CSC 495.002 – Lecture 14
Privacy Perceptions: Privacy Attitudes

Dr. Özgür Kafalı

North Carolina State University
Department of Computer Science

Fall 2017

Westin Categories

- How Westin's survey evolved over time
- Limitations/criticism to Westin's privacy index
- Alternative studies to Westin

Problem Definition

- What changes in privacy behavior and attitudes among different cultures?
- What factors cause such changes?
- Mental models of humans for privacy decision making

Networked Privacy

- A model of privacy that is networked
- “Networked privacy invokes the constellation of audience dynamics, social norms, and technical functionality that affect the processes of information disclosure, concealment, obscurity, and interpretation within a networked public.”
- Privacy in social networks cannot only be controlled by individuals
  - Not completely depend on individual choices
  - No absolute control over own data

Marwick and Boyd. Networked privacy: How teenagers negotiate context in social media. New Media & Society, 16(7):1051–1067, 2014

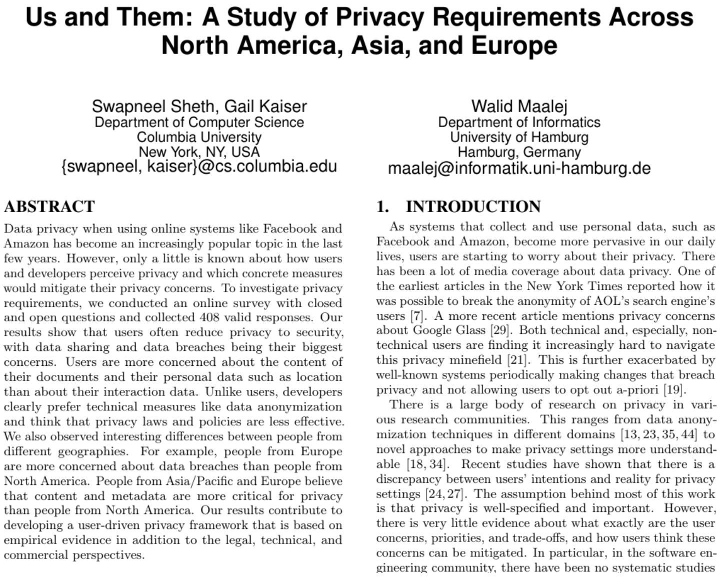
Transitional and Contextual Privacy

- Transitional: Even if a user makes a picture available to only three friends, these friends can easily disseminate it further
- Contextual: Privacy violations occur once information slips to a different context with different norms

Studies

- Four studies
- Privacy attitudes among countries and cultures [2 papers]
  - America
  - Europe
  - Asia
- mTurk workers vs the public
- User mental models about privacy

Us and Them



Us and Them: A Study of Privacy Requirements Across North America, Asia, and Europe

Swapneel Sheth, Gail Kaiser
Department of Computer Science
Columbia University
New York, NY, USA
{swapneel, kaiser}@cs.columbia.edu

Walid Maalej
Department of Informatics
University of Hamburg
Hamburg, Germany
maalej@informatik.uni-hamburg.de

ABSTRACT

Data privacy when using online systems like Facebook and Amazon has become an increasingly popular topic in the last few years. However, only a little is known about how users and developers perceive privacy and which concrete measures would mitigate their privacy concerns. To investigate privacy requirements, we conducted an online survey with closed and open questions and collected 408 valid responses. Our results show that users often reduce privacy to security, with data sharing and data breaches being their biggest concerns. Users are more concerned about the content of their documents and their personal data such as location than about their interaction data. Unlike users, developers clearly prefer technical measures like data anonymization and think that privacy laws and policies are less effective. We also observed interesting differences between people from different geographies. For example, people from Europe are more concerned about data breaches than people from North America. People from Asia/Pacific and Europe believe that content and metadata are more critical for privacy than people from North America. Our results contribute to developing a user-driven privacy framework that is based on empirical evidence in addition to the legal, technical, and commercial perspectives.

1. INTRODUCTION

As systems that collect and use personal data, such as Facebook and Amazon, become more pervasive in our daily lives, users are starting to worry about their privacy. There has been a lot of media coverage about data privacy. One of the earliest articles in the New York Times reported how it was possible to break the anonymity of AOL's search engine's users [7]. A more recent article mentions privacy concerns about Google Glass [29]. Both technical and, especially, non-technical users are finding it increasingly hard to navigate this privacy minefield [21]. This is further exacerbated by well-known systems periodically making changes that breach privacy and not allowing users to opt out a-priori [19].

There is a large body of research on privacy in various research communities. This ranges from data anonymization techniques in different domains [13, 23, 35, 44] to novel approaches to make privacy settings more understandable [18, 34]. Recent studies have shown that there is a discrepancy between users' intentions and reality for privacy settings [24, 27]. The assumption behind most of this work is that privacy is well-specified and important. However, there is very little evidence about what exactly are the user concerns, priorities, and trade-offs, and how users think these concerns can be mitigated. In particular, in the software engineering community, there have been no systematic studies

Sheth et al. Us and Them: A Study of Privacy Requirements Across North America, Asia, and Europe. International Conference on Software Engineering (ICSE), pages 859–870, 2014

Objective and Research Questions

- Objective: Understand (data) privacy expectations from modern software systems

- What are developers' and users' perceptions of privacy?
- Does experience in software development have any impact on privacy requirements?
- Does geography have any impact on privacy requirements?

Contributions

- Demonstrate general trends on how users understand privacy
- Understand how users assess privacy concerns
- Identify privacy expectations
- Provide insights into how software engineers can analyze privacy concerns of users

Methodology: Participants

	Developers	Users
North America	85	44
Europe	116	65
Asia	61	30
South America	3	2
Africa	2	0

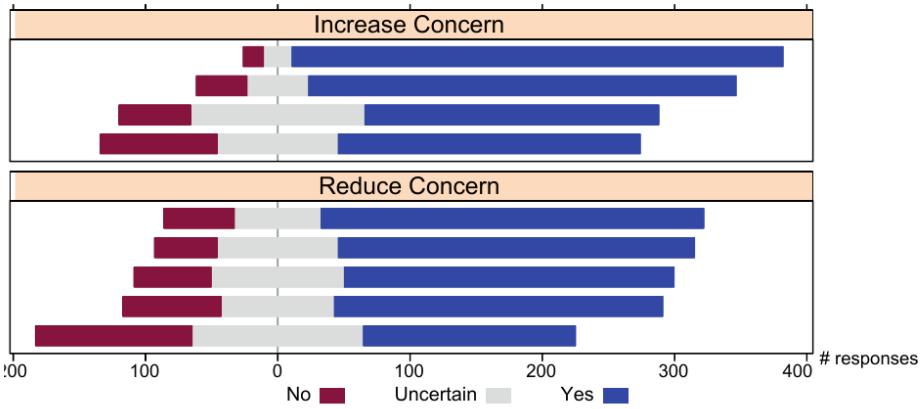
Factors to Increase Privacy Concerns

- Data Aggregation: The system discovers additional information about the user by aggregating data over a long period of time
- Data Distortion: The system might misrepresent the data or user intent
- Data Sharing: The collected data might be given to third parties for purposes like advertising
- Data Breaches: Malicious users might get access to sensitive data about other users

