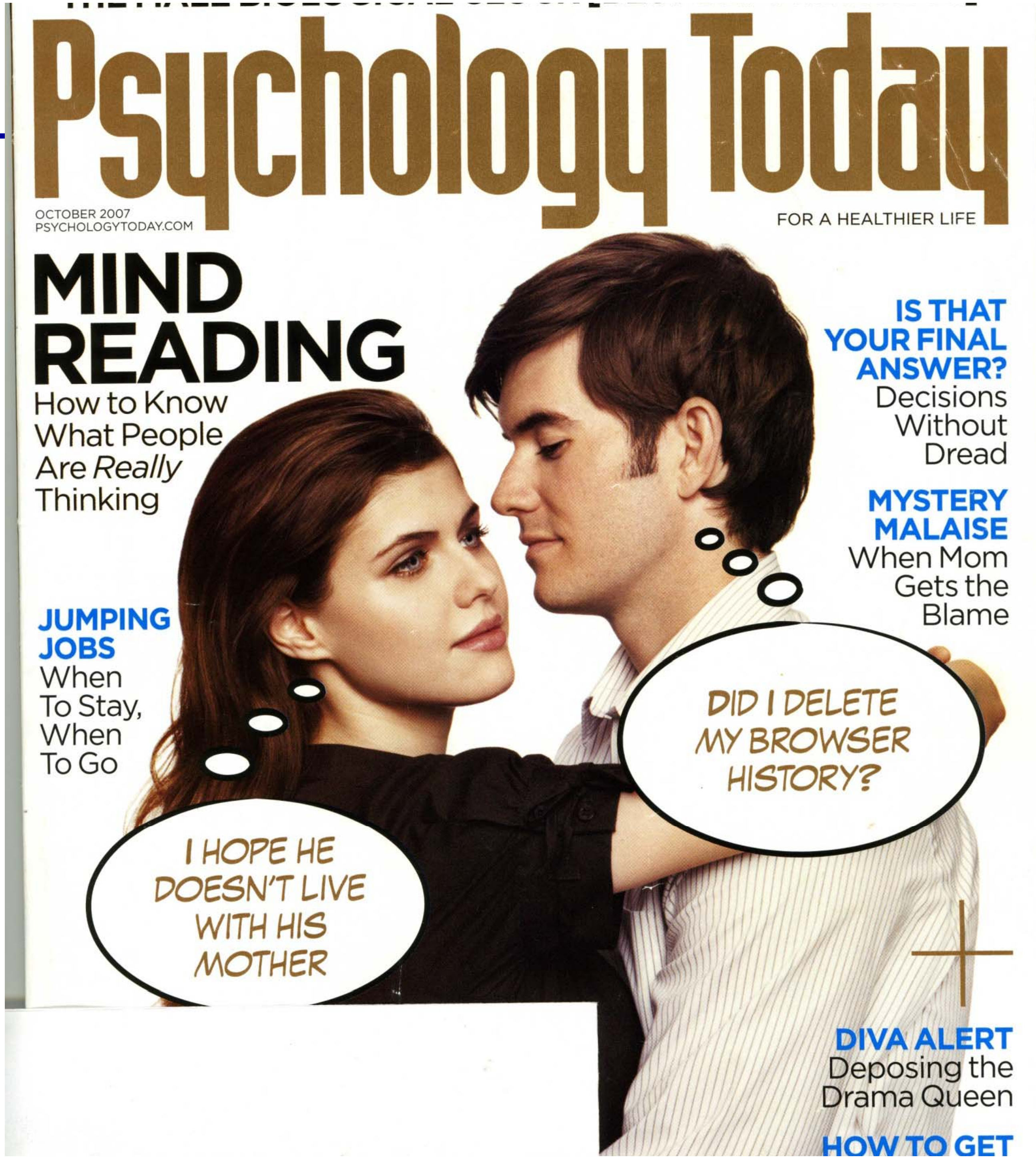
Security and Privacy
A Modern Perspective

Emmett Witchel

First Bytes Teachers Workshop

7/9/9

Thanks to Vitaly Shmatikov, James Hamilton


Psychology Today
OCTOBER 2007
PSYCHOLOGYTODAY.COM
FOR A HEALTHIER LIFE
MIND READING
How to Know What People Are *Really* Thinking
IS THAT YOUR FINAL ANSWER?
Decisions Without Dread
MYSTERY MALAISE
When Mom Gets the Blame
JUMPING JOBS
When To Stay, When To Go
DID I DELETE MY BROWSER HISTORY?
I HOPE HE DOESN'T LIVE WITH HIS MOTHER
DIVA ALERT
Deposing the Drama Queen
HOW TO GET

Outline

- Motivation
- Background & definitions for security
  - Cryptographic operations for security
- Netflix deanonymization attack
- Anonymity and privacy of social networks
- Just a touch of cloud computing
- Mandatory access control
- Differential privacy – interactive privacy

The problem

Potential solutions

Exposure to a modern view of security

Where is security headed?

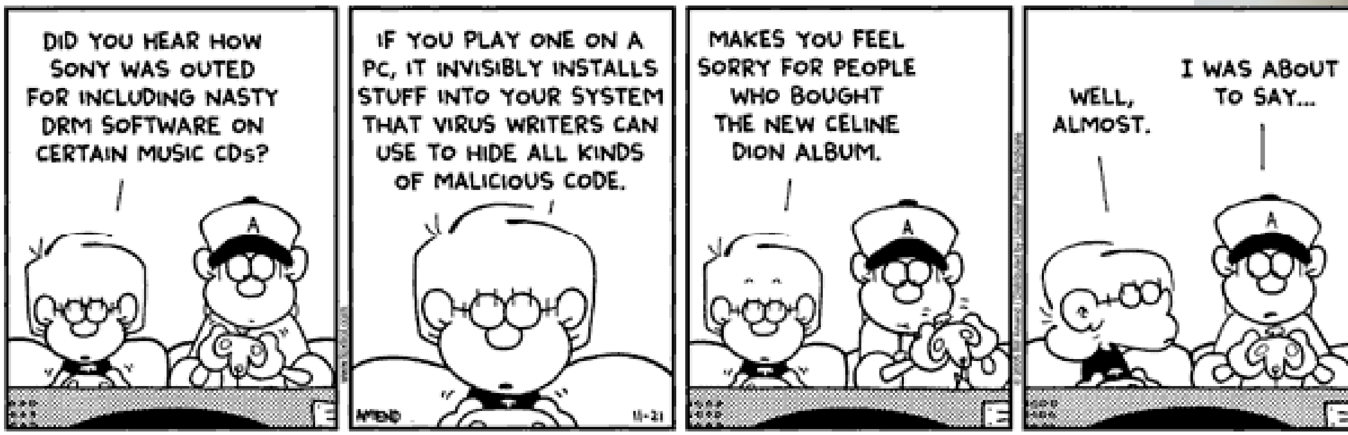
Leaking information

- Stealing 26.5 million veteran’s data
- Data on laptop stolen from employee’s home (5/06)
  - Veterans’ names
  - Social Security numbers
  - Dates of birth
- Exposure to identity theft

- CardSystems exposes data of 40 million cards (2005)
  - Data on 70,000 cards downloaded from ftp server

These are attacks on privacy (confidentiality, anonymity)

The Sony rootkit



- “Protected” albums included
  - Billie Holiday
  - Louis Armstrong
  - Switchfoot
  - The Dead 60’s
  - Flatt & Scruggs, etc.
- Rootkits modify files to infiltrate & hide
  - System configuration files
  - Drivers (executable files)

The Sony rootkit

- Sony's rootkit enforced DRM but exposed computer
  - CDs recalled
  - Classified as spyware by anti-virus software
  - Rootkit removal software distrubuted
  - Removal software had exposure vulnerability
  - New removal software distrubuted
- Sony sued by
  - Texas
  - New York
  - California

This is an attack on integrity

The Problem

- Types of misuse
  - Accidental
  - Intentional (malicious)

- Protection and security objective
  - Protect against/prevent misuse

- Three key components:
  - Authentication: Verify user identity
  - Integrity: Data has not been written by unauthorized entity
  - Privacy: Data has not been read by unauthorized entity

Have you used an anonymizing service?

1. Yes, for email
2. Yes, for web browsing
3. Yes for a pseudonymous service (craigslist)
4. Yes, for something else
5. No

What are your security goals?

- Authentication
  - User is who s/he says they are.
  - Example: Certificate authority (verisign)
- Integrity
  - Adversary can not change contents of message
  - But not necessarily private
  - Example: secure checksum
- Privacy (confidentiality)
  - Adversary can not read your message
  - If adversary eventually breaks your system can they decode all stored communication?
  - Example: Anonymous remailer (how to reply?)
- Authorization, repudiation (or non-repudiation), forward security (crack now, not crack future), backward security (crack now, not cracked past)

What About Security in Distributed Systems?

