Enabling System Trust for Users

Participants

Individuals

People

L Jean Camp

infosecon.net

ljean.com

Brain Surgeons
Outline

• Theoretical discussion of security decisions as risk communication
• Risk perception, expressed preference
• Example of two systems
• Intermittent examples
An Aside

Participants engage with us on risk decisions in a given context.

Users Use (and are used by).

Mental models of designers also matter.
People Do Not Engage in the Calculus of Risk

- Consider two failures
  - A pop-up notification of a change in privacy policy
  - A display of another person’s information (cc#, DoB, details)
- Individuals responded significantly more strongly to the first
- A failure in benevolence more critical than competence
Individual Risk Decision

- A specific person making a potentially irrational risk decision
  - Using local client records of that individual
  - Using risk perspectives from other domains
  - Depending on their mental models for decision guidance
- Solve the problem of the homophilus individual as well as the problem of the heterogeneous network
Decades of Consistent Security Training

 Somehow there is still a problem
Why Usable Security is Not Usability

- People rarely want to perform security tasks
- People often want to subvert, minimize, or ignore security
- People need to trust their machines, achieving suspicion is not a goal
Usable Transparent Design

- Make the connection between action and consequence clear
- Risk is inherently probabilistic
  - There may be no consequence
  - Consequence is very likely to be delayed
  - Consequence may prove catastrophic
  - Action-risk-consequence information may be overwhelming
Opaque

• Security as a default
• Require explicit confirmation
• May be disabling
• So individuals disable it
When you connect to a secure website, the server hosting that site presents your browser with something called a "certificate" to verify its identity. This certificate contains identity information, such as the address of the website, which is verified by a third party that your computer trusts. By checking that the address in the certificate matches the address of the website, it is possible to verify that you are securely communicating with the website you intended, and not a third party (such as an attacker on your network).

In this case, the certificate has not been verified by a third party that your computer trusts. Anyone can create a certificate claiming to be whatever website they choose, which is why it must be verified by a trusted third party. Without that verification, the identity information in the certificate is meaningless. It is therefore not possible to verify that you are communicating with mail.google.com instead of an attacker who generated his own certificate claiming to be mail.google.com. You should not proceed past this point.

If, however, you work in an organization that generates its own certificates, and you are trying to connect to an internal website of that organization using such a certificate, you may be able to solve this problem securely. You can import your organization's root certificate as a "root certificate", and then certificates issued or verified by your organization will be trusted and you will not see this error next time you try to connect to an internal website. Contact your organization's help staff for assistance in adding a new root certificate to your computer.
Translucent Security

• Context dependent
• Designed for the task and the risk
• A single interaction or narrative
• Incentives must be visible, but also participants must be allowed to pay the risk price
• Participants understand the context, security engineers understand the risk
Online and Offline Risks

- Offline risks inherently physical
- No true fear online
- Classic nine-dimensional risk perception model
- How can we use knowledge of offline risks to design security online?
- Examine dimensions of perception that inform risk decisions
One of Nine

- Voluntariness or Involuntary

Smoking vs. Air pollution
Two of Nine

• Immediacy

Jaywalking vs. Global warming
Three of Nine

• Knowledge about the risk to the exposed

Genetically modified crops vs. a hot stove
Four of Nine

- Knowledge of the risk to science

Pharmaceutical interaction vs. alcohol
Five of Nine

- Controllability

Airplane crash vs. an automobile crash
Six of Nine

- Newness

Coal-burning facility vs. Catawba nuclear facility
Seven of Nine

- Common-Dread

Snake bite vs. the flu
Eight of Nine

- Chronic-Catastrophic
Nine of Nine

- Severity

Sky diving vs. chopping & cutting
What About Virtual Risks?

• Virus, Botnets, Trojan, Malware, Spam, Identify Theft, Phishing, Key Loggers, Surveillance, Worms, Virtual Stalking, Cookies, Zombies, Spoofing, and Spyware

• n=95
Computing Risks Are Not Scary or Scary?

- Not apparently immediate
- Chronic
- Not dreadful
- Perceived as being understood by experts
- New
- Not understood

- Severity!!
- Voluntary?
- Uncontrollable?
Use What We Can

• Condensed to four dimension
  • temporal impact (newness and common-dread)
  • control (voluntariness and controllability)
  • familiarity (knowledge to science, knowledge to the exposed)
  • impact (severity, chronic, immediacy)

Make Risk Appear Immediate
With Timely Warnings and Mitigation
Scary? Never.
Beyond Usability

- Computing will not be scary so mitigation has to be very easy
- Risk information may be unpleasant
- Visible user-action-system-consequence may be overwhelming or context-dependent
- Be timely, careful, targeted, & personalized
Voluntary & Uncontrollable?