Factors to Reduce Privacy Concerns

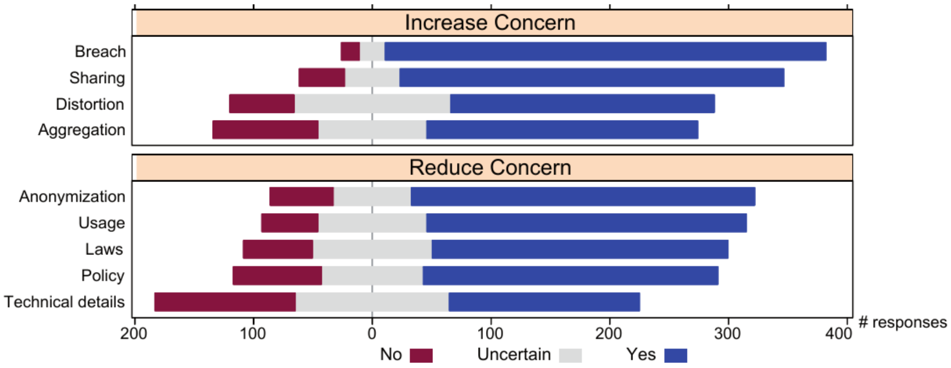
- Privacy Policy, License Agreements: Describing what the system will/won’t do with the data
- Privacy Laws: Describing which national law the system is compliant with (e.g., HIPAA in the US, European privacy laws)
- Anonymizing all data: Ensuring that none of the data has any personal identifiers
- Technical Details: Describing the algorithms/source code of the system in order to achieve higher trust (e.g., encryption of data)
- Details on usage: Describe, e.g., in a table how different data are used

Exercise: Factors to Increase and Reduce Concerns



- Increase: Aggregation, Distortion, Sharing, Breaches
- Reduce: Policy, Laws, Anonymization, Technical details, Usage

Factors to Increase and Reduce Concerns

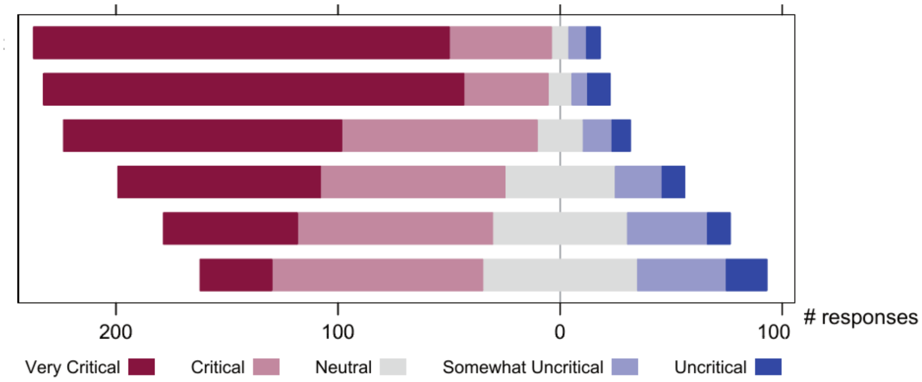


Privacy concerns	p-values
Sharing > Aggregation	$p = 1.231e^{-12}$
Sharing > Distortion	$p = 6.036e^{-14}$
Breach > Aggregation	$p < 2.2e^{-16}$
Breach > Distortion	$p < 2.2e^{-16}$

Exercise: Additional Concerns

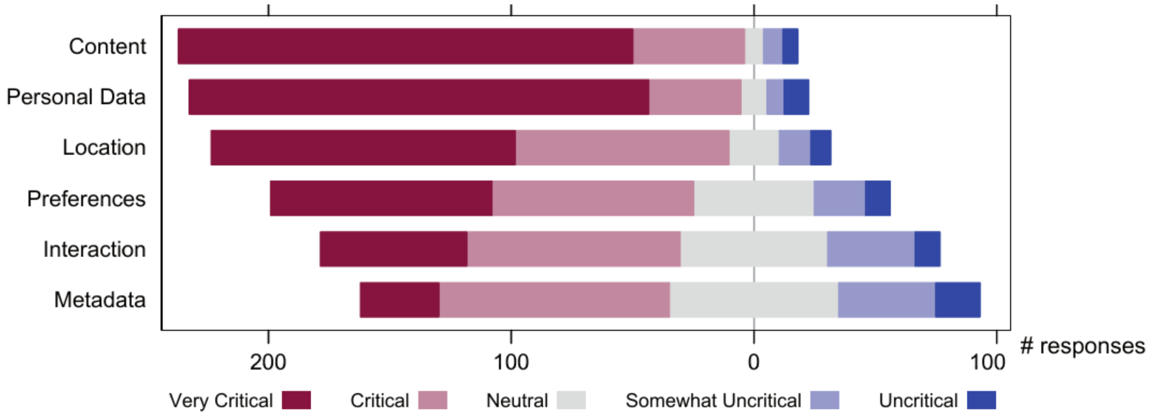
- Authorities and intelligence services: Government access [13/66]
- Malicious software or sharing data over APIs: Google Analytics [9/66]
- Unusable and nontransparent policies: Long, convoluted, hard to read [7/66]
- Lack of control: Options to delete data [7/66]

Exercise: Criticality of Types of Data



- Content (email body), metadata (date), interaction (mouse click to open email), location (city email is sent from), name (email address), user preferences (email settings)

Criticality of Types of Data



Giving up Privacy

- Would you accept less privacy for the following:
- Monetary discounts
- Added functionality of the system
- Fewer ads

- 37% accepts less privacy for added functionality
- 21% accepts less privacy for monetary discounts
- 14% accepts less privacy for fewer ads

Developers vs Users

- Data distortion: 49% of developers vs 65% of users
- Data aggregation: 52% of developers vs 63% of users
- Data criticality: Name, personal data, and interaction are more critical for developers
- Added functionality: 43% of developers give up privacy

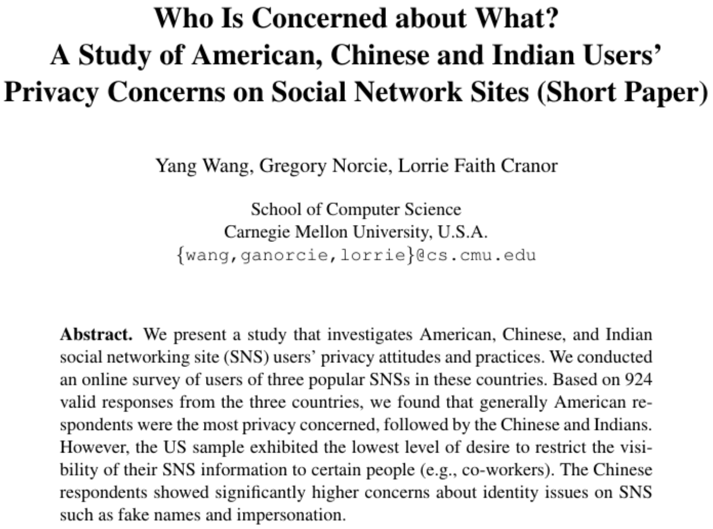
Role of Geography

- America thinks all types of data are less critical than Europe and Asia
- No statistically significant difference between Europe and Asia
- Added functionality: 51% of Europe does not give up, only 24% for America
- Europe feels that providing usage details is more effective than laws and policies
- America feels these options are all equal

Components of Privacy Framework

- Anonymization
- Data usage details
- Default encryption
- Fine-grained control
- Time and space limited storage
- Policies and laws

Who is Concerned About What?



