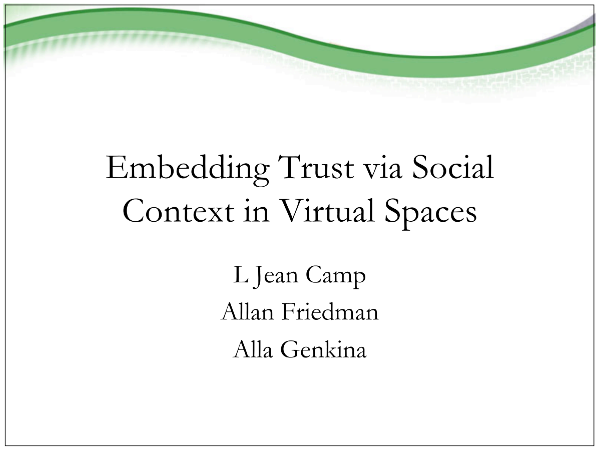
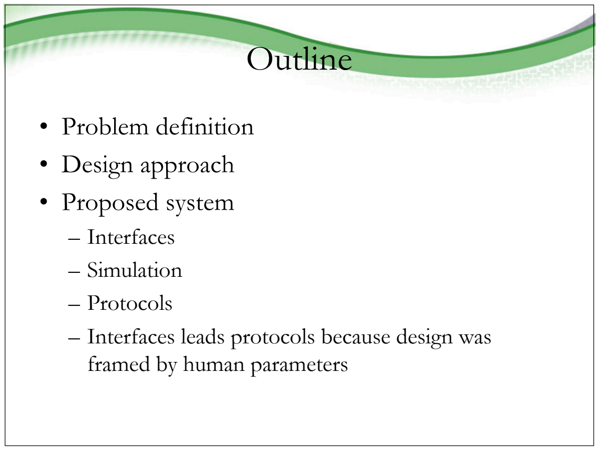


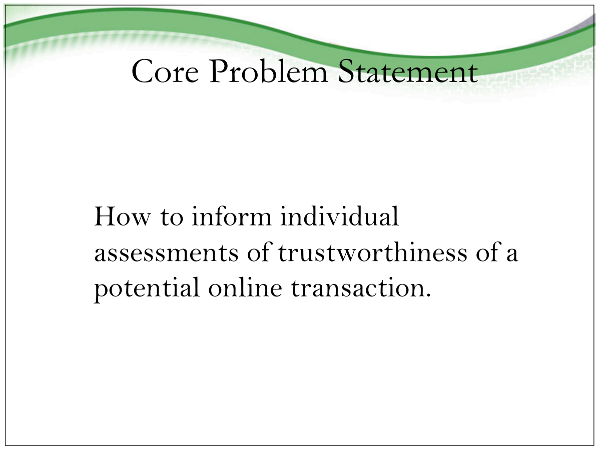
Embedding Trust via Social Context in Virtual Spaces

L Jean Camp
Allan Friedman
Alla Genkina



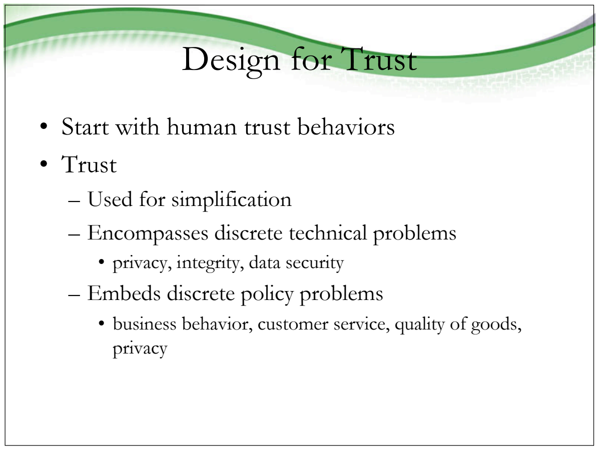
Outline

- Problem definition
- Design approach
- Proposed system
  - Interfaces
  - Simulation
  - Protocols
  - Interfaces leads protocols because design was framed by human parameters



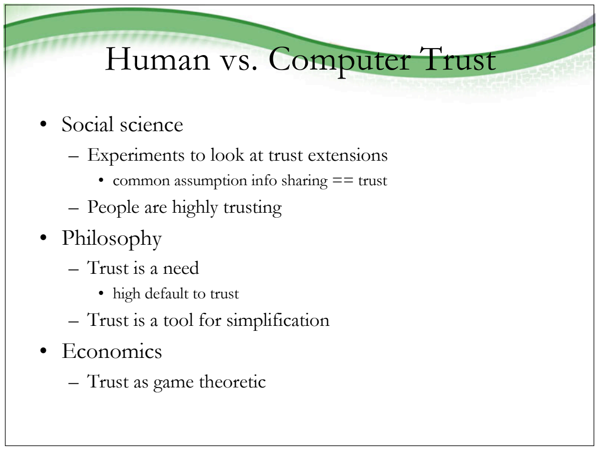
Core Problem Statement

How to inform individual assessments of trustworthiness of a potential online transaction.



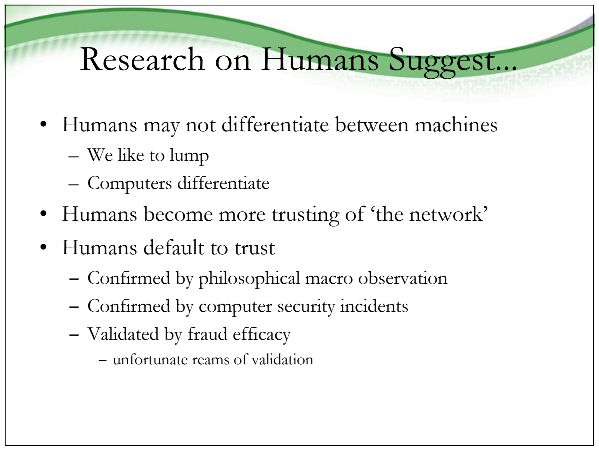
Design for Trust

- Start with human trust behaviors
- Trust
  - Used for simplification
  - Encompasses discrete technical problems
    - privacy, integrity, data security
  - Embeds discrete policy problems
    - business behavior, customer service, quality of goods, privacy