- Three challenges
  - Authentication
    - Verify user identity
  - Integrity
    - Verify that the communication has not been tempered with
  - Privacy
    - Protect access to communication across hosts
- Solution: Encryption
  - Achieves all these goals
  - Transform data that can easily reversed given the correct key (and hard to reverse without the key)
- Two common approaches
  - Private key encryption
  - Public key encryption
- Cryptographic hash
  - Hash is a fixed sized byte string which represents arbitrary length data. Hard to find two messages with same hash.
  - If m != m' then H(m) != H(m') with high probability. H(m) is 256 bits

Private Key (Symmetric Key) Encryption

- Basic idea:
  - {Plain text}^K → cipher text
  - {Cipher text}^K → plain text
  - As long as key K stays secret, we can get authentication, secrecy and integrity
- Infrastructure: Authentication server (example: kerberos)
  - Maintains a list of passwords; provides a key for two parties to communicate
- Basic steps (using secure server S)
  - A → S {Hi! I would like a key for AB}
  - S → A {Use Kab {This is A! Use Kab}^Kb}^Ka
  - A→ B {This is A! Use Kab}^Kb
  - Master keys (Ka and Kb) distributed out-of-band and stored securely at clients (the bootstrap problem)
- Refinements
  - Generate temporary keys to communicate between clients and authentication server

Public Key Encryption

- Basic idea:
  - Separate authentication from secrecy
  - Each key is a pair: K-public and K-private
  - {Plain text}^K-private → cipher text
  - {Cipher text}^K-public → plain text
  - K-private is kept a secret; K-public is distributed
- Examples:
  - {I'm Emmett}^K-private
    - Everyone can read it, but only I can send it (authentication)
  - {Hi, Emmett}^K-public
    - Anyone can send it but only I can read it (secrecy)
- Two-party communication
  - A → B {I'm A {use Kab}^K-privateA}^K-publicB
  - No need for an authentication server
  - Question: how do you trust the "public key" server?
    - Trusted server: {K-publicA}^K-privateS

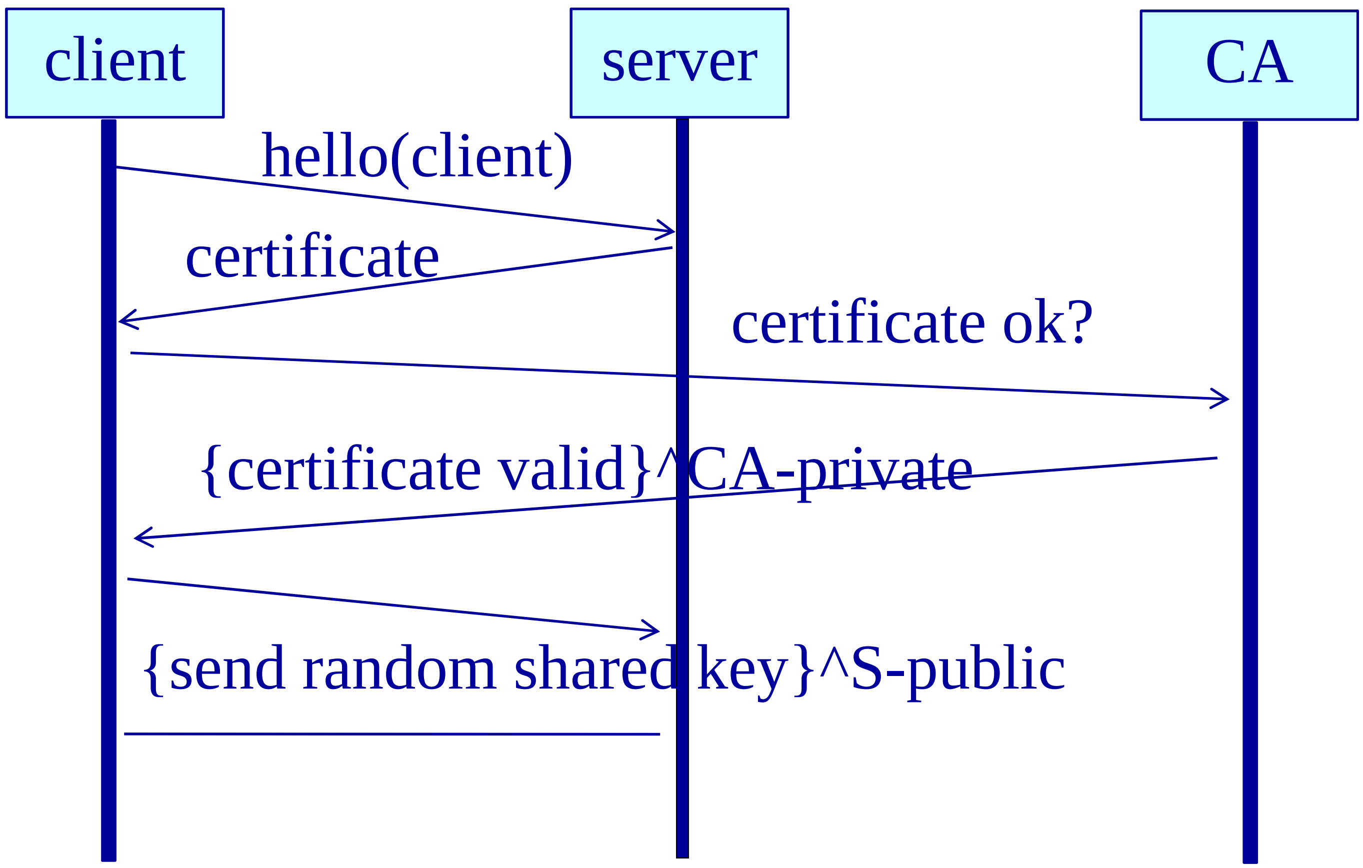
Implementing your security goals

- Authentication (requires public key infrastructure)
  - {I'm Emmett}^K-private
- Integrity (Digital signature)
  - {SHA-256 hash of message I just sent is …}^K-private
- Privacy (confidentiality)
  - Public keys to exchange a secret
  - Use shared-key cryptography (for speed)
  - Strategy used by ssh
- Forward/backward security
  - Rotate shared keys every hour
- Repudiation
  - Public list of cracked keys

When you visit a website using an http URL, which property are you missing?

1. Authentication (server to user)
2. Authentication (user to server)
3. Integrity
4. Privacy
5. None

Securing HTTP: HTTPS (HTTP+SSL/TLS)



switch to encrypted
connection using shared key

When you visit a website using an https URL, which property are you missing?

1. Authentication (server to user)
2. Authentication (user to server)
3. Integrity
4. Privacy
5. None

Authentication

- Objective: Verify user identity
- Common approach:
  - Passwords: shared secret between two parties
  - Present password to verify identity

1. How can the system maintain a copy of passwords?
   - Encryption: Transformation that is difficult to reverse without right key
   - Example: Unix /etc/passwd file contains encrypted passwords
   - When you type password, system encrypts it and then compared encrypted versions

Authentication (Cont'd.)

2. Passwords must be long and obscure
   - Paradox:
     - Short passwords are easy to crack
     - Long passwords – users write down to remember ➔ vulnerable
   - Original Unix:
     - 5 letter, lower case password
     - Exhaustive search requires 26^5 = 12 million comparisons
     - Today: < 1us to compare a password ➔ 12 seconds to crack a password
   - Choice of passwords
     - English words: Shakespeare's vocabulary: 30K words
     - All English words, fictional characters, place names, words reversed, … still too few words
     - (Partial) solution: More complex passwords
       - At least 8 characters long, with upper/lower case, numbers, and special characters

Alternatives/enhancements to Passwords

- Easier to remember passwords (visual recognition)
- Two-factor authentication
  - Password and some other channel, e.g., physical device with key that changes every minute
  - http://www.schneier.com/essay-083.html
  - What about a fake bank web site? (man in the middle)
  - Local Trojan program records second factor
- Biometrics
  - Fingerprint, retinal scan
  - What if I have a cut? What if someone wants my finger?
- Facial recognition

Password security

- Instead of hashing your password, I will hash your password concatenated with a random salt. Then I store the unhashed salt along with the hash.
  - (password . salt)^H salt
- What attack does this address?

1. Brute force password guessing for all accounts.
2. Brute force password guessing for one account.
3. Trojan horse password value
4. Man-in-the-middle attack when user gives password at login prompt.

Authorization

- Objective:
  - Specify access rights: who can do what?
- Access control: formalize all permissions in the system

	File1	File2	File3	…
User A	RW	R	--	…
User B	--	RW	RW	..
User C	RW	RW	RW	…

- Problem:
  - Potentially huge number of users, objects that dynamically change ➔ impractical
- Access control lists
  - Store permissions for all users with objects
  - Unix approach: three categories of access rights (owner, group, world)
  - Recent systems: more flexible with respect to group creation
- Privileged user (becomes security hole)
  - Administrator in windows, root in Unix
  - Principle of least privlege

Dweeb Nolife develops a file system that responds to requests with digitally signed packets of data from a content provider. Any untrusted machine can serve the data and clients can verify that the packets they receive were signed. So utexas.edu can give signed copies of the read-only portions of its web site to untrusted servers. Dweeb's FS provides which property?