**Who Is Concerned about What?
A Study of American, Chinese and Indian Users'
Privacy Concerns on Social Network Sites (Short Paper)**

Yang Wang, Gregory Norcie, Lorrie Faith Cranor

School of Computer Science
Carnegie Mellon University, U.S.A.
{wang,ganorcie,lorrie}@cs.cmu.edu

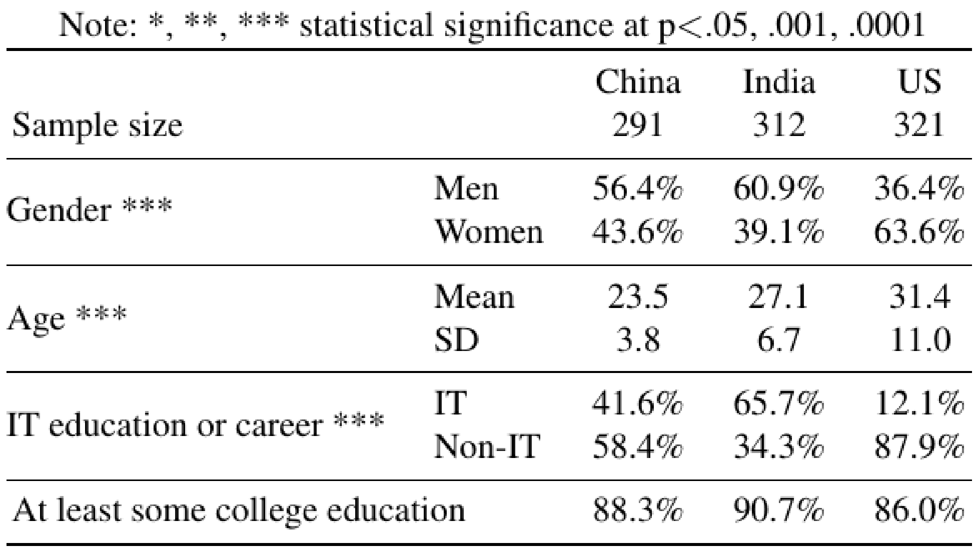
Abstract. We present a study that investigates American, Chinese, and Indian social networking site (SNS) users' privacy attitudes and practices. We conducted an online survey of users of three popular SNSs in these countries. Based on 924 valid responses from the three countries, we found that generally American respondents were the most privacy concerned, followed by the Chinese and Indians. However, the US sample exhibited the lowest level of desire to restrict the visibility of their SNS information to certain people (e.g., co-workers). The Chinese respondents showed significantly higher concerns about identity issues on SNS such as fake names and impersonation.

Wang et al. Who is Concerned About What? A Study of American, Chinese and Indian Users' Privacy Concerns on Social Network Sites. International Conference on Trust and Trustworthy Computing, pages 146–153, 2011

Survey Design

- Countries: US, China, India
- Social network sites:
  - Facebook: Highest traffic in US and India
  - Renren.com and Kaixin001.com: Domestic sites for China
- Common features: Profiles, walls, photo sharing, games, third party app development

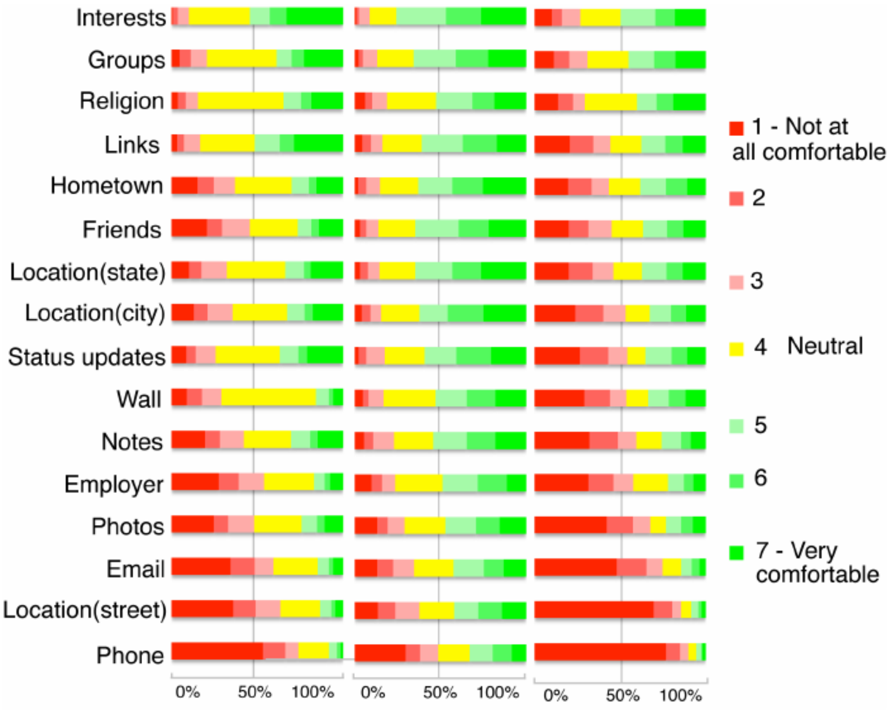
Demographics



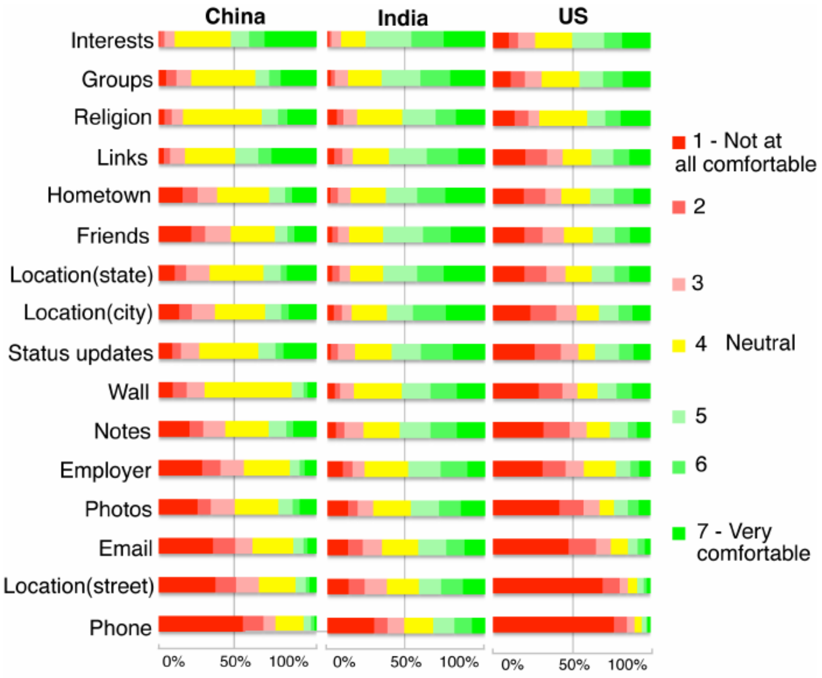
Note: *, **, *** statistical significance at p<.05, .001, .0001

		China	India	US
Sample size		291	312	321
Gender ***	Men	56.4%	60.9%	36.4%
	Women	43.6%	39.1%	63.6%
Age ***	Mean	23.5	27.1	31.4
	SD	3.8	6.7	11.0
IT education or career ***	IT	41.6%	65.7%	12.1%
	Non-IT	58.4%	34.3%	87.9%
At least some college education		88.3%	90.7%	86.0%