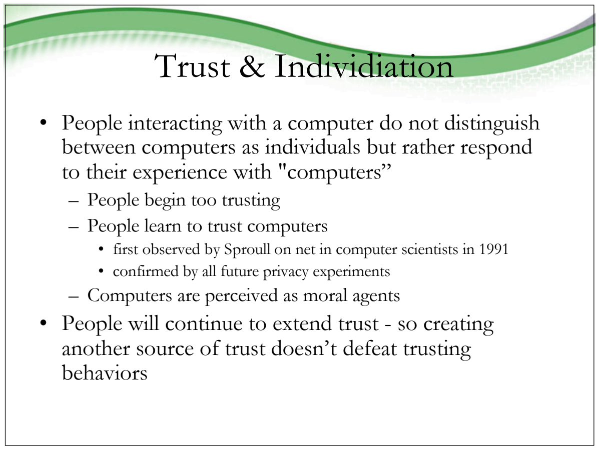
Human vs. Computer Trust

- Social science
  - Experiments to look at trust extensions
    - common assumption info sharing == trust
  - People are highly trusting
- Philosophy
  - Trust is a need
    - high default to trust
  - Trust is a tool for simplification
- Economics
  - Trust as game theoretic



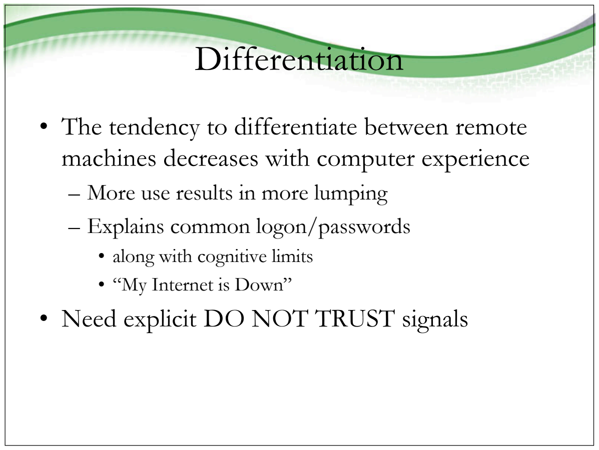
Research on Humans Suggest...

- Humans may not differentiate between machines
  - We like to lump
  - Computers differentiate
- Humans become more trusting of ‘the network’
- Humans default to trust
  - Confirmed by philosophical macro observation
  - Confirmed by computer security incidents
  - Validated by fraud efficacy
    - unfortunate reams of validation



Trust & Individiation

- People interacting with a computer do not distinguish between computers as individuals but rather respond to their experience with "computers”
  - People begin too trusting
  - People learn to trust computers
    - first observed by Sproull on net in computer scientists in 1991
    - confirmed by all future privacy experiments
  - Computers are perceived as moral agents
- People will continue to extend trust - so creating another source of trust doesn’t defeat trusting behaviors



Differentiation

- The tendency to differentiate between remote machines decreases with computer experience
  - More use results in more lumping
  - Explains common logon/passwords
    - along with cognitive limits
    - “My Internet is Down”
- Need explicit DO NOT TRUST signals

Observations

- Users are bad security managers
  - PGP, P3P, passwords, ….
- Security should necessarily be a default
- Surveys illustrate a continuing confusion of privacy & security
  - educate All Net Users
  - build upon this

Computer Security is Built for Machines

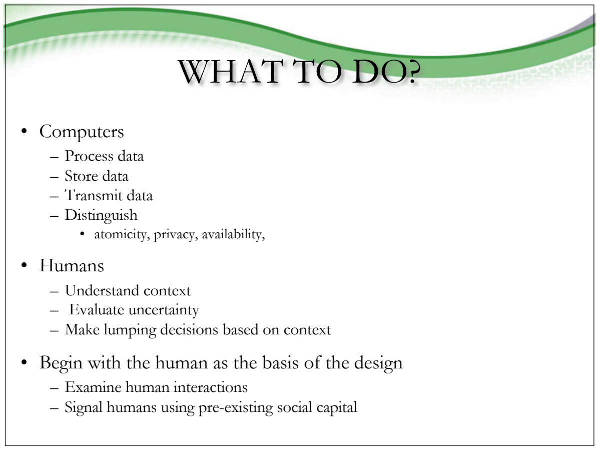
- Passwords
  - Humans are a bad source of entropy
- SSL
  - Two categories: secure and not secure
  - Does not encourage individiation
  - Computer security should seek to differentiate machines
    - Every site should include a unique graphic with the lock

Privacy standards are built for machines

- P3P assumes
  - All merchants trustworthy w.r.t. their own policies
  - Assumes increasingly sophisticated user
    - e.g., preference expression and negotiation
  - One standard for all transactions
- PGP
  - Monotonic increase in trust
  - No reset
  - No decrease in rate of trust extension
    - to compensate for increasing trust
  - No global or local reset
    - e.g. change in status

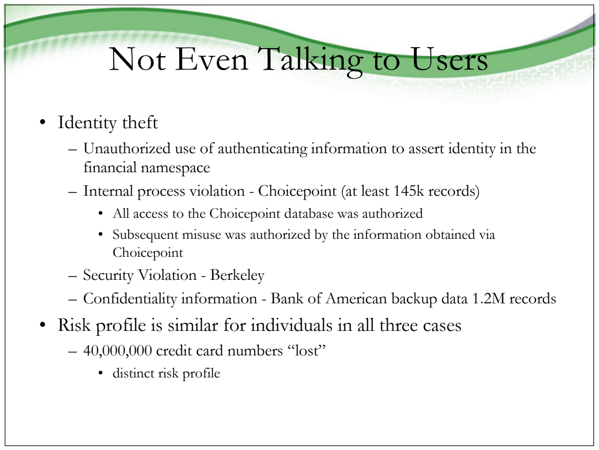
Key revocation is built for Machines

- CRL tend to be single level
- Different levels of revocation are needed
  - Falsified initial credential
    - all past transactions suspect
  - Change in status
    - future transactions prohibited
  - Refusal of renewal
    - current systems adequate
- CRL should reflect the entire systems in which they work, including the social system



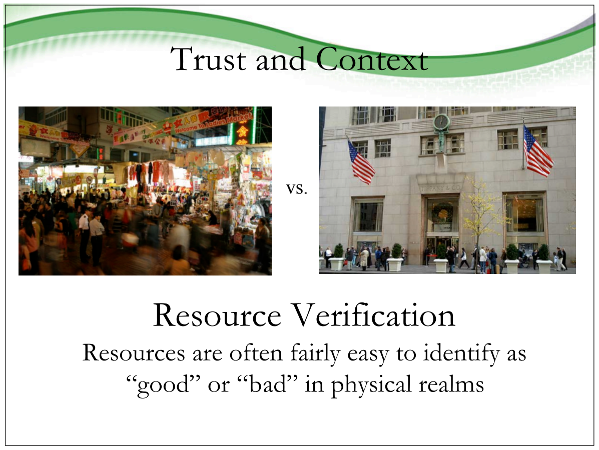
WHAT TO DO?