1. Authentication of file system users
2. Integrity of file system contents
3. Privacy of file system data & metadata
4. Authorization of access to data & metadata

Outline

- Motivation
- Background & definitions for security
  - Cryptographic operations for security
- Netflix deanonymization attack
- Anonymity and privacy of social networks

The problem

- Just a touch of cloud computing
- Mandatory access control
- Differential privacy – interactive privacy

Netflix Prize Dataset

- Netflix: online movie rental service
- In October 2006, released real movie ratings of 500,000 subscribers
  - 10% of all Netflix users as of late 2005
  - Names removed
  - Information may be perturbed
  - Numerical ratings as well as dates
  - Average user rated over 200 movies
- Task is to predict how a user will rate a movie
  - Beat Netflix's algorithm (called Cinematch) by 10%
  - You get 1 million dollars

Netflix Prize

- Dataset properties
  - 17,770 movies
  - 480K people
  - 100M ratings
  - 3M unknowns
- 40,000+ teams
- 185 countries
- $1M for 10% gain

Netflix Prize

Home | Rules | Leaderboard | Register | Update | Submit | Download

Leaderboard

Display top 20 leaders.

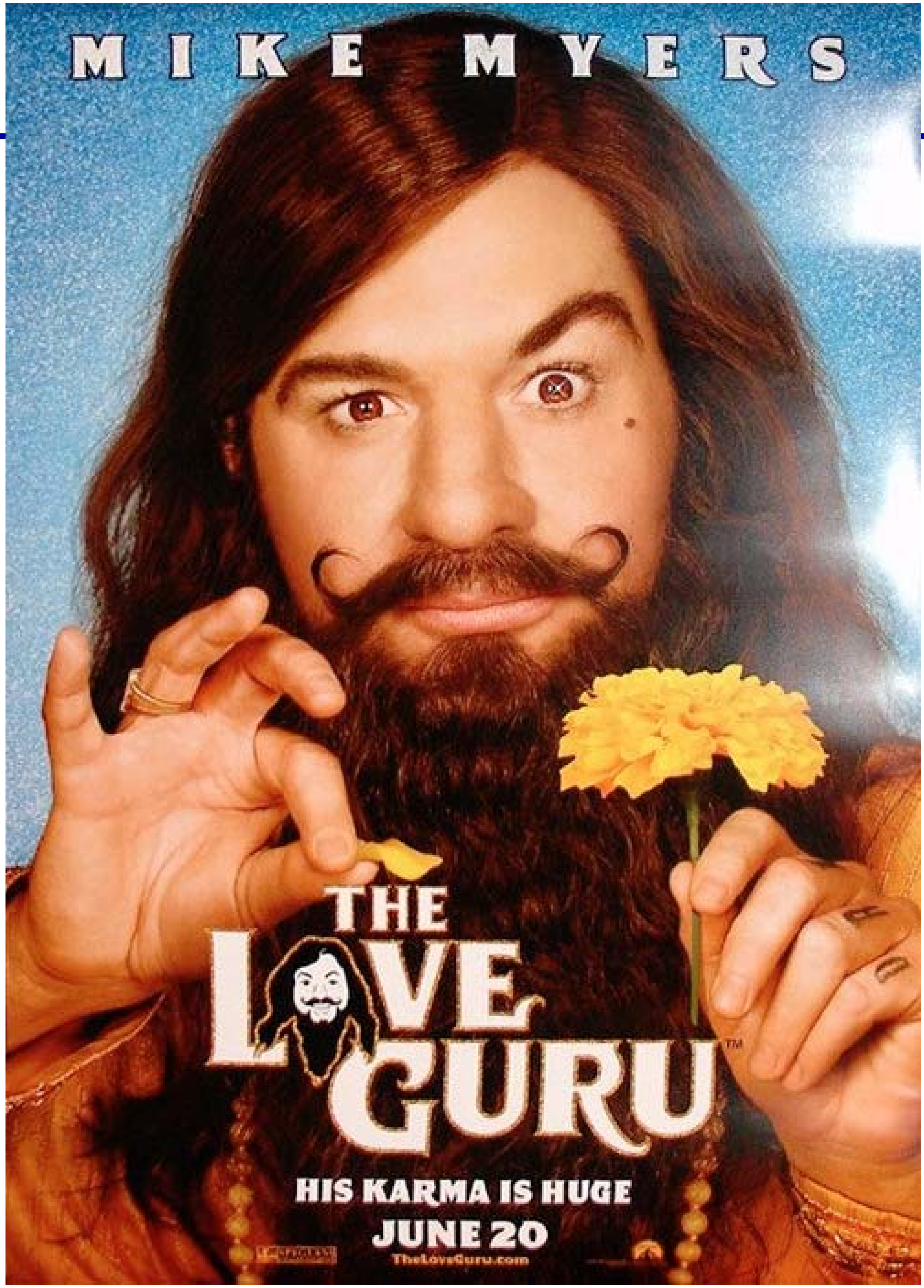
Rank	Team Name	Best Score	% Improvement	Last Submit Time
1	BellKor's Pragmatic Chaos	0.8558	10.05	2009-07-08 18:29:25
Grand Prize - RMSE <= 0.8563				
2	Grand Prize Team	0.8572	9.90	2009-07-07 21:37:25
3	Opera Solutions and Vandelay United	0.8576	9.86	2009-07-07 22:49:58
4	xlvector	0.8579	9.83	2009-07-08 08:36:52
5	PragmaticTheory	0.8582	9.80	2009-07-08 22:31:31
6	Vandelay Industries !	0.8584	9.78	2009-07-08 12:15:35
7	BellKor in BigChaos	0.8590	9.71	2009-07-08 06:55:44
8	Team ESP	0.8598	9.63	2009-07-08 08:03:14
9	BigChaos	0.8613	9.47	2009-06-23 23:06:52
10	Opera Solutions	0.8614	9.46	2009-07-02 17:32:37
11	BellKor	0.8615	9.45	2009-07-08 18:58:03
Progress Prize 2008 - RMSE = 0.8616 - Winning Team: BellKor in BigChaos				
12	space drop	0.8621	9.39	2009-07-09 05:59:48
13	Feeds2	0.8624	9.35	2009-07-09 07:25:14
14	Gravity	0.8634	9.25	2009-04-22 18:31:32
15	BruceDengDaoCiYiYou	0.8638	9.21	2009-06-27 00:55:55
16	pengpengzhou	0.8638	9.21	2009-06-27 01:06:43
17	majia2	0.8638	9.21	2009-07-07 07:13:18
18	Ces	0.8642	9.17	2009-07-07 03:14:03
19	We are the Borg	0.8643	9.15	2009-07-06 22:48:59
20	Just a guy in a garage	0.8650	9.08	2009-07-06 16:12:33
Progress Prize 2007 - RMSE = 0.8712 - Winning Team: KorBell				
Cinematch score on quiz subset - RMSE = 0.9514				

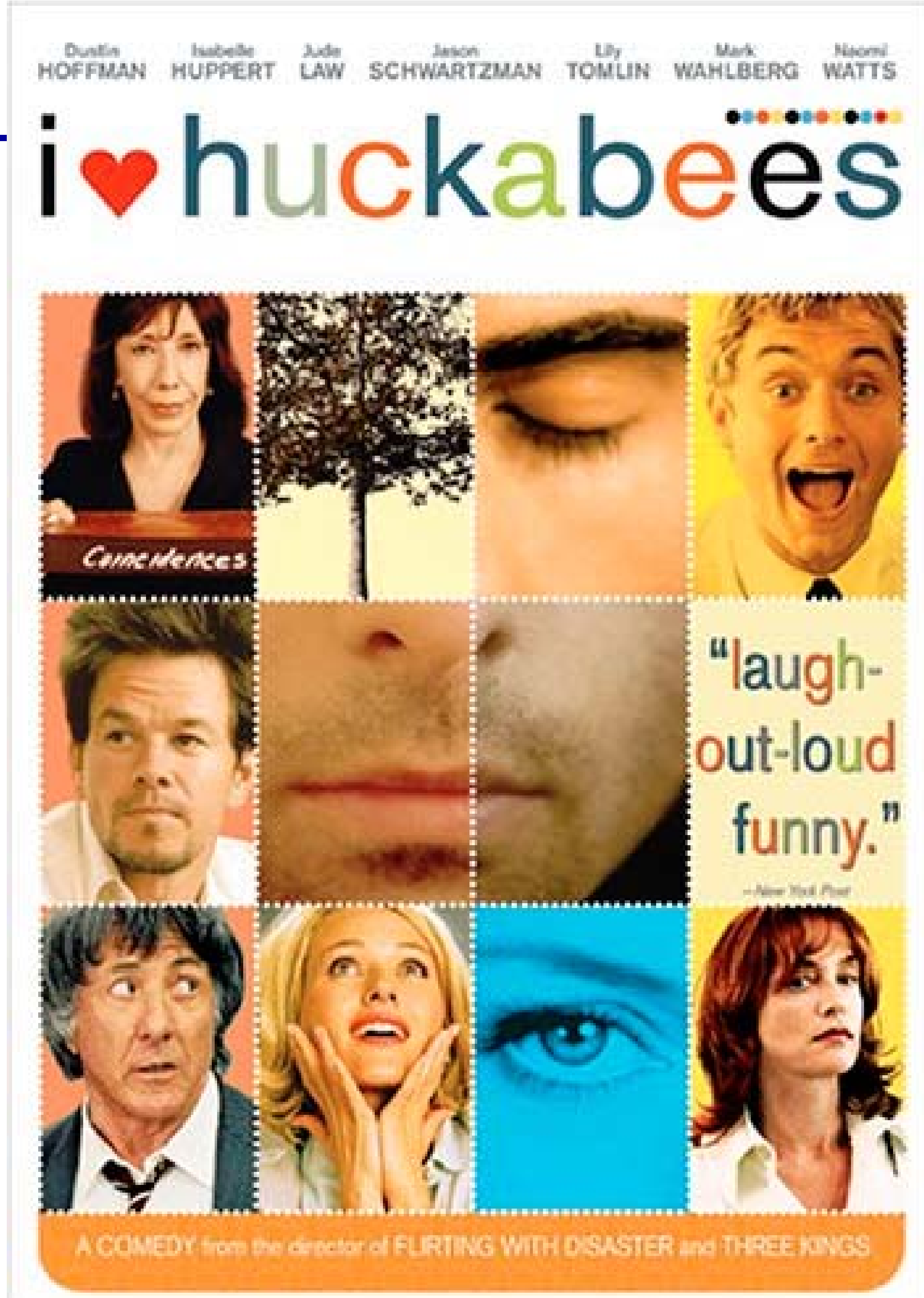
There are currently 50289 contestants on 40922 teams from 185 different countries.
We have received 42524 valid submissions from 4921 different teams; 217 submissions in the last 24 hours.

