlessly
 - Persistent storage of exceptions
- Recent telemetry study: they work pretty well

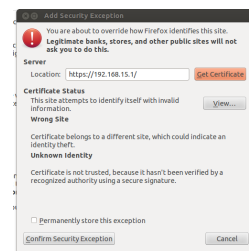
Modern Firefox warning



Modern Firefox warning (2)



Modern Firefox warning (3)



Spam-advertised purchases

- "Replica" Rolex watches, herbal V!@gr@, etc.
- This business is clearly unscrupulous; if I pay, will I get anything at all?
- Empirical answer: yes, almost always
 - Not a scam, a black market
 - Importance of credit-card bank relationships

Advance fee fraud

- "Why do Nigerian Scammers say they are from Nigeria?" (Herley, WEIS 2012)
- Short answer: false positives
 - Sending spam is cheap
 - But, luring victims is expensive
 - Scammer wants to minimize victims who respond but ultimately don't pay

Trusted UI

- Tricky to ask users to make trust decisions based on UI appearance
 - Lock icon in browser, etc.
- Attacking code can draw lookalike indicators
 - Lock favicon
 - Picture-in-picture attack

Smartphone app permissions

- Smartphone OSES have more fine-grained per-application permissions
 - Access to GPS, microphone
 - Access to address book
 - Make calls
- Phone also has more tempting targets
- Users install more apps from small providers

Permissions manifest

- Android approach: present listed of requested permissions at install time
- Can be hard question to answer hypothetically
 - Users may have hard time understanding implications
- User choices seem to put low value on privacy

Time-of-use checks

- iOS approach: for narrower set of permissions, ask on each use
- Proper context makes decisions clearer
- But, have to avoid asking about common things
- iOS app store is also more closely curated

Trusted UI for privileged actions

- Trusted UI works better when asking permission (e.g., Oakland'12)
- Say, "take picture" button in phone app
 - Requested by app
 - Drawn and interpreted by OS
 - OS well positioned to be sure click is real
- Little value to attacker in drawing fake button

Outline

Usability and security

Announcements intermission

Usable security example areas

AI/LLM safety and security, cont'd

DNSSEC

Emergent risks

- Scaling LLMs have often shown novel capabilities
 - Which ones are most concerning in amplifying AI risk?
- Planning, pursuing goals (positive applications too)
- Self-replication (e.g., compare computer worm)
- Real world influence and deception
 - Example: TaskRabbit to solve a CAPTCHA

Medium-term concerns

- Economic disruption
 - E.g., widespread job losses and unemployment
- Acceleration: positive feedback increasing the rate of AI development
 - Reckless competition towards AI goals
 - AI facilitating science and technological development

Some reasons alignment is hard

- Humans already can't agree among themselves on universal values
- Human desires have a lot of implicit side conditions and unstated restrictions
- We don't understand many details of how LLMs work internally
- If AIs become smarter than people, why would they want to obey us?

Hypothetical endpoints

- Paperclip maximizer
 - Seemingly simple goal + great capability = deeply undesirable result
- Will super-human AIs treat humans the way humans have treated non-human animals?
 - Extreme loss of agency is possible without destruction
 - Many different example animals and possible perspectives
 - Too close of an analogy may be unrealistic, since AI may be much less like us than animals are

Precaution and p(doom)

- A trending conversation topic is comparing estimates on the probability of a catastrophic outcome from AI
- Surprisingly many people working in AI have a significant p(doom)
 - Progress is inevitable, or it would be worse without me
- Choosing not to pursue technology because of downside risks is rare
 - Compare: nuclear weapons and energy

Outline

Usability and security

Announcements intermission

Usable security example areas

AI/LLM safety and security, cont'd

DNSSEC

DNS: trusted but vulnerable

- Almost every higher-level service interacts with DNS
- UDP protocol with no authentication or crypto
 - Lots of attacks possible
- Problems known for a long time, but challenge to fix compatibly

DNSSEC goals and non-goals

- + Authenticity of positive replies
- + Authenticity of negative replies
- + Integrity
- Confidentiality
- Availability

First cut: signatures and certificates

- Each resource record gets an RRSIG signature
 - E.g., A record for one name→address mapping
 - Observe: signature often larger than data
- Signature validation keys in DNSKEY RRs
- Recursive chain up to the root (or other "anchor")

Add more indirection

- DNS needs to scale to very large flat domains like .com
- Facilitated by having single DS RR in parent indicating delegation
- Chain to root now includes DSEs as well

Negative answers

- Also don't want attackers to spoof non-existence
 - Gratuitous denial of service, force fallback, etc.
- But don't want to sign "x does not exist" for all x
- Solution 1, NSEC: "there is no name between acacia and baobab"

Preventing zone enumeration

- Many domains would not like people enumerating all their entries
- DNS is public, but "not that public"
- Unfortunately NSEC makes this trivial
- Compromise: NSEC3 uses password-like salt and repeated hash, allows opt-out

DANE: linking TLS to DNSSEC

- "DNS-based Authentication of Named Entities"
- DNS contains hash of TLS cert, don't need CAs
- How is DNSSEC's tree of certs better than TLS's?

Signing the root

- Political problem: many already distrust US-centered nature of DNS infrastructure
- Practical problem: must be very secure with no single point of failure
- Finally accomplished in 2010
 - Solution involves 'key ceremonies', international committees, smart cards, safe deposit boxes, etc.

Deployment

- Standard deployment problem: all cost and no benefit to being first mover
- Servers working on it, mostly top-down
- Clients: still less than 20%
- Will probably be common for a while: insecure connection to secure resolver

What about privacy?

- Users increasingly want privacy for their DNS queries as well
- Older DNSCurve and DNSCrypt protocols were not standardized
- More recent "DNS over TLS" and "DNS over HTTPS" are RFCs
- DNS over HTTPS in major browsers might have serious centralization effects