- Computers
  - Process data
  - Store data
  - Transmit data
  - Distinguish
    - atomicity, privacy, availability,
- Humans
  - Understand context
  - Evaluate uncertainty
  - Make lumping decisions based on context
- Begin with the human as the basis of the design
  - Examine human interactions
  - Signal humans using pre-existing social capital



Not Even Talking to Users

- Identity theft
  - Unauthorized use of authenticating information to assert identity in the financial namespace
  - Internal process violation - Choicepoint (at least 145k records)
    - All access to the Choicepoint database was authorized
    - Subsequent misuse was authorized by the information obtained via Choicepoint
  - Security Violation - Berkeley
  - Confidentiality information - Bank of American backup data 1.2M records
- Risk profile is similar for individuals in all three cases
  - 40,000,000 credit card numbers "lost"
    - distinct risk profile

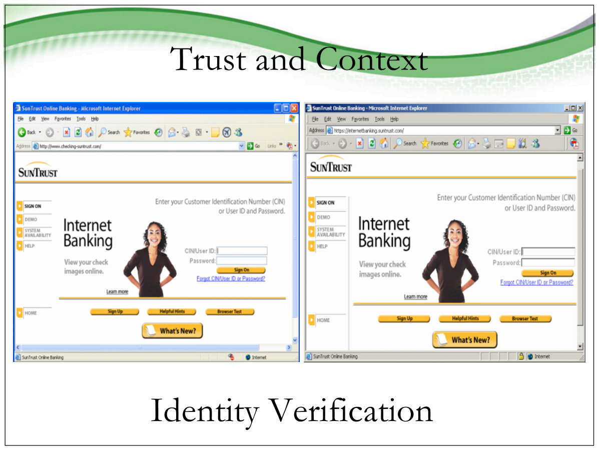


Trust and Context

vs.

Resource Verification

Resources are often fairly easy to identify as "good" or "bad" in physical realms

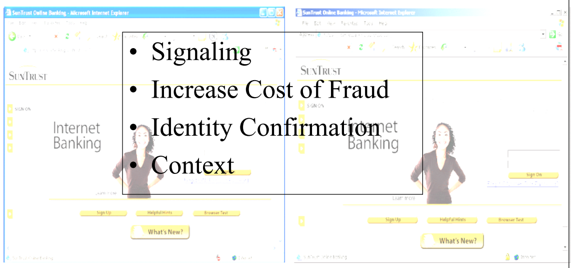


Trust and Context

SunTrust Online Banking - Microsoft Internet Explorer
SUNTRUST
Internet Banking
View your check images online.
Enter your Customer Identification Number (CIN) or User ID and Password.
CIN/User ID:
Password:
Sign On
Forgot CIN/User ID or Password?
Sign Up
Helpful Hints
Browser Test
What's New?

Identity Verification

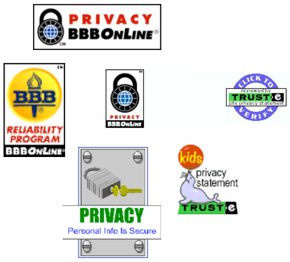
Possible Solutions



- Signaling
- Increase Cost of Fraud
- Identity Confirmation
- Context

Contextual information is needed for trust to reduce complexity.

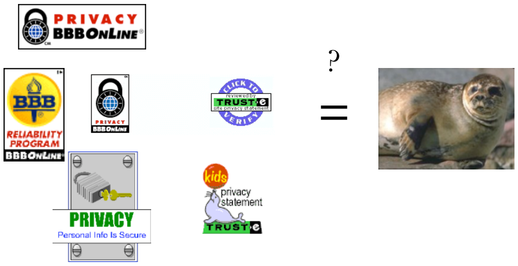
Signaling


PRIVACY BBBOnLine
BBB RELIABILITY PROGRAM BBBOnLine
PRIVACY BBBOnLine
CLICK TO VERIFY TRUSTe
PRIVACY Personal Info Is Secure
kids privacy statement TRUSTe

Seals

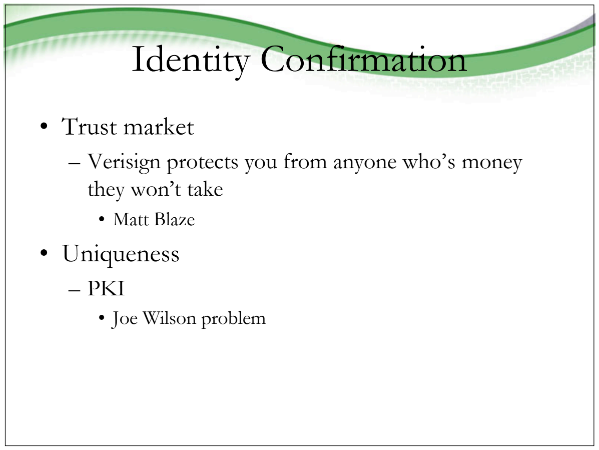
Traditional mechanisms to communicate trustworthiness.