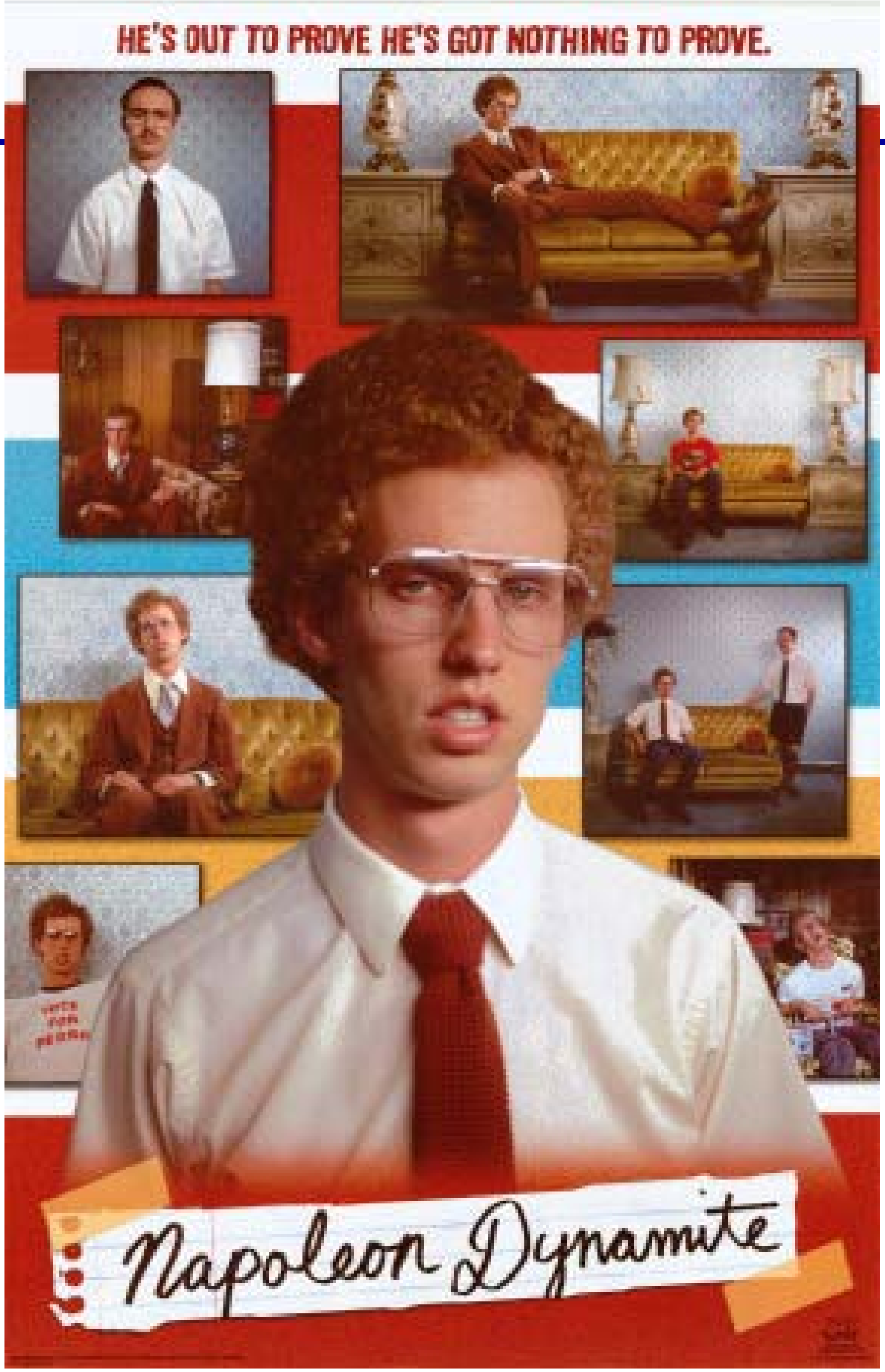
Questions about interpreting the leaderboard? Please read this.


WELCOME TO A WORLD WITHOUT RULES.
CHRISTIAN BALE
MICHAEL CAINE
HEATH LEDGER
GARY OLDMAN
AARON ECKHART
MAGGIE GYLLENHAAL
AND MORGAN FREEMAN
THE DARK KNIGHT
JULY 18
WWW.THEDARKKNIGHT.COM
SYNCOPY

Disney · PIXAR
WALL·E
© Disney


MIKE MYERS
THE
LOVE
GURU™
HIS KARMA IS HUGE
JUNE 20
TheLoveGuru.com


Dustin HOFFMAN
Isabelle HUPPERT
Jude LAW
Jason SCHWARTZMAN
Lily TOMLIN
Mark WAHLBERG
Naomi WATTS
i ♥ huckabees
Coincidences
"laugh-out-loud funny."
—New York Post
A COMEDY from the director of FLIRTING WITH DISASTER and THREE KINGS


HE'S OUT TO PROVE HE'S GOT NOTHING TO PROVE.
Napoleon Dynamite

How do you rate a movie?

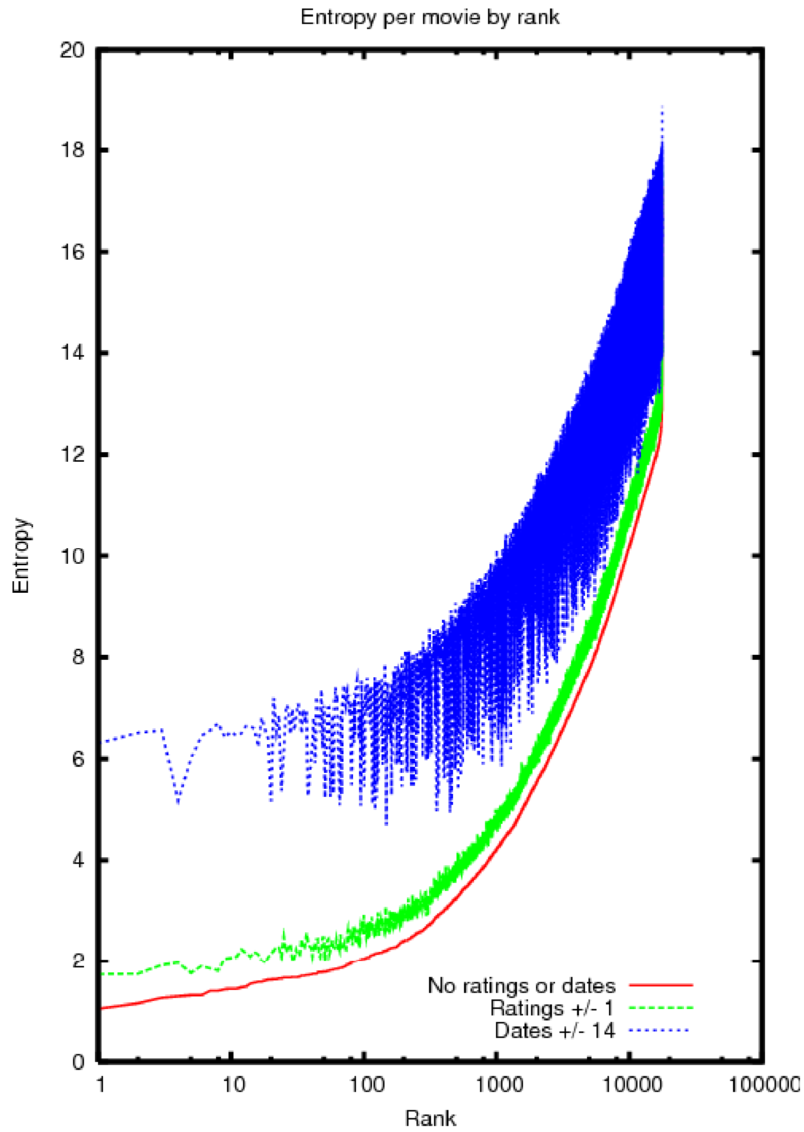
- Report global average
  - I predict you will rate this movie 3.6 (1-5 scale)
  - Algorithm is 15% worse than Cinematch
- Report movie average (Movie effects)
  - Dark knight: 4.3
  - Wall-E: 4.2
  - The Love Guru: 2.8
  - I heart Huckabees: 3.2
  - Napoleon Dynamite: 3.4
  - Algorithm is 10% worse than Cinematch

How do you rate a movie?