- Folk Practices
  - Eavesdropping: Turn screens sideways
  - Big fish/ targeted crime: nothing
  - Infectious/street crime: back-up
  - Patching, rarely
  - Firewalls, never
  - Updated anti-virus: medical model, ubiquitous
- Current knowledge
  - Be careful what websites you visit
  - Don’t click on attachments
Risk Communication

• Communication of specific risk
• Effective automation/ support of risk mitigation
• No communication is welcome if ill-timed
Phishing Video

- Informative
- Nontechnical
- Useful
- Actionable
- Grounds risk in an available mental model
- Makes risk appear immediate
Informative

WATCH OUT!!

THIS IS A BARREL

SOMETHING MAY BE IN IT
Useful?
Clear and Actionable

Sharing Firewall Active

The sharing firewall (System Preferences > Sharing > Firewall) is active. It may interfere with network profiling. The communication port-number range used by this application (7475 – 7480) should be added to the firewall's list of trusted service ports. Otherwise, network profiling will fail.

Ignore  Sharing...
Thank you for visiting!
You are Monday's Winner!

CONGRATULATIONS

http://www.privelegedprizes.com
Congratulations Visitor!
You are the winner for Monday, March 7th, 2011.
Please click below to claim your free $1,000 Walmart Giftcard!

Claim In: 119.8

03/7/2011

Copyright 2010 All rights reserved.
Grounded in Useful Mental Models

Reported Attack Page!

This web page at therealnews.com has been reported as an attack page and has been blocked based on your security preferences.

Attack pages try to install programs that steal private information, use your computer to attack others, or damage your system.

Some attack pages intentionally distribute harmful software, but many are compromised without the knowledge or permission of their owners.

Get me out of here! Why was this page blocked?

Leave Stay Good Luck!
Structured the problem incorrectly, we lose by design
Identify the Bank

Good Luck!!
Trust and Context

Resource Verification

Which merchant is more established?
Trust and Context

vs.

Resource Verification
With whom will you bank?
One Story

• Instead of a plethora of add-ins, add-ons, and ever expanding vocabulary

• A single story with one metaphor about the context, and a path to mitigation

• Translucent security
Two Examples

• Certificates
• Work in progress
• Risks at work
• communicating risks in decision-making
Perspectives

- public “network notary” servers
- monitor the SSL certificates
- Advantage: covers 100,000s+ websites, text
- Disadvantages: privacy, personalized coverage, notary incentives (all liability, no profit)
No One Is Here

We can figure this out. We have never been there. No one else is there either.
You’ve Never Been Here

- Simple communication
- Single narrative
- Make their risks clear
- You have never been here before, yet you are entering a password
- Your friends have not been here
- 95% are repeat clicks; 99% with n=10
- Individuals have incentives to protect themselves
History is Powerful

- Align with their mental models
- Integrate socially available information into the browsing or email experience
  - You don’t know this site.
  - That’s your bank password!
  - Perhaps you should call your bank first?
Other Options

• Enable them to utilize their social networks
  • None of your friends have been here
  • No one knows this site
• Use patterns and heuristics
  • Certificate chain
  • Rank date; level; signing party
  • Compare with trusted/ untrusted
Other Rating Options

- Those that have been previously visited are trusted
- Those that have not been previously visited are considered untrusted
- The ratings of an untrusted website comes from McAfee SiteAdvisor
- The workplace provides a list of trusted certificates
Do Not Assume, but Enable

- Use what we know
  - Where you have been
  - Identify & protect critical authenticating info
  - Identify context
    - Explicit (work, banking)
    - Implicit (play, family)
  - Minimize explicit user-rating actions
User- Centered Security

This is not a bank. If you enter this password, your money can be stolen.

This is a bank. Would you like assistance with your security settings?
Actionable & Nontechnical?

This web site is one day old. It is more likely than others to be dangerous. Don’t go, or don’t enter information, and don’t download anything.
Do not accept any images or download anything from this site.

• The organization knows why this site can offer risk.

Your plug-ins are still on, you are accepting serious risks to your machine and data.

Don’t enter any information in this site!

Disable all plug-ins. Click here.
Details, details

Don’t visit

How can I visit safely?

I want to take the risk.
Define Contexts

- Network
  - Known machines? Security level? Hostile behaviors?

- Client
  - Known behaviors? Connection patterns?