Signaling Requires Malicious Party to Cooperate


PRIVACY BBBOnLine
BBB RELIABILITY PROGRAM BBBOnLine
PRIVACY BBBOnLine
CLICK TO VERIFY TRUSTe
PRIVACY Personal Info Is Secure
kids privacy statement TRUSTe
?
=

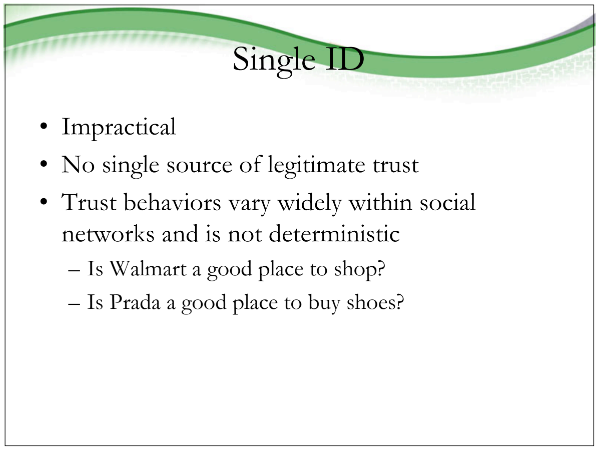
Economics Requires Payment

- Reliable payment
- New opportunities for fraud
- Perversely increase incentive?
- Setting uniform prices or negotiating prices
  - Price discrimination is the opposite of privacy
- Fraud management in a new realm



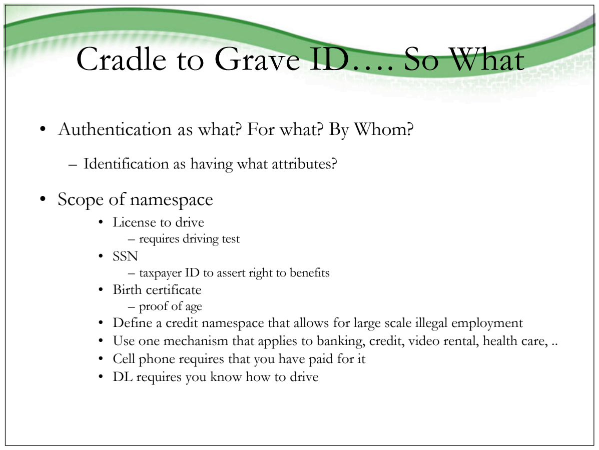
Identity Confirmation

- Trust market
  - Verisign protects you from anyone who’s money they won’t take
    - Matt Blaze
- Uniqueness
  - PKI
    - Joe Wilson problem



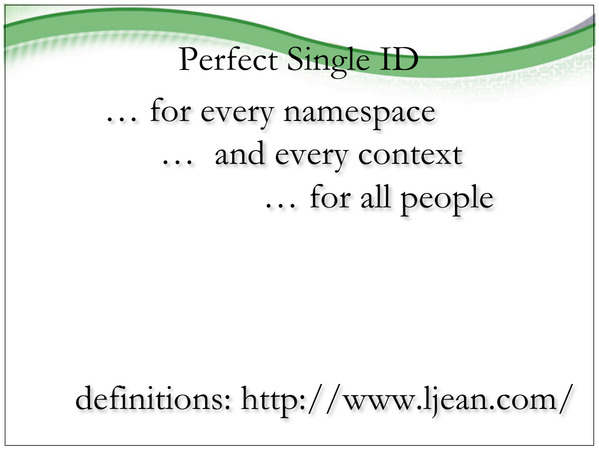
Single ID

- Impractical
- No single source of legitimate trust
- Trust behaviors vary widely within social networks and is not deterministic
  - Is Walmart a good place to shop?
  - Is Prada a good place to buy shoes?



Cradle to Grave ID…. So What

- Authentication as what? For what? By Whom?
  - Identification as having what attributes?
- Scope of namespace
  - License to drive
    - requires driving test
  - SSN
    - taxpayer ID to assert right to benefits
  - Birth certificate
    - proof of age
  - Define a credit namespace that allows for large scale illegal employment
  - Use one mechanism that applies to banking, credit, video rental, health care, ..
  - Cell phone requires that you have paid for it
  - DL requires you know how to drive



Perfect Single ID

… for every namespace

… and every context

… for all people

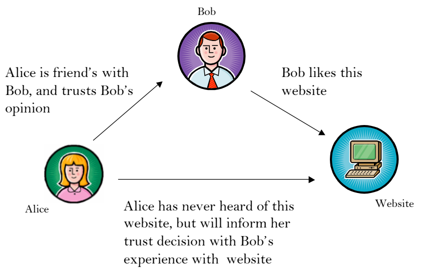
definitions: http://www.ljean.com/

… or solve the problem at hand by enabling contextually rational trust behavior

Embedding Browsing in Social Context

- First trust challenge
  - Enabling trust to allow entry onto the net
    - pre {84, 87, 91} the Net was a "trust environment" because of entry barriers
  - Enabling monetary flows
- Second trust challenge
  - Providing meaning trust information
    - TrustE, BBB, Verisign
  - Namespaces for specific trust assertions
    - Christian, gay friendly, responsible merchants
  - Requires a common understanding of the limits of the namespaces
    - Transitivity
    - Automated trust decisions
    - Consistency across contexts or explicit definition of context
      - E.g., purchase a book
        - » On divorce; impotency; effective job searching; number theory

Embedding Trust via Context



Alice is friend's with Bob, and trusts Bob's opinion

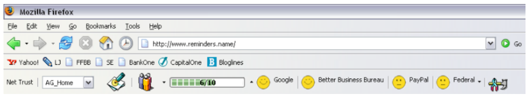
Bob likes this website

Alice has never heard of this website, but will inform her trust decision with Bob's experience with website

Net Trust View

Using a user's **social network** (known as a buddy list) as well as user-selected **centralized authorities** (known as broadcasters) the Net Trust system displays meaningful information to the user so they can make an educated decision about the trustworthiness of a website.

The Net Trust Toolbar


Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://www.reminders.name/
Go
Yahoo! LJ FFBB SE BankOne CapitalOne Bloglines
Net Trust
AG_Home
6/10
Google
Better Business Bureau
PayPal
Federal

Usability

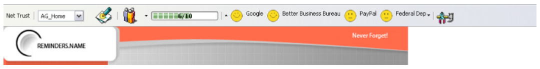
Initial Usability Testing

- **25 Students**
- **Undergraduates/Graduates**
- **Informatics Department, Indiana University**

First Results: Reminders.name

Without Toolbar: 60% say they do not trust this site

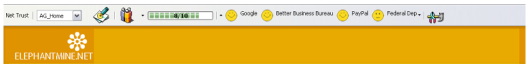
With Toolbar: 42% say they do not trust this site



Second: Elephantmine.net

Without Toolbar: 52% say they do not trust this site

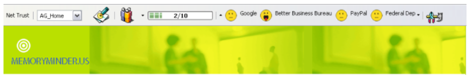
With Toolbar: 24% say they do not trust this site