- Report global average [-15%]
- Report movie average (Movie effects) [-10%]
- User effects
  - Find each user’s average
  - Subtract average from each rating
  - Corrects for curmudgeons and Pollyannas
- Movie + User effects is 5% worse than Cinematch
- More sophisticated techniques use covariance matrix

Netflix Dataset: Attributes

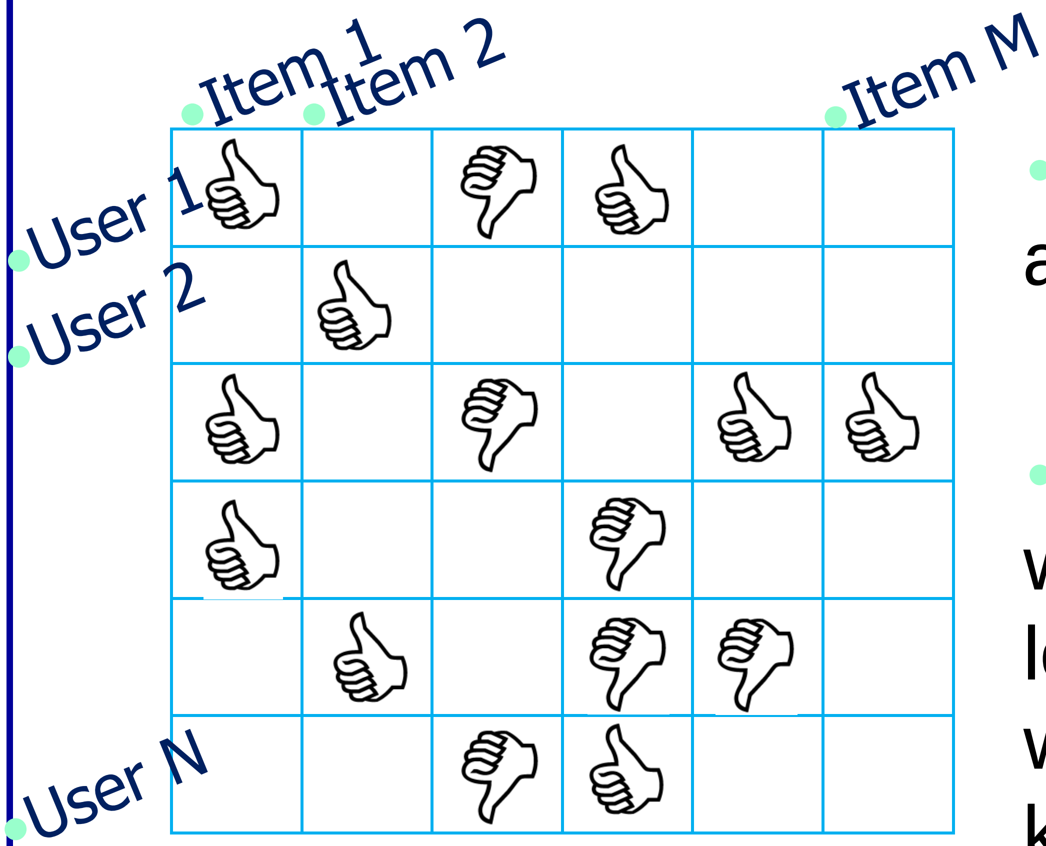
- Most popular movie rated by almost half the users!
- Least popular: 4 users
- Most users rank movies outside top 100/500/1000



Confounding prediction

- Some movies are quirky
  - I Heart Huckabees
  - Napoleon Dynamite
  - Lost In Translation
  - These movies have intermediate average, but high standard deviation
- Users polarize on these movies
- Lovers and Haters hard to determine
  - The Dark Knight might predict X-men II
  - Hard to find predictors for some movies
- Maybe use social networks to weight ratings

Why is Netflix database private?



- Provides some anonymity

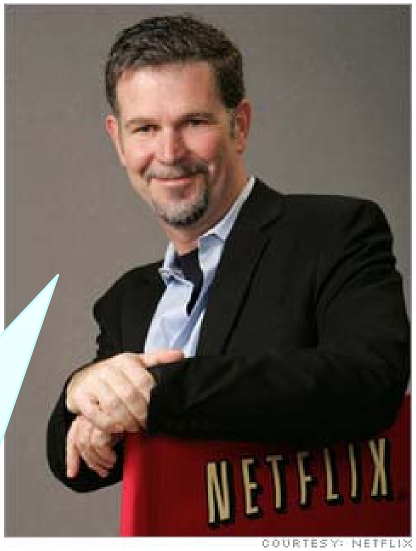
- Privacy question: what can the adversary learn by combining with background knowledge?

- No explicit identifiers

Netflix's Take on Privacy

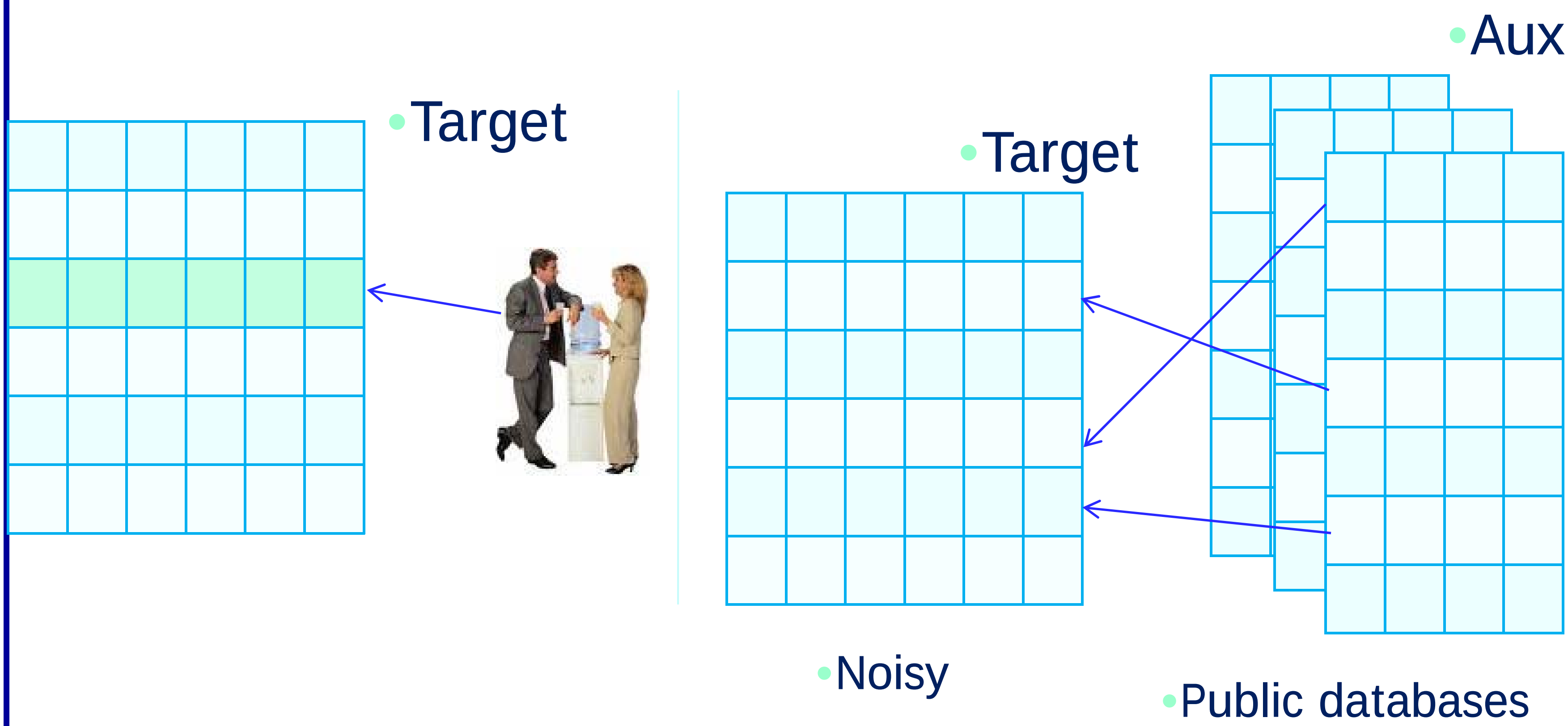
Even if, for example, you knew all your own ratings and their dates you probably couldn't identify them reliably in the data because only a small sample was included (less than one-tenth of our complete dataset) and that data was subject to perturbation. Of course, since you know all your own ratings that really isn't a privacy problem is it?