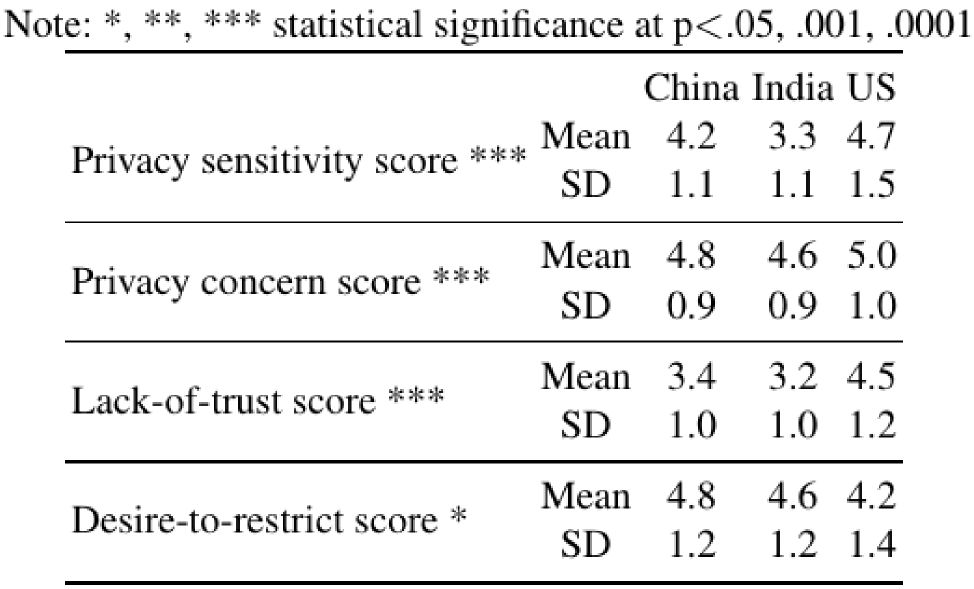
Exercise: Privacy Attitudes



Exercise: Privacy Attitudes



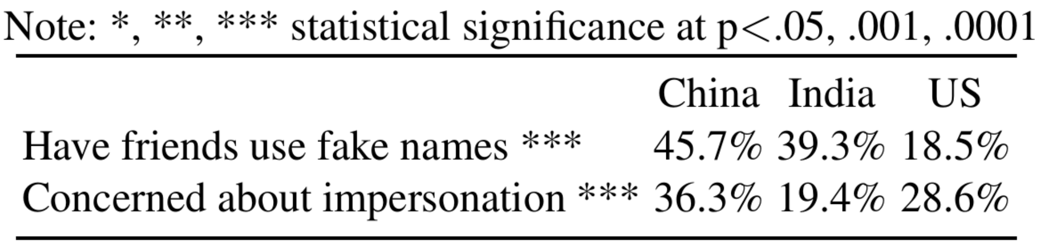
Attitude Dimensions



Note: *, **, *** statistical significance at p<.05, .001, .0001

		China	India	US
Privacy sensitivity score ***	Mean	4.2	3.3	4.7
	SD	1.1	1.1	1.5
Privacy concern score ***	Mean	4.8	4.6	5.0
	SD	0.9	0.9	1.0
Lack-of-trust score ***	Mean	3.4	3.2	4.5
	SD	1.0	1.0	1.2
Desire-to-restrict score *	Mean	4.8	4.6	4.2
	SD	1.2	1.2	1.4

Fake Names and Impersonation



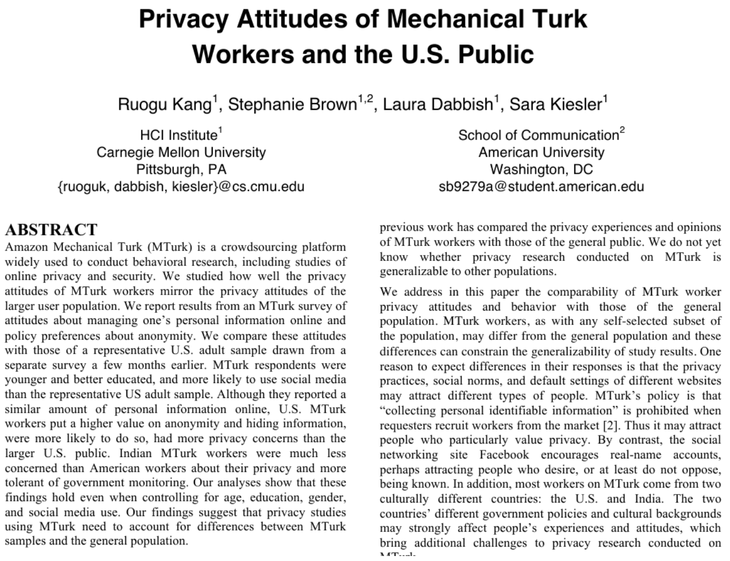
Note: *, **, *** statistical significance at p<.05, .001, .0001

	China	India	US
Have friends use fake names ***	45.7%	39.3%	18.5%
Concerned about impersonation ***	36.3%	19.4%	28.6%

Implications

- Recurring pattern: US > China > India
- Previous research on Chinese social network users
  - Venue for meeting new people and entertainment
  - Generally not privacy concerned
  - However, particularly concerned about identity issues

Mechanical Turk Workers vs US Public



Privacy Attitudes of Mechanical Turk Workers and the U.S. Public

Ruogu Kang[1], Stephanie Brown[1,2], Laura Dabbish[1], Sara Kiesler[1]

HCI Institute[1]
Carnegie Mellon University
Pittsburgh, PA
{ruoguk, dabbish, kiesler}@cs.cmu.edu

School of Communication[2]
American University
Washington, DC
sb9279a@student.american.edu

ABSTRACT

Amazon Mechanical Turk (MTurk) is a crowdsourcing platform widely used to conduct behavioral research, including studies of online privacy and security. We studied how well the privacy attitudes of MTurk workers mirror the privacy attitudes of the larger user population. We report results from an MTurk survey of attitudes about managing one's personal information online and policy preferences about anonymity. We compare these attitudes with those of a representative U.S. adult sample drawn from a separate survey a few months earlier. MTurk respondents were younger and better educated, and more likely to use social media than the representative US adult sample. Although they reported a similar amount of personal information online, U.S. MTurk workers put a higher value on anonymity and hiding information, were more likely to do so, had more privacy concerns than the larger U.S. public. Indian MTurk workers were much less concerned than American workers about their privacy and more tolerant of government monitoring. Our analyses show that these findings hold even when controlling for age, education, gender, and social media use. Our findings suggest that privacy studies using MTurk need to account for differences between MTurk samples and the general population.

previous work has compared the privacy experiences and opinions of MTurk workers with those of the general public. We do not yet know whether privacy research conducted on MTurk is generalizable to other populations.