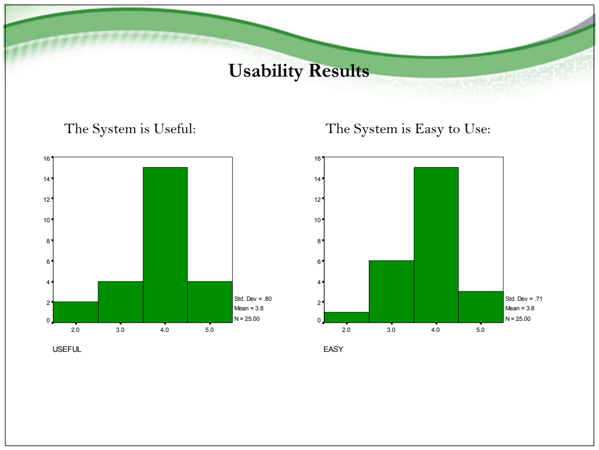
Third: MemoryMinders.us

Without Toolbar: 80% say they do not trust this site

With Toolbar: 76% say they do not trust this site



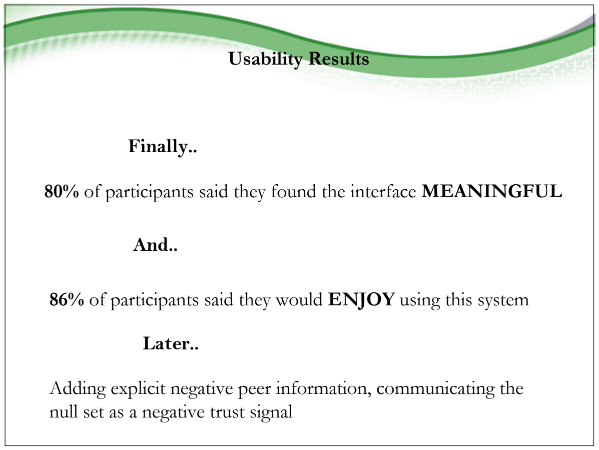
Notice that positive peer feedback dominates the negative feedback from the Better Business Bureau



Usability Results

The System is Useful:

The System is Easy to Use:



Usability Results

Finally..

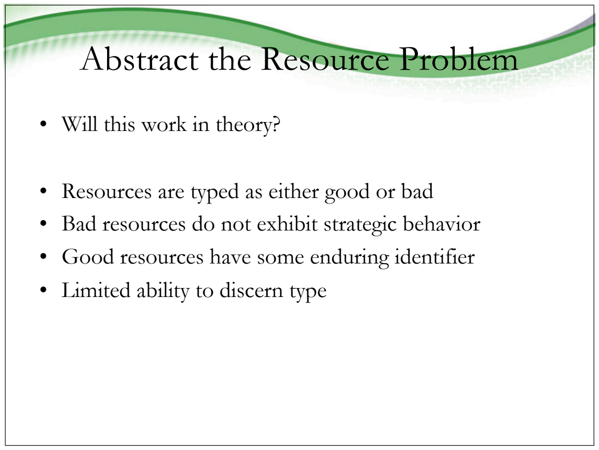
80% of participants said they found the interface **MEANINGFUL**

And..

86% of participants said they would **ENJOY** using this system

Later..

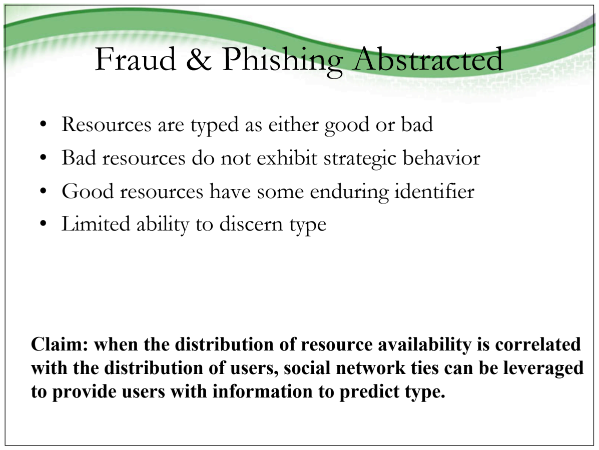
Adding explicit negative peer information, communicating the null set as a negative trust signal



Abstract the Resource Problem

- Will this work in theory?

- Resources are typed as either good or bad
- Bad resources do not exhibit strategic behavior
- Good resources have some enduring identifier
- Limited ability to discern type



Fraud & Phishing Abstracted

- Resources are typed as either good or bad
- Bad resources do not exhibit strategic behavior
- Good resources have some enduring identifier
- Limited ability to discern type

Claim: when the distribution of resource availability is correlated with the distribution of users, social network ties can be leveraged to provide users with information to predict type.

Simulation Summary

- Very simple model of networked actors deciding whether or not to visit resource
  - Network: extend Jin, Girvan & Newman (2000) to include homophily
- Decision rule: a function of number of neighbors who have also visited that resource

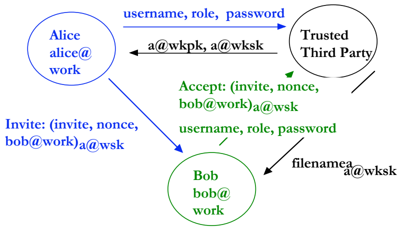
Simulation Results

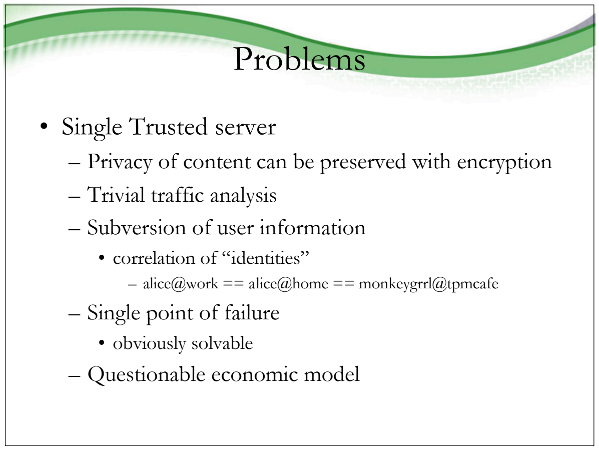
- Under basic conditions, networked actors are very good at rejecting bad resources without avoiding good resources.
- A mechanism is needed to seed the network with good information.
- The network amplifies the power of individual detection abilities.
- Temporal signatures of bad resources (phishing) can be detected.
- BUT: non-savvy actors cannot achieve perfect (95%+) results without exogenous information sources.