-- Netflix Prize FAQ



Background Knowledge (Aux. Info.)

Information available to adversary outside of normal data release process



De-anonymization Objective

- Fix some target record r in the original dataset
- Goal: learn as much about r as possible
- Subtler than “find r in the released database”

- Background knowledge is noisy
- Released records may be perturbed
- Only a sample of records has been released
- False matches

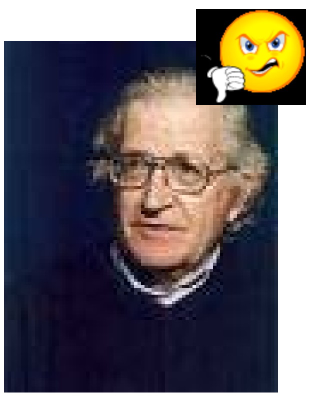
Narayanan & Shmatikov 2008

Earth's Biggest Movie Database

martinwilliamrandall	
Email	martinwilliamrandall@yahoo.co.uk
Biography	i went to st peters & st pauls primary from 1982 to 1985

Using IMDb as Aux

- Extremely noisy, some data missing
- Most IMDb users are not in the Netflix dataset
- Here is what we learn from the Netflix record of one IMDb user (not in his IMDb profile)




De-anonymizing the Netflix Dataset

- Average subscriber has 214 dated ratings
- Two is enough to reduce to 8 candidate records
- Four is enough to identify uniquely (on average)
- Works even better with relatively rare ratings
    - “The Astro-Zombies” rather than “Star Wars”

- Fat Tail effect helps here:
- most people watch obscure movies (really!)

More linking attacks

- Profile 1 in IMDb

- Profile 2 in AIDS survivors online

Anonymity vs. Privacy

- Anonymity is insufficient for privacy

- Anonymity is necessary for privacy

- Anonymity is unachievable in practice

- Re-identification attack → anonymity breach → privacy breach

  - Just ask Justice Scalia
    “It is silly to think that every single
    datum about my life is private”

Beyond recommendations…

- Adaptive systems reveal information about users

Outline

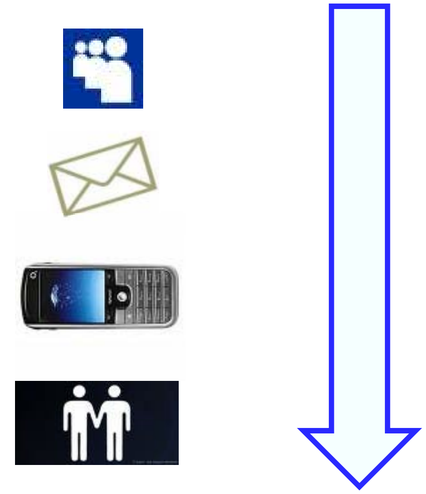
- Motivation
- Background & definitions for security
  - Cryptographic operations for security
- Netflix deanonymization attack
- Anonymity and privacy of social networks

  The problem
- Just a touch of cloud computing
- Mandatory access control
- Differential privacy – interactive privacy

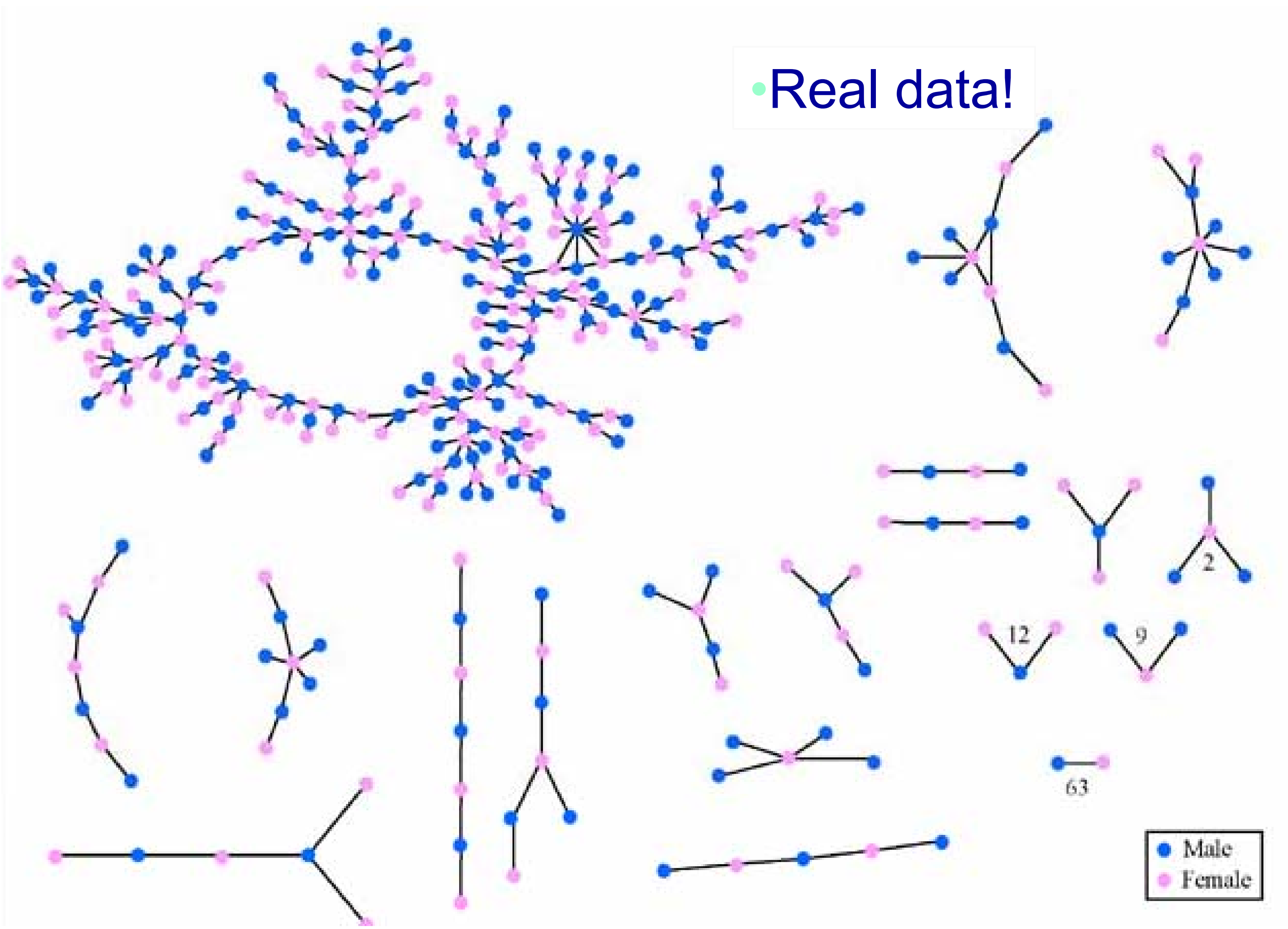
Social Networks

- Online social network services
- Email, instant messenger
- Phone call graphs
- Plain old real-life relationships