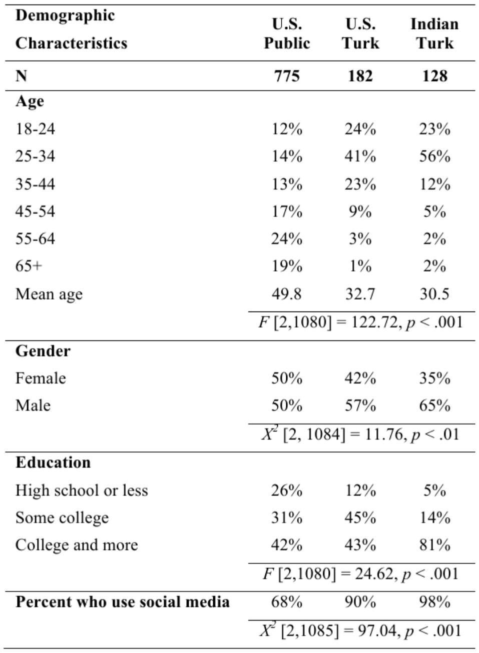
We address in this paper the comparability of MTurk worker privacy attitudes and behavior with those of the general population. MTurk workers, as with any self-selected subset of the population, may differ from the general population and these differences can constrain the generalizability of study results. One reason to expect differences in their responses is that the privacy practices, social norms, and default settings of different websites may attract different types of people. MTurk's policy is that "collecting personal identifiable information" is prohibited when requesters recruit workers from the market [2]. Thus it may attract people who particularly value privacy. By contrast, the social networking site Facebook encourages real-name accounts, perhaps attracting people who desire, or at least do not oppose, being known. In addition, most workers on MTurk come from two culturally different countries: the U.S. and India. The two countries' different government policies and cultural backgrounds may strongly affect people's experiences and attitudes, which bring additional challenges to privacy research conducted on MTurk.

Kang et al. Privacy Attitudes of Mechanical Turk Workers and the U.S. Public. Symposium on Usable Privacy and Security (SOUPS), pages 37–49, 2014

mTurk Workers

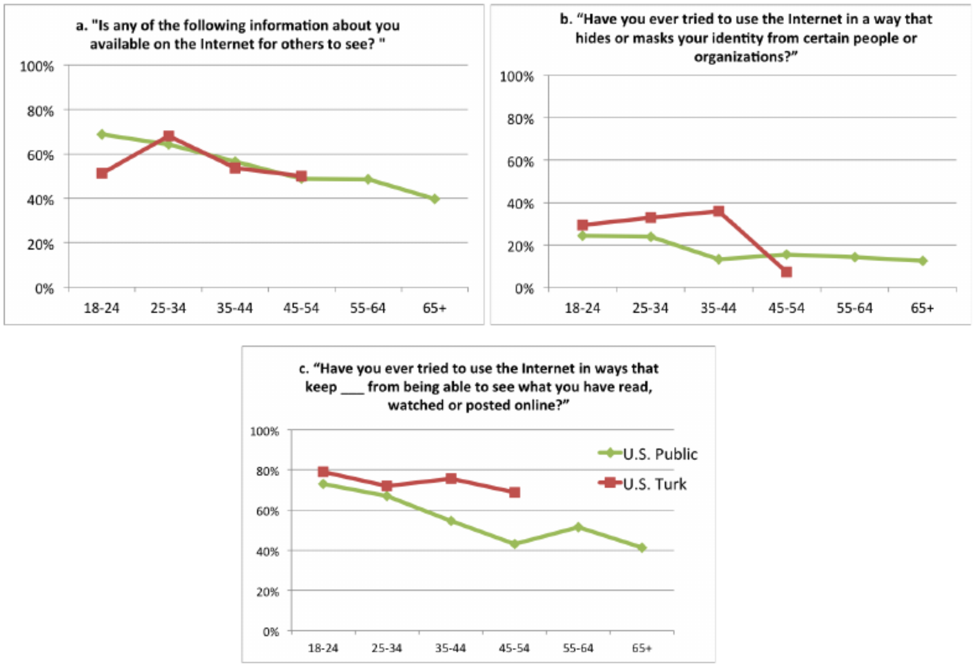
- Chosen an anonymous, flexible worksite
- Compared to the general population
  - Better educated
  - More liberal
  - Younger

Demographics: US Public vs mTurk

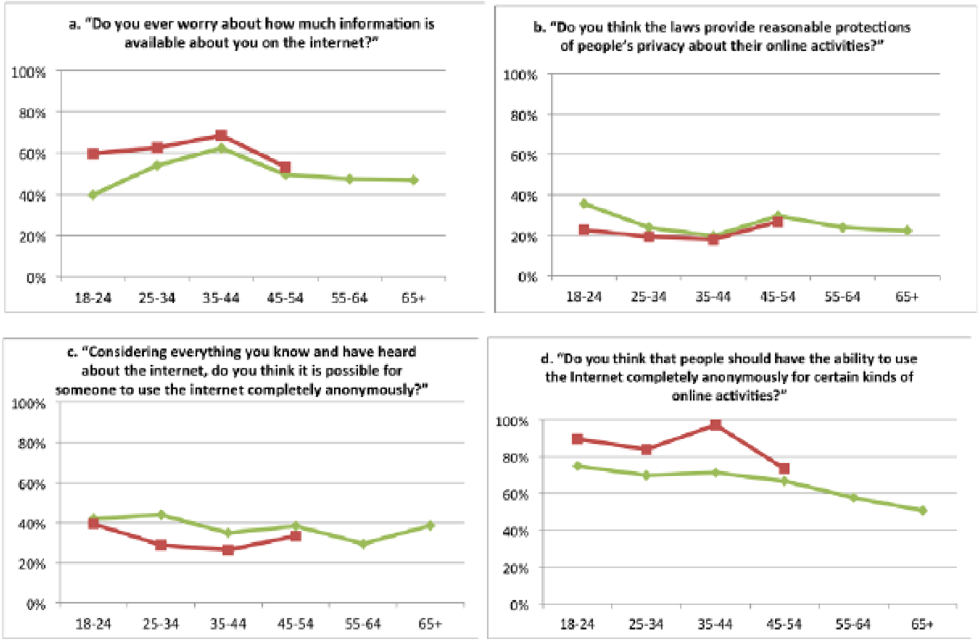


Demographic Characteristics	**U.S. Public**	**U.S. Turk**	**Indian Turk**
N	**775**	**182**	**128**
Age			
18-24	12%	24%	23%
25-34	14%	41%	56%
35-44	13%	23%	12%
45-54	17%	9%	5%
55-64	24%	3%	2%
65+	19%	1%	2%
Mean age	49.8	32.7	30.5
	F [2,1080] = 122.72, p < .001		
Gender			
Female	50%	42%	35%
Male	50%	57%	65%
	X^2 [2, 1084] = 11.76, p < .01		
Education			
High school or less	26%	12%	5%
Some college	31%	45%	14%
College and more	42%	43%	81%
	F [2,1080] = 24.62, p < .001		
Percent who use social media	68%	90%	98%
	X^2 [2,1085] = 97.04, p < .001		