The Theory is Good

- Complies and leverages human trust behaviors
- Simulations suggest potential value
- How to build it?

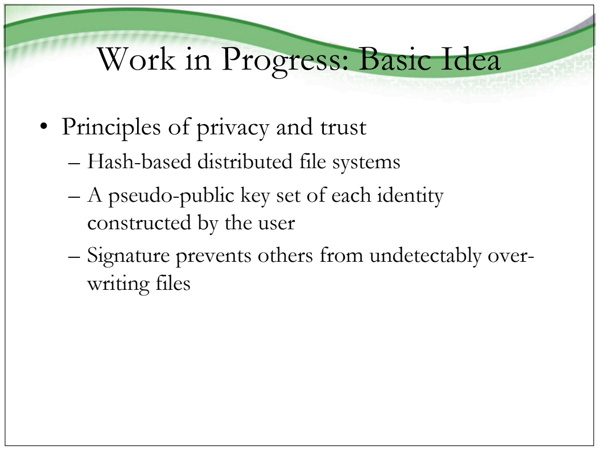
Server Implementation


username, role, password
Alice
alice@
work
a@wkpk, a@wksk
Trusted
Third Party
Accept: (invite, nonce, bob@work)a@wsk
Invite: (invite, nonce, bob@work)a@wsk
username, role, password
filenamea@wksk
Bob
bob@
work



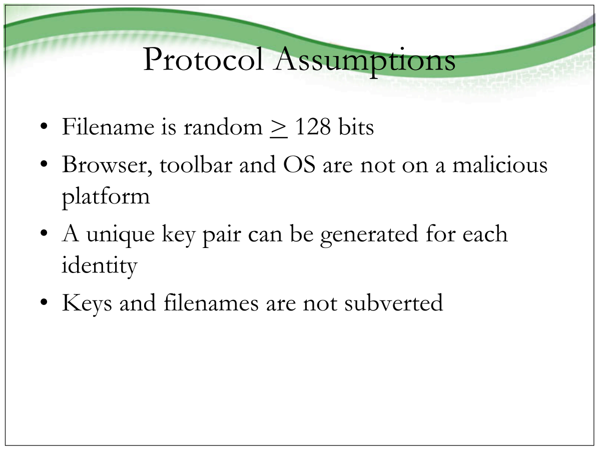
Problems

- Single Trusted server
  - Privacy of content can be preserved with encryption
  - Trivial traffic analysis
  - Subversion of user information
    - correlation of “identities”
      - alice@work == alice@home == monkeygrrl@tpmcafe
  - Single point of failure
    - obviously solvable
  - Questionable economic model



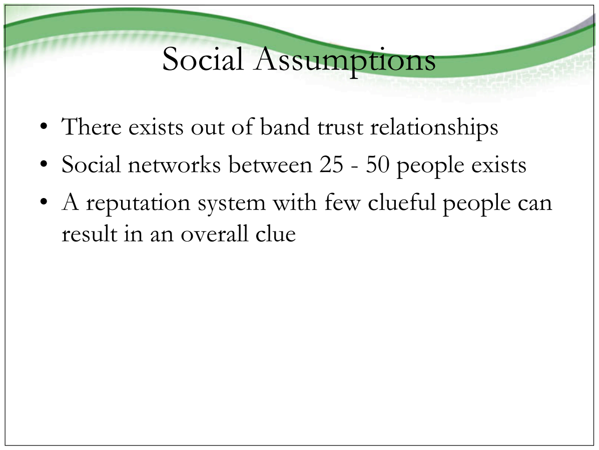
Work in Progress: Basic Idea

- Principles of privacy and trust
  - Hash-based distributed file systems
  - A pseudo-public key set of each identity constructed by the user
  - Signature prevents others from undetectably over-writing files



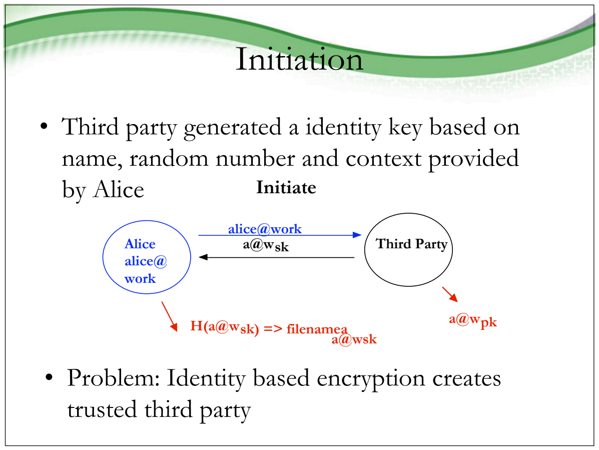
Protocol Assumptions

- Filename is random ≥ 128 bits
- Browser, toolbar and OS are not on a malicious platform
- A unique key pair can be generated for each identity
- Keys and filenames are not subverted



Social Assumptions

- There exists out of band trust relationships
- Social networks between 25 - 50 people exists
- A reputation system with few clueful people can result in an overall clue



Initiation

- Third party generated a identity key based on name, random number and context provided by Alice