- Sensitivity



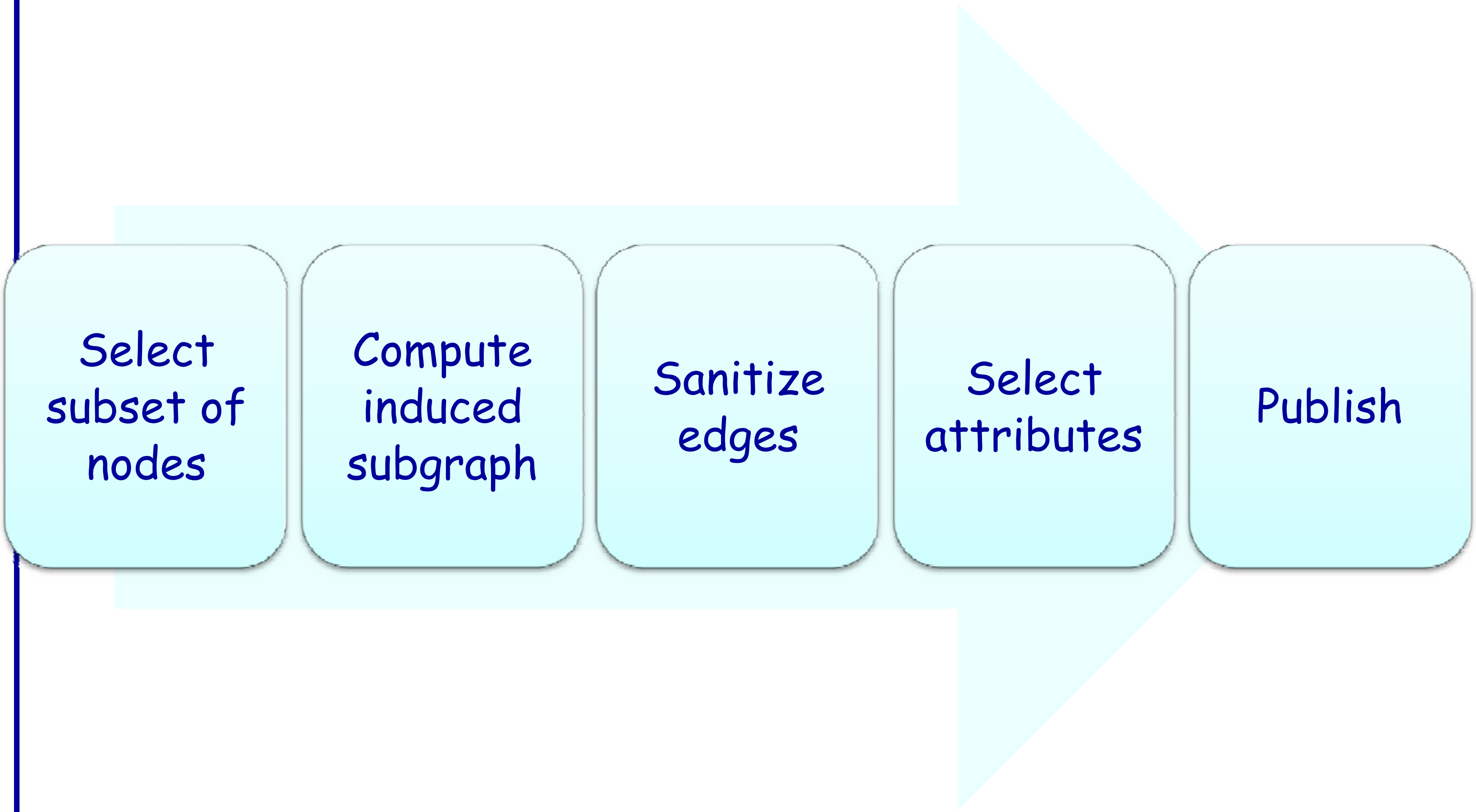
“Jefferson High”: Romantic and Sexual Network



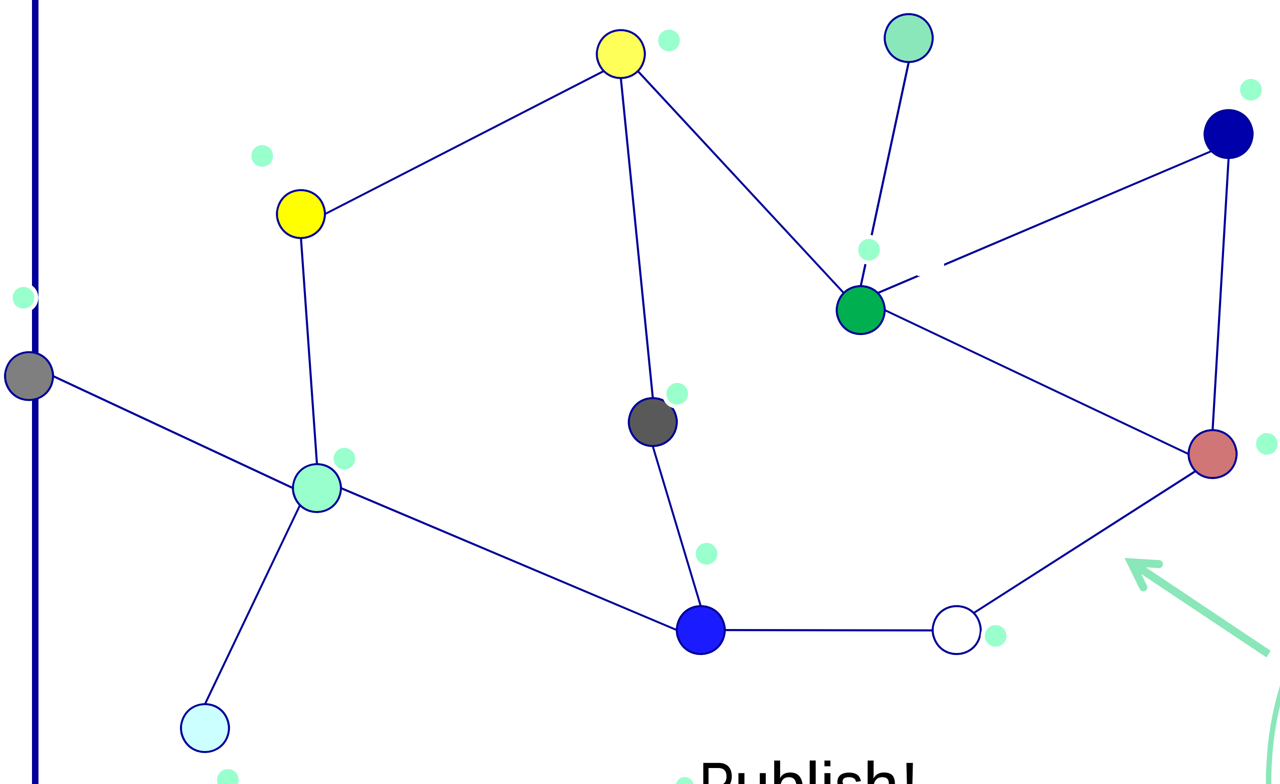
"Jefferson High" romantic dataset

- James Moody at Ohio State
- 1,000 students over 18 months in 1995
  - 537 were sexually active (those were graphed)
- Network is like rural phone lines
  - Main trunk line to individual houses
  - Many adult sexual networks are hub & spoke
  - Easier to control disease without hubs
- One component links 288 students (52%)
  - But 37 degrees of separation maximum
- 63 simple pairs
- Little cycling
  - No "sloppy seconds"

Social Networks: Data Release



Attack Model



- Publish!

- Large-scale
- Background
- Knowledge

Motivating Scenario: Overlapping Networks

- Social networks A and B have overlapping memberships
- Owner of A releases anonymized, sanitized graph
  - say, to enable targeted advertising
- Can owner of B learn sensitive information from released graph A'?

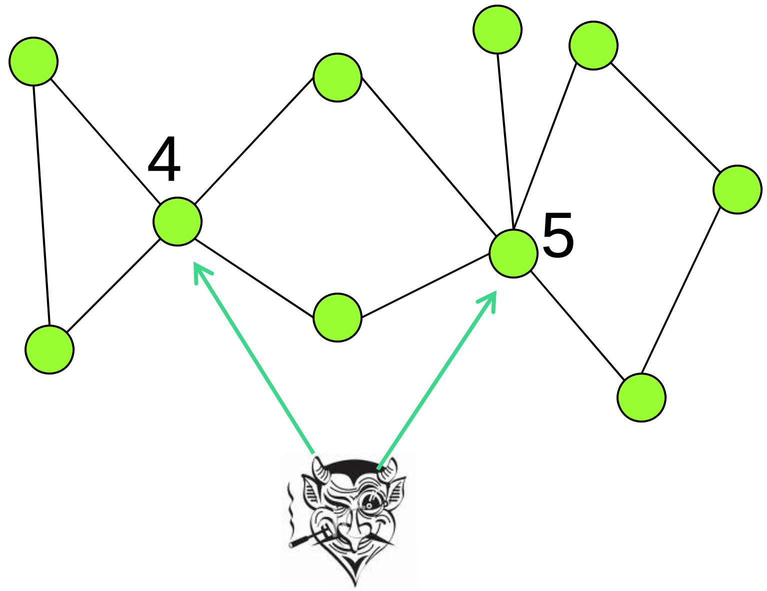
Re-identification: Two-stage Paradigm

Re-identifying target graph =
Mapping between Aux and target nodes

- Seed identification:
  - Detailed knowledge about small number of nodes
  - Relatively precise
  - Link neighborhood constant
  - In my top 5 call and email list…..my wife
- Propagation: similar to infection model
  - Successively build mappings
  - Use other auxiliary information
    - I’m on facebook and flickr from 8pm-10pm
- Intuition: no two random graphs are the same
  - Assuming enough nodes, of course

Seed Identification: Background Knowledge

- How:
  - Creating sybil nodes
  - Bribing
  - Phishing
  - Hacked machines
  - Stolen cellphones



What: List of neighbors

- Degree
- Number of common neighbors of two nodes

- Degrees: (4,5)
- Common nbrs: (2)

Preliminary Results

- Datasets: flickr twitter
- 27,000 common nodes
- Only 15% edge overlap
- 150 seeds
- 32% re-identified as measured by centrality
  - 12% error rate

How do I view the web?

- Everything you put on the web is
  - Permanent
  - Public
- Check out my embarassing question on comp.lang.perl in 1994

Outline

- Motivation
- Background & definitions for security
  - Cryptographic operations for security
- Netflix deanonymization attack
- Anonymity and privacy of social networks
- Just a touch of cloud computing
- Mandatory access control
- Differential privacy – interactive privacy

The problem

Potential solutions

What is cloud computing?