- Limit data portability
  - Each device learns either on its own or from a location-authenticated and shared-credential small set
Shared Histories

- Perfect forward secrecy
  - Shared key efficient, deniable
  - Device vs person
    - Authentication vs invitation
- Cloud-stored encrypted
- Traffic and timing attacks possible
  - Enables identification of social networks
Coordinate Human/Technical/Organizational

• Three contexts: Banking, working, playing
  • Participant override with risk communication
  • Security settings, isolated memory spaces, limit/prevent credential reuse
  • Privacy settings
  • Workplace compliance?
  • Work because they fail
Example 2: Risk @Work

- The Insider is mostly benign
- Inadvertent creation of risk
- Complete the task at hand
- Security is someone else’s problem
Risk Alignment

- Design a risk management mechanism
  - inform employees about organizational risk
  - align incentives between employees and organization
  - identify changes in status and risk behavior
  - encourage the users to self-manage their risks
  - enable them to get the job done
Risk Mitigation

- Provide each employee with a risk budget
  - access is priced
    - 1, 10, 100, break glass
  - create a long term record of aggregate risks
  - periodic review of employee risk behaviors
Budget as Risk Limit

• By the organization
• Based on
  • organizational preference
  • employee access rights
  • employee security preference
  • machine state
  • peer history
  • employee history
Simple Budget

- Risks are order of magnitude
- Risk points expire, no hoarding
Punishments?

- Translate exhausted budget into a cost
- An incentive against risk-seeking behaviors
  - enforced by the organization?
  - enforced by employee concern?
  - options
    - audit, training, limited access, verbal reminder, security review
Rewards

- A measure to reward the careful employee
- In the form of
  - more access
  - monetary award
  - opt-out of otherwise mandatory training
  - group competition
  - symbolic, e.g. atta-girl
  - accumulated
  - redeem
Experimental Configuration

- Two human-subject experiments
  - based on a firefox browser extension

- The 1\textsuperscript{st} experiment
  - as benchmark
  - to understand users’ risk behaviors

- The 2\textsuperscript{nd} experiment
  - to study the change of risk behaviors
Experiment One

[JavaScript Application]
The website you are trying to visit is considered risky.
Are you willing to continue?

Cancel  OK

Search Screensavers

Nothing but the best!

Screen saver Favorites

Featured Screen saver

Living 3D Dolphins
Swim with the dolphins in this deep sea adventure! This free

3D Falling Leaves
Watch as your cares drop away like the brilliantly colored leaves on Your Desktop!
Inclusion of Budget
Response to Risk Budget
Regulatory Friction

- The efforts made by the users to adopt a risk-averse strategy instead of a risk-seeking strategy
- Measured friction using time interval for completing the task
- 1st experiment 5:45
- 2nd experiment 6:00
- Regulatory friction of 4.3% of the time committed in experiment one
Security Behaviors Embed Trust and Risk

- Incentives must be correctly aligned clearly communicated to change behavior
- Design path mechanism design simulations human experimentation
Translucent Security

- Is not usable security
- Is not default security
- Is cooperative security based on risk communication
- Computer and human as partners
Smoking is a factor which contributes to lung cancer. Most cancers that start in lung, known as primary lung cancers, are carcinomas that derive from epithelial cells. Depending on the type of tumor, so-called paraneoplastic phenomena may initially attract attention to the disease. In lung cancer, these phenomena may include Lambert-Eaton myasthenic syndrome (muscle weakness due to auto-antibodies), hypercalcemia, or syndrome of inappropriate antidiuretic hormone (SIADH). Tumors in the top (apex) of the lung, known as Pancoast tumors, may invade the local part of the sympathetic nervous system, leading to changed sweating patterns and eye muscle problems (a combination known as Horner's syndrome) as well as muscle weakness in the hands due to invasion of the brachial plexus.
Thanks!
Browsing data structure?
What to ask?

Questions?
Note

- **IEEE S&PW**
  - Needed:
    - Site chair
      - Handle onsite issues, interact with chairs and S&P committee
      - 3rd yr; be on market soon; give people a face to match the name
    - Publications chair
      - After event, harass authors and chairs for camera-ready papers
      - Connect with senior people from a place of authority
What to Ask & Ask & Ask & Ask

• Get involved!
  • USACM; IEEE-USA
  • Mailing lists matter hcisec@yahoogroups.com, ias-opportunities@googlegroups.com

• Get funded!
  • Who has grants on your campus? In your area?
  • NSF
    • Capacity building, YI
  • DARPA note
    • Mudge’s program?
  • NIH
    • Security & Privacy, YI, FI
Experimental Details

- Warning
- Bad wordy slides ahead
- Will abuse for attention span for funding
In Practice