Initiate

Alice alice@work

alice@work

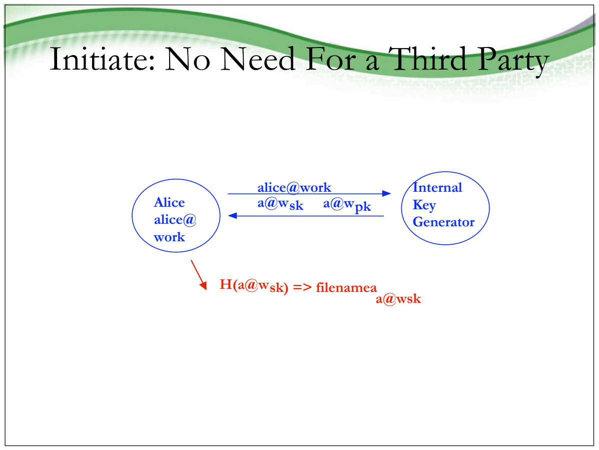
a@w$_{sk}$

Third Party

H(a@w$_{sk}$) => filenamea$_{a@wsk}$

a@w$_{pk}$

- Problem: Identity based encryption creates trusted third party



Initiate: No Need For a Third Party

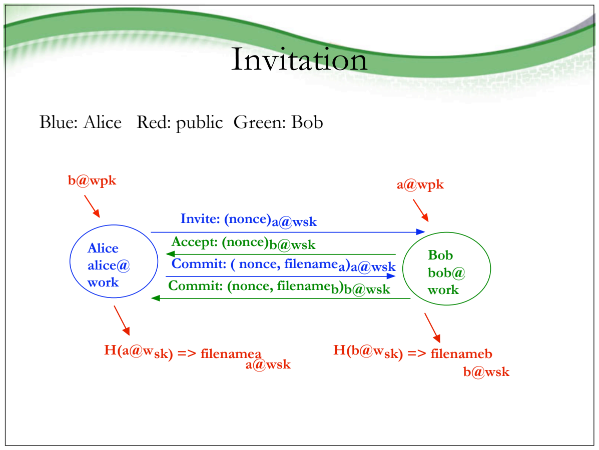
Alice alice@ work

alice@work

a@w$_{sk}$ a@w$_{pk}$

Internal Key Generator

H(a@w$_{sk}$) => filenamea$_{a@wsk}$



Invitation

Blue: Alice Red: public Green: Bob

b@wpk

a@wpk

Invite: (nonce)$_{a@wsk}$

Accept: (nonce)$_{b@wsk}$

Commit: (nonce, filename$_a$)$_{a@wsk}$

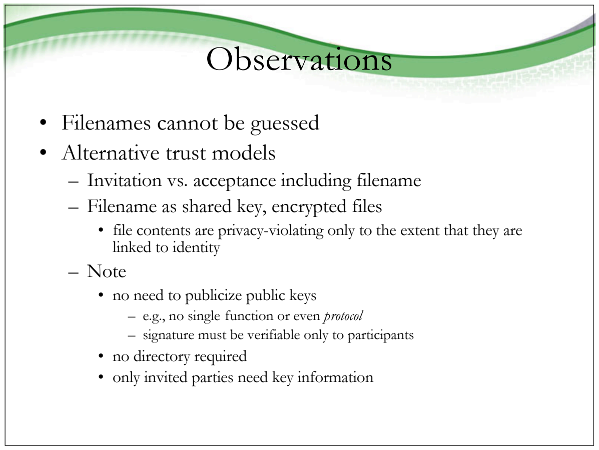
Commit: (nonce, filename$_b$)$_{b@wsk}$

Alice alice@ work

Bob bob@ work

H(a@w$_{sk}$) => filenamea$_{a@wsk}$

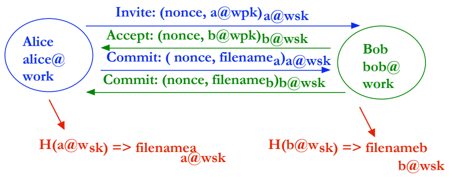
H(b@w$_{sk}$) => filenameb$_{b@wsk}$



Observations

- Filenames cannot be guessed
- Alternative trust models
  - Invitation vs. acceptance including filename
  - Filename as shared key, encrypted files
    - file contents are privacy-violating only to the extent that they are linked to identity
  - Note
    - no need to publicize public keys
      - e.g., no single function or even *protocol*
      - signature must be verifiable only to participants
    - no directory required
    - only invited parties need key information

Invitation to a Network



Invite: (nonce, a@wpk)$_{a@wsk}$

Accept: (nonce, b@wpk)$_{b@wsk}$

Commit: (nonce, filename$_a$)$_{a@wsk}$

Commit: (nonce, filename$_b$)$_{b@wsk}$

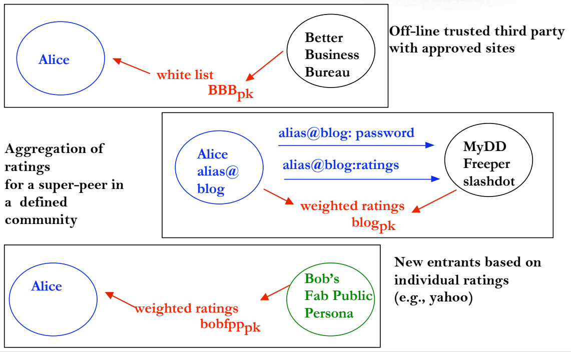
Alice alice@ work

Bob bob@ work

H(a@w$_{sk}$) => filename$_{a\,a@wsk}$

H(b@w$_{sk}$) => filename$_{b\,b@wsk}$

Broadcasting Options



Alice

white list BBB$_{pk}$

Better Business Bureau

Off-line trusted third party with approved sites

Aggregation of ratings for a super-peer in a defined community

Alice alias@ blog

alias@blog: password

alias@blog:ratings

MyDD Freeper slashdot

weighted ratings blog$_{pk}$

Alice

weighted ratings bobfpp$_{pk}$

Bob's Fab Public Persona

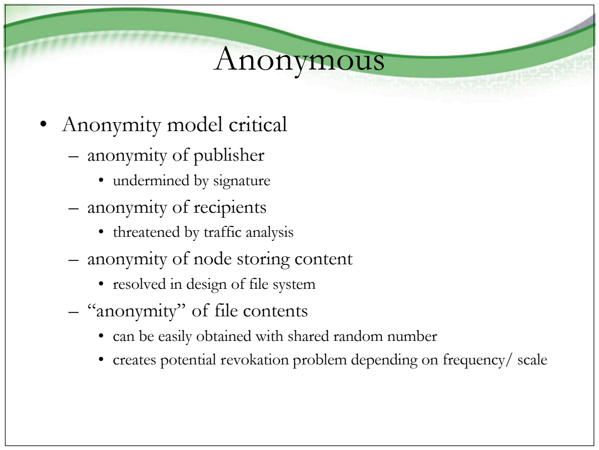
New entrants based on individual ratings (e.g., yahoo)

The Distributed File System

- Lookup protocol need only map filename
  - key:identity affiliations are stored only in the client
- The storage protocol must store, replicate, cache, and retrieve data
- Authentication of data occurs in the client