Personal Information: US Public vs US mTurk



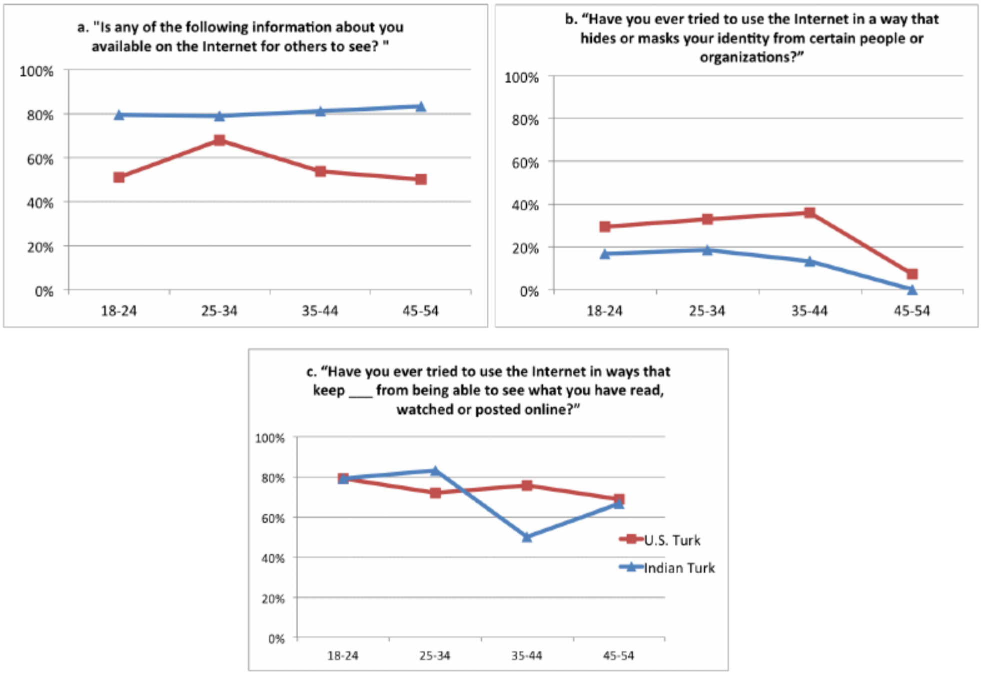
Privacy Preferences: US Public vs US mTurk



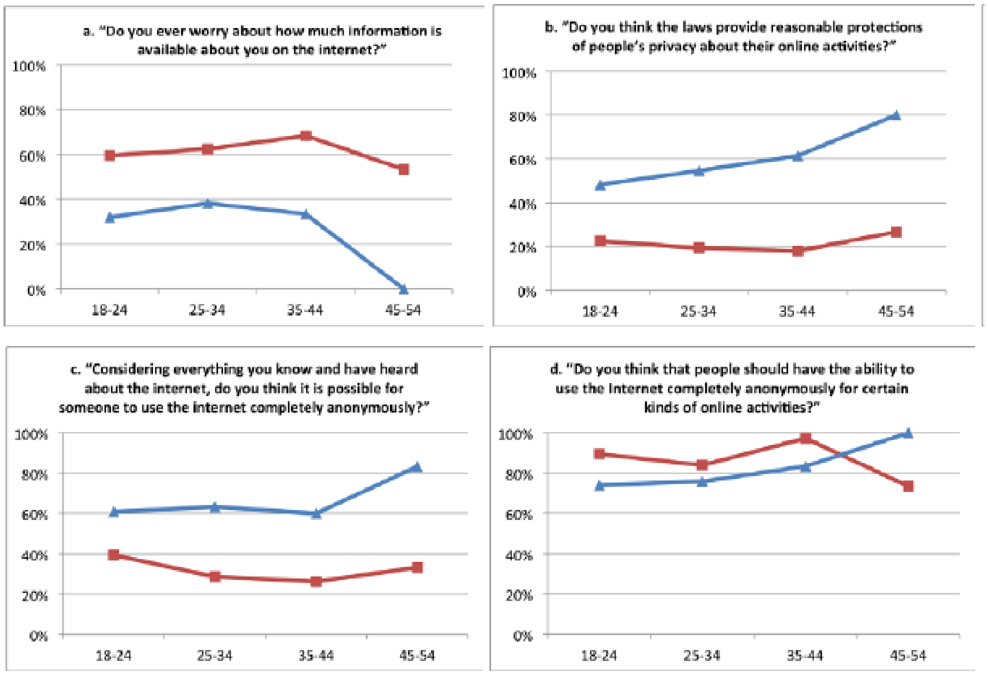
Findings

- On four of seven items US mTurk workers differs from US sample
- Even when demographic variables and social media use are taken into account
- Show similar trends based on age

Personal Information: US mTurk vs India mTurk



Privacy Preferences: US mTurk vs India mTurk



a. "Do you ever worry about how much information is available about you on the internet?"

b. "Do you think the laws provide reasonable protections of people's privacy about their online activities?"

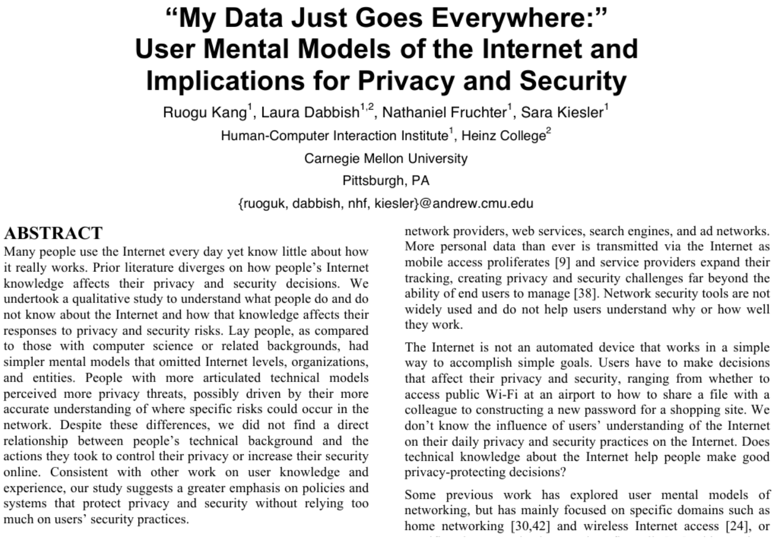
c. "Considering everything you know and have heard about the internet, do you think it is possible for someone to use the internet completely anonymously?"

d. "Do you think that people should have the ability to use the Internet completely anonymously for certain kinds of online activities?"

Implications

- mTurk workers more tech savvy than the general public
- Limitation: Survey for the US public did not ask questions about technology use

My Data Just Goes Everywhere



"My Data Just Goes Everywhere:" User Mental Models of the Internet and Implications for Privacy and Security

Ruogu Kang[1], Laura Dabbish[1,2], Nathaniel Fruchter[1], Sara Kiesler[1]
Human-Computer Interaction Institute[1], Heinz College[2]
Carnegie Mellon University
Pittsburgh, PA
{ruoguk, dabbish, nhf, kiesler}@andrew.cmu.edu

ABSTRACT

Many people use the Internet every day yet know little about how it really works. Prior literature diverges on how people's Internet knowledge affects their privacy and security decisions. We undertook a qualitative study to understand what people do and do not know about the Internet and how that knowledge affects their responses to privacy and security risks. Lay people, as compared to those with computer science or related backgrounds, had simpler mental models that omitted Internet levels, organizations, and entities. People with more articulated technical models perceived more privacy threats, possibly driven by their more accurate understanding of where specific risks could occur in the network. Despite these differences, we did not find a direct relationship between people's technical background and the actions they took to control their privacy or increase their security online. Consistent with other work on user knowledge and experience, our study suggests a greater emphasis on policies and systems that protect privacy and security without relying too much on users' security practices.

network providers, web services, search engines, and ad networks. More personal data than ever is transmitted via the Internet as mobile access proliferates [9] and service providers expand their tracking, creating privacy and security challenges far beyond the ability of end users to manage [38]. Network security tools are not widely used and do not help users understand why or how well they work.