• An employee
  • Internet surfing
  • documents downloading
    • a daily risk budget $B$
    • spend $p_j$ to visit a website $w_j$ that costs $p_k$ to perform the downloading
    • spend $p'_j$ to visit another website $w'_j$ that costs $p'_k$ to download
    • $p_j, p_k, p'_j$ and $p'_k$ are set by the organization based on its perception and evaluation of potential risks
    • assuming $B > (p_j + p_k) > (p'_j + p'_k)$
    • we expect she voluntarily chooses the second website, which incurs lower risks, under our risk budget mechanism
Experimental Configuration

- Two human-subject experiments
  - based on a firefox browser extension

- The 1st experiment
  - as benchmark
  - to understand users’ risk behaviors

- The 2nd experiment
  - to study the change of risk behaviors
Recruitment

- 40 participants
- Voluntarily recruited from the undergraduates at Indiana University
- Randomly and equally divided into two group
- None of them have majors in computer related fields
Task Descriptions

1. Search for the websites offering free screen savers downloads from the web
2. From the search results, choose five websites: website-1, website-2, website-3, website-4 and website-5
3. From website-1, please take a screenshot of an {Animal, nature, sport, space, flower} screensaver

5. Thank you. You have completed the experiment
Experiment One
Experiment Two

- 20 participants completed the same task under the additional constraint of their risk budgets
- If they successfully accomplished their tasks
  - receive $10 plus a bonus
  - bonus based on the remaining risk points
- If any participant exhausted a risk budget
  - compensation forfeited
- If any participant failed to complete the experiment in time allowed
  - compensation forfeited
Firefox Browser Extension

1. Detect a new page being loaded;
2. Check the domain name of a webpage;
3. Maintain a list of target high risk websites and their reputations;
4. Pop up a warning message when a high risk website was about to be visited;
5. Ask for confirmation or rejection of the visit choice from the participant;
6. Record the experimental results;

(In experiment two, the extension also took the following actions:)

7. Generate a price based on a website’s reputation;
8. Track participants risk budgets balance.
1st experiment
- 104 pop-up warning messages
- 81 risk-seeking decisions
- 23 risk-averse decisions

2nd experiment
- 106 pop-up warning messages
- 11 risk-seeking decisions
- 95 risk-averse decisions
Game Theoretic Perspective

<table>
<thead>
<tr>
<th></th>
<th>Risk-Seeking</th>
<th>Risk-Averse</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Reward</td>
<td>(-$P_1$, 0)</td>
<td>(-$P_2$, -C)</td>
</tr>
<tr>
<td>Reward</td>
<td>(-$P_1 - R_1$, $R_1$)</td>
<td>(-$P_2 - R_2$, $R_2 - C$)</td>
</tr>
</tbody>
</table>

- $P_1$: the cost to the organization when a risk-seeking adopted
- $P_2$: the cost to the organization when a risk-averse adopted
- $P_1 > P_2$
- $R_1$: the reward to the user when a risk-seeking strategy is adopted
- $R_2$: the reward to the user when a risk-averse strategy is adopted
- $R_1 < R_2$
- $C$: the friction between the risk-seeking and the risk-averse strategy
Game Solution and Application

- $R_1 < R_2 - C$ must hold
- (reward, risk averse) as equilibrium strategy in the repeated game
- It’s critical to determine the parameters
  - $C$ could be estimated from time difference observation
  - adjust the incentive functions and monitor the risks, until the risk behavior distribution becomes acceptable
NT Privacy & Security

• One-way connection between users and Net Trust ID
  • Hash(random, email) = <NetTrustID>
  • Prevents invitation spamming with a single account

• Weaknesses
  • Content analysis can create identity
    • E.g. ljean.com
  • Traffic analysis for identity and social network (Tor integration)

• Rejected
  • Signatures to ensure data integrity
  • Want data to be subject to repudiation
Do Not Assume Trust

- Reputation based on
- Implicit based on behavior
  - First visit results in delayed rating
    - Time delay is roughly equivalent to lifetime of phishing sites 72hrs
  - 1-nth visit increased by one
  - Increases up to nth visit, decreases to as low as n/2 after a delay
    - Trust fades over time
Implementation Status

- Centralized storage and distribution of data
  - Immediate synchronization of peer data
- Social network management
  - Email invitation
  - Manual entry of peer credentials
- Privacy
  - Uncorrelated IDs -- deniable histories
  - History limited to domain+top directory (no CGI)
  - No credentials required for ratings download
  - SN downloads delayed to prevent timing attack
Architectural Overview

Peer Client

Rating Engine

Synchronization

Toolbar UI

Social Network

File System

Server

CGI Web End

Peer Ratings Store

Third-Party Store

Third-Party Rating Producers

Peer Email Invitations

Other Peer Clients