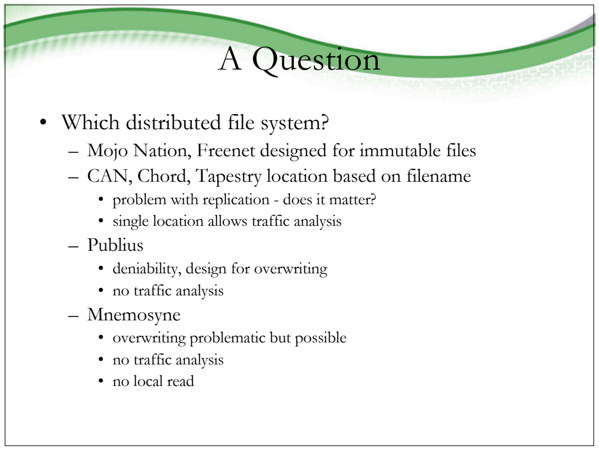
Current Questions

- Trade-offs of distributed availability within social network
  - geographical assumptions?
  - availability vs. efficiency
  - survivability
- Is the social network topology significantly different to enable increases in efficiency?
- Validate adversarial model



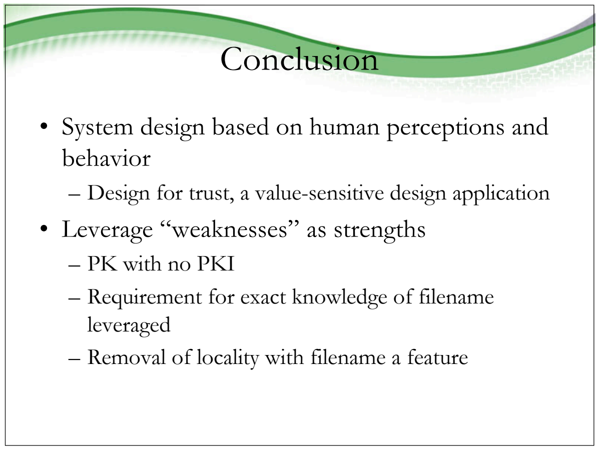
Anonymous

- Anonymity model critical
  - anonymity of publisher
    - undermined by signature
  - anonymity of recipients
    - threatened by traffic analysis
  - anonymity of node storing content
    - resolved in design of file system
  - “anonymity” of file contents
    - can be easily obtained with shared random number
    - creates potential revokation problem depending on frequency/ scale



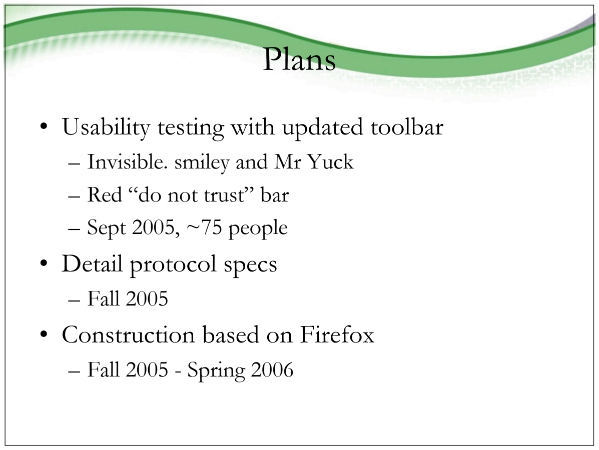
A Question

- Which distributed file system?
  - Mojo Nation, Freenet designed for immutable files
  - CAN, Chord, Tapestry location based on filename
    - problem with replication - does it matter?
    - single location allows traffic analysis
  - Publius
    - deniability, design for overwriting
    - no traffic analysis
  - Mnemosyne
    - overwriting problematic but possible
    - no traffic analysis
    - no local read



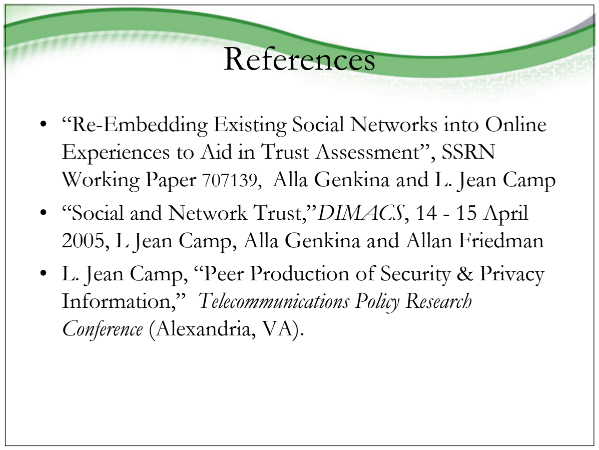
Conclusion

- System design based on human perceptions and behavior
  - Design for trust, a value-sensitive design application
- Leverage “weaknesses” as strengths
  - PK with no PKI
  - Requirement for exact knowledge of filename leveraged
  - Removal of locality with filename a feature



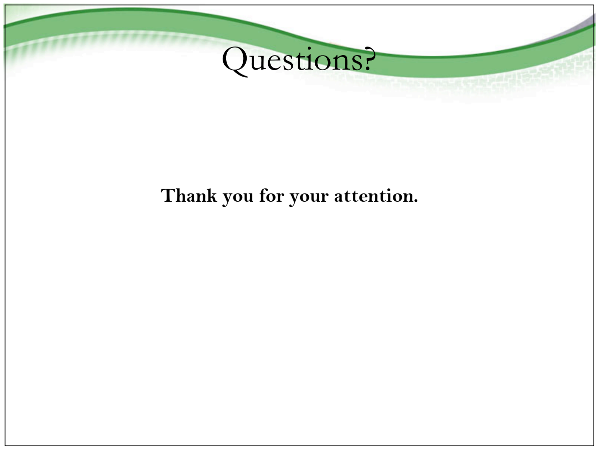
Plans

- Usability testing with updated toolbar
  - Invisible. smiley and Mr Yuck
  - Red “do not trust” bar
  - Sept 2005, ~75 people
- Detail protocol specs
  - Fall 2005
- Construction based on Firefox
  - Fall 2005 - Spring 2006



References

- "Re-Embedding Existing Social Networks into Online Experiences to Aid in Trust Assessment", SSRN Working Paper 707139, Alla Genkina and L. Jean Camp
- "Social and Network Trust,"*DIMACS*, 14 - 15 April 2005, L Jean Camp, Alla Genkina and Allan Friedman
- L. Jean Camp, "Peer Production of Security & Privacy Information," *Telecommunications Policy Research Conference* (Alexandria, VA).



Questions?

Thank you for your attention.