The Internet is not an automated device that works in a simple way to accomplish simple goals. Users have to make decisions that affect their privacy and security, ranging from whether to access public Wi-Fi at an airport to how to share a file with a colleague to constructing a new password for a shopping site. We don't know the influence of users' understanding of the Internet on their daily privacy and security practices on the Internet. Does technical knowledge about the Internet help people make good privacy-protecting decisions?

Some previous work has explored user mental models of networking, but has mainly focused on specific domains such as home networking [30,42] and wireless Internet access [24], or

Kang et al. My Data Just Goes Everywhere: User Mental Models of the Internet and Implications for Privacy and Security. Symposium On Usable Privacy and Security (SOUPS), pages 39–52, 2015

Objective

- Explore the correlation between people's technology background and their privacy related actions

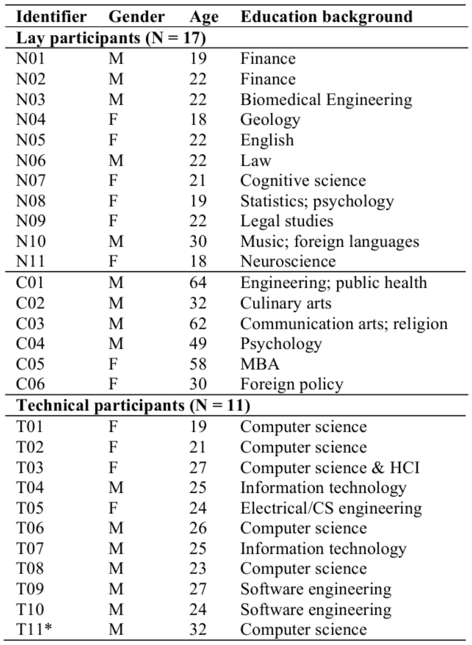
Knowledge of Internet

- Declarative knowledge: Knowledge about facts and terms (e.g., privacy settings, tagging, bcc)
- Procedural knowledge: How to take actions and complete tasks
- Technical familiarity
- Awareness of institutional practices
- Policy understanding
- User skills
- Awareness of security threats and tools

Mental Models

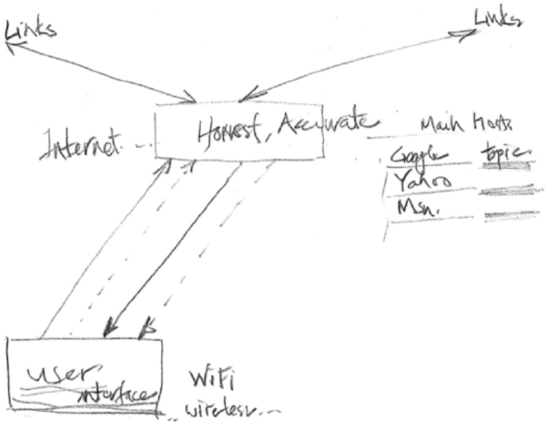
- Semi-structured interviews to understand users' mental models
- Participants draw diagrams about certain concepts
- Experts (faculty members in computer networking and security) review their drawings

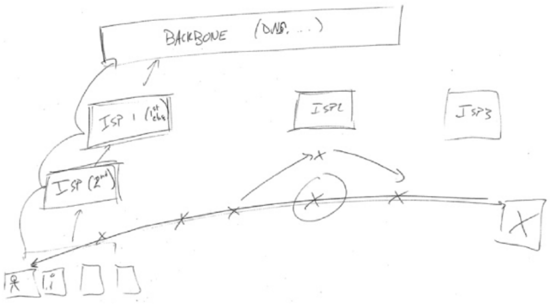
Participants



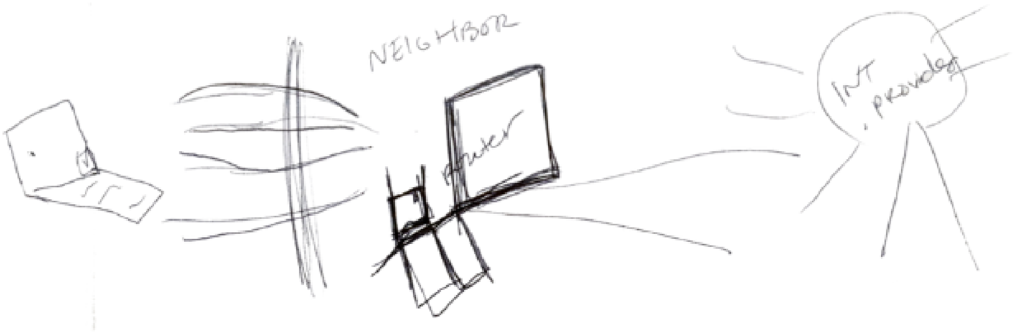
Identifier	Gender	Age	Education background
Lay participants (N = 17)			
N01	M	19	Finance
N02	M	22	Finance
N03	M	22	Biomedical Engineering
N04	F	18	Geology
N05	F	22	English
N06	M	22	Law
N07	F	21	Cognitive science
N08	F	19	Statistics; psychology
N09	F	22	Legal studies
N10	M	30	Music; foreign languages
N11	F	18	Neuroscience
C01	M	64	Engineering; public health
C02	M	32	Culinary arts
C03	M	62	Communication arts; religion
C04	M	49	Psychology
C05	F	58	MBA
C06	F	30	Foreign policy
Technical participants (N = 11)			
T01	F	19	Computer science
T02	F	21	Computer science
T03	F	27	Computer science & HCI
T04	M	25	Information technology
T05	F	24	Electrical/CS engineering
T06	M	26	Computer science
T07	M	25	Information technology
T08	M	23	Computer science
T09	M	27	Software engineering
T10	M	24	Software engineering
T11*	M	32	Computer science

Exercise: Internet as Service

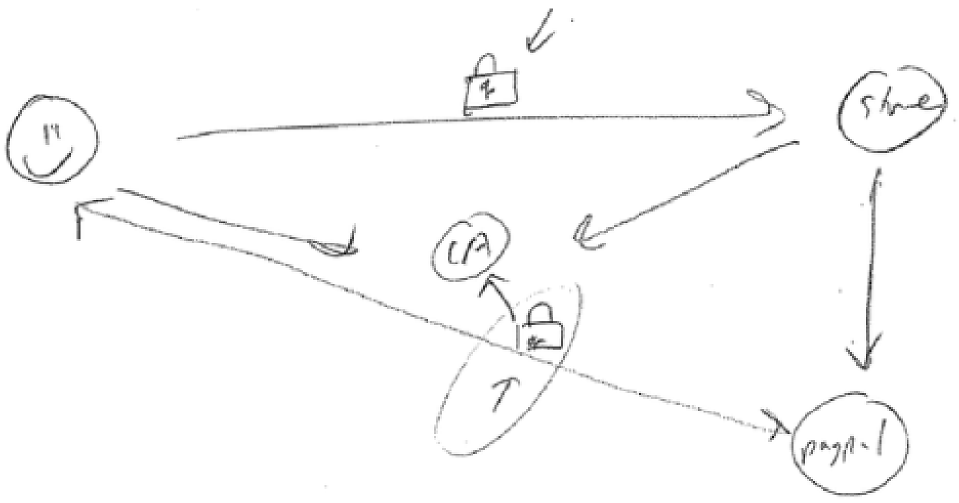




How to Use Neighbor’s WiFi

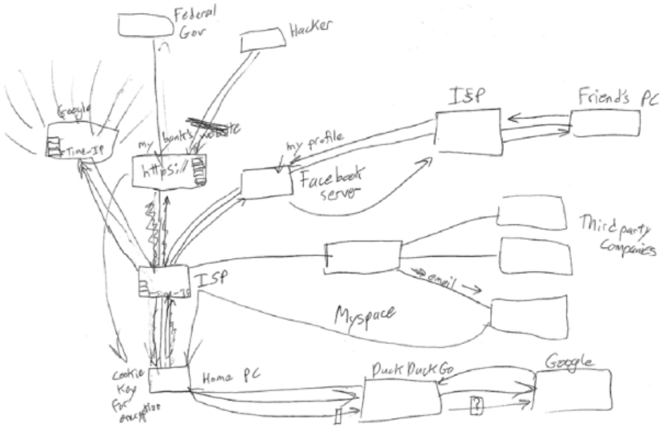


Exercise: Making Online Payment to a Shoe Store

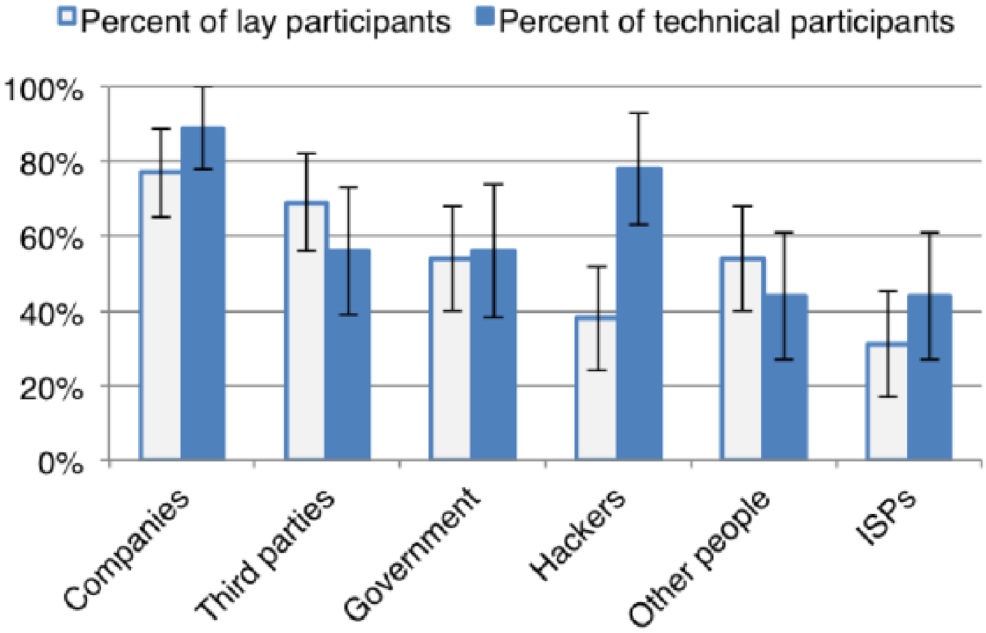


Who Can See My Data?

worldwide info
Earth
Facebook



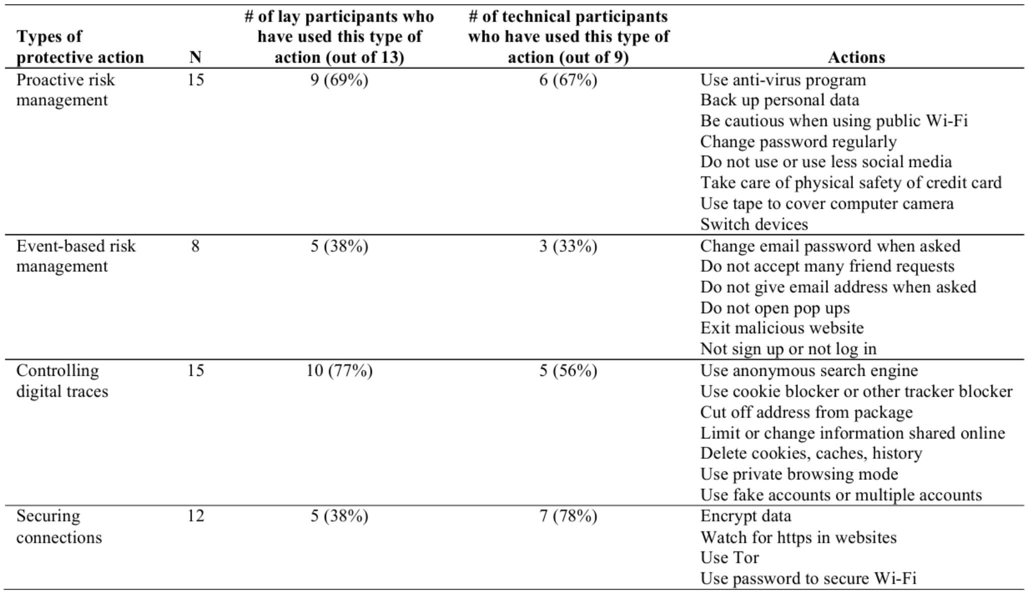
Access to Data: Non technical vs Technical



Quotes from Participants

- Non technical: “I don’t put [my credit card info] in when there’s not like that little lock up on top of the screen. I think it’s pretty secure.”
- Technical: “The Internet is not designed to be private.”
  “At the end of the day you’re relying on correct implementations of logically sound security protocols, and historically most implementations aren’t correct and most protocols aren’t logically sound. So, it’s just a question of an arms race of who’s paying more attention.”

Protective Actions



Types of protective action	N	# of lay participants who have used this type of action (out of 13)	# of technical participants who have used this type of action (out of 9)	Actions
Proactive risk management	15	9 (69%)	6 (67%)	Use anti-virus program Back up personal data Be cautious when using public Wi-Fi Change password regularly Do not use or use less social media Take care of physical safety of credit card Use tape to cover computer camera Switch devices
Event-based risk management	8	5 (38%)	3 (33%)	Change email password when asked Do not accept many friend requests Do not give email address when asked Do not open pop ups Exit malicious website Not sign up or not log in
Controlling digital traces	15	10 (77%)	5 (56%)	Use anonymous search engine Use cookie blocker or other tracker blocker Cut off address from package Limit or change information shared online Delete cookies, caches, history Use private browsing mode Use fake accounts or multiple accounts
Securing connections	12	5 (38%)	7 (78%)	Encrypt data Watch for https in websites Use Tor Use password to secure Wi-Fi

What Prevents People from Taking Privacy Actions?

- I have nothing to hide
- Doing so would sacrifice effectiveness or convenience
- Poor usability of privacy protection tools
- Lack of procedural knowledge

Shoot all the Drones

- News article: http://www.wdrb.com/story/29650818/hillview-man-arrested-for-shooting-down-drone-cites-right-to-privacy
- Links are also on the course website

Things to Look For

- Root cause: What went wrong?
- If it was not intentional, what was the original aim?
- Affected parties
- Implications and similar problems
- Mitigation (using methods we have seen): Prevention, detection, recovery

- Take 10 minutes to look at the incident on your own

- Now discuss with your neighbor
- Also take a look at the summary report: https://drive.google.com/file/d/0B3m-I0YVAv0EcnJiZUttTFhWeGs/view