- **Cloud computing** is where dynamically scalable and often virtualized resources are provided as a service over the Internet (thanks, wikipedia!)
- Infrastructure as a service (IaaS)
  - Amazon's EC2 (elastic compute cloud)
- Platform as a service (PaaS)
  - Google gears
  - Microsoft azure
- Software as a service (SaaS)
  - gmail
  - facebook
  - flickr

Services Economies of Scale

- Substantial economies of scale possible
- 2006 comparison of very large service with small/mid-sized: (~1000 servers):

- High cost of entry
  - Physical plant expensive: 15MW roughly $200M
- Summary: significant economies of scale but at very high cost of entry
  - Small number of large players likely outcome

- Thanks, James Hamilton, amazon

Services Different from Enterprises

- **Enterprise Approach:**
  - Largest cost is people -- scales roughly with servers (~100:1 common)
  - Enterprise interests center around consolidation & utilization
    - Consolidate workload onto fewer, larger systems
    - Large SANs for storage & large routers for networking

- **Internet-Scale Services Approach:**
  - Largest costs is server & storage H/W
    - Typically followed by cooling, power distribution, power
    - Networking varies from very low to dominant depending upon service
    - People costs under 10% & often under 5% (>1000+:1 server:admin)
  - Services interests center around work-done-per-$ (or joule)

- **Observations:**
  - People costs shift from top to nearly irrelevant.
  - Expect high-scale service techniques to spread to enterprise
  - Focus instead on work done/$ & work done/joule

Thanks, James Hamilton, amazon

Outline

- Motivation
- Background & definitions for security
  - Cryptographic operations for security
- Netflix deanonymization attack
- Anonymity and privacy of social networks
- Just a touch of cloud computing
- Mandatory access control
- Differential privacy – interactive privacy

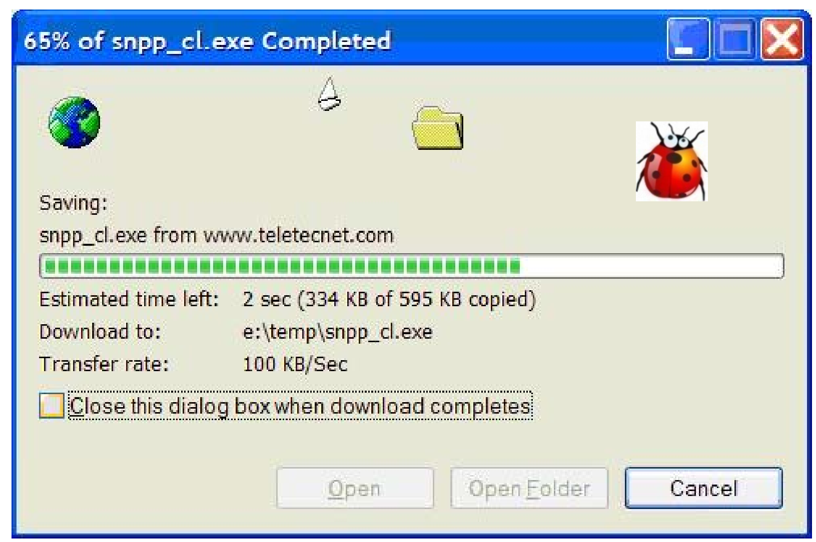
Potential solutions

Mandatory access control (MAC)

- System-wide, enforced rules on data propagation
- Problem with discretionary access control
  - I give permission to alice to read my data
  - Now Alice can do anything with my data!
  - Make a deal with the Chinese
- Facebook third party applications
  - The Facebook Platform Developer Terms of Service prohibit third party applications from storing certain information for longer than 24 hours, and Facebook takes action on developers who are found to be violating this.
- MAC prevents transitive data leaks

Untrusted code on trusted data

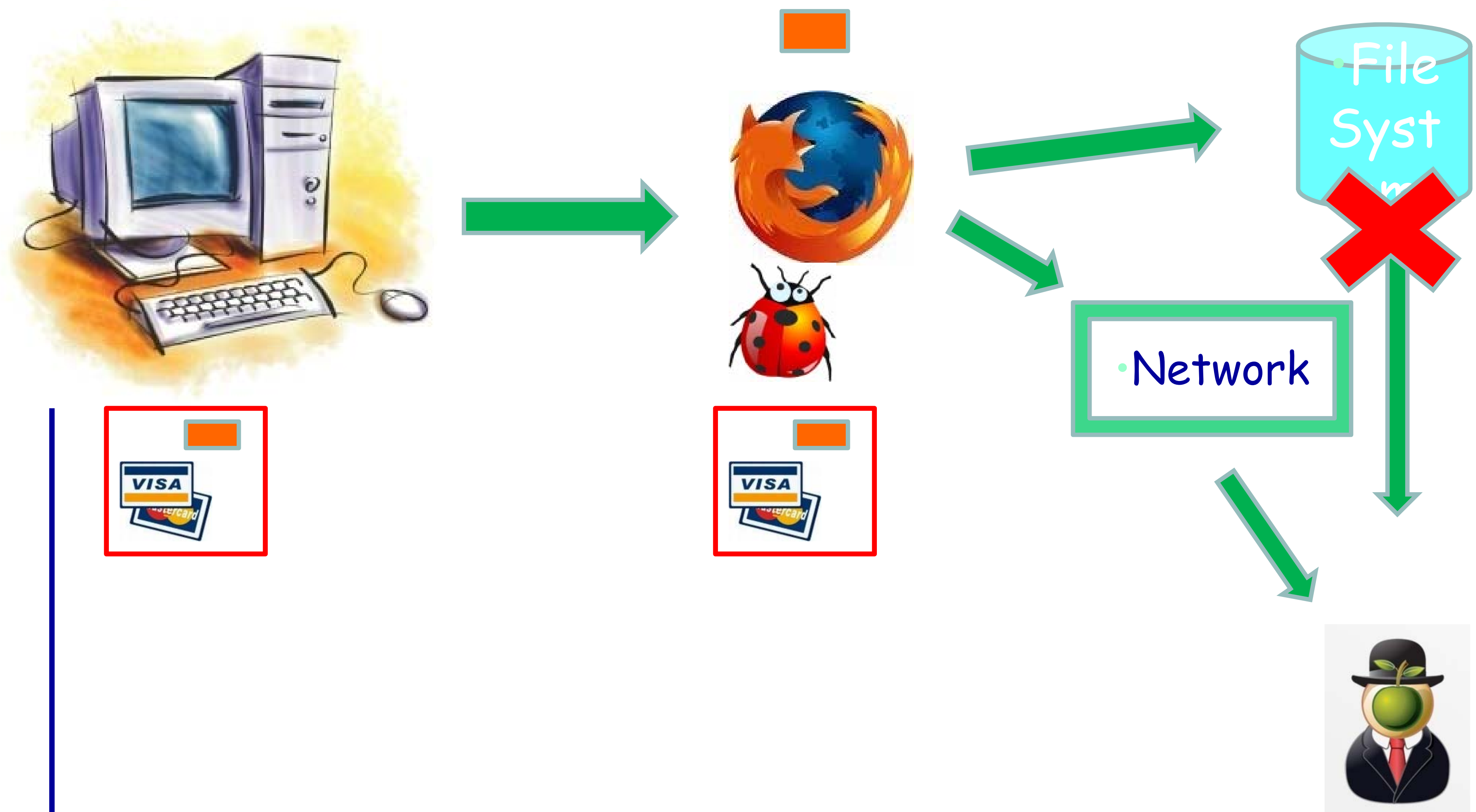
- Your computer holds trusted and sensitive data
  - Credit card number, SSN, personal calendar…
- But not every program you run is trusted
  - Bugs in code, malicious plugins…



Security model

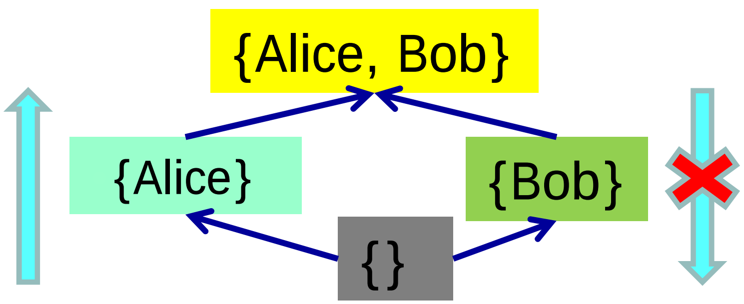
- Decentralized Information Flow Control (DIFC) [Myers and Liskov '97]
  - An example of a mandatory access control system
- Associate labels with the data
- System tracks the flow of data and the labels
- Access and distribution of data depends on labels
  - Firefox may read the credit card number
  - But firefox may **not** send it to the outside world

Control thy data (and its fate)



DIFC Implementation

- How do we rethink and rewrite code for security?
  - Hopefully not many changes…
- Users create a lattice of labels
- Associate labels with the data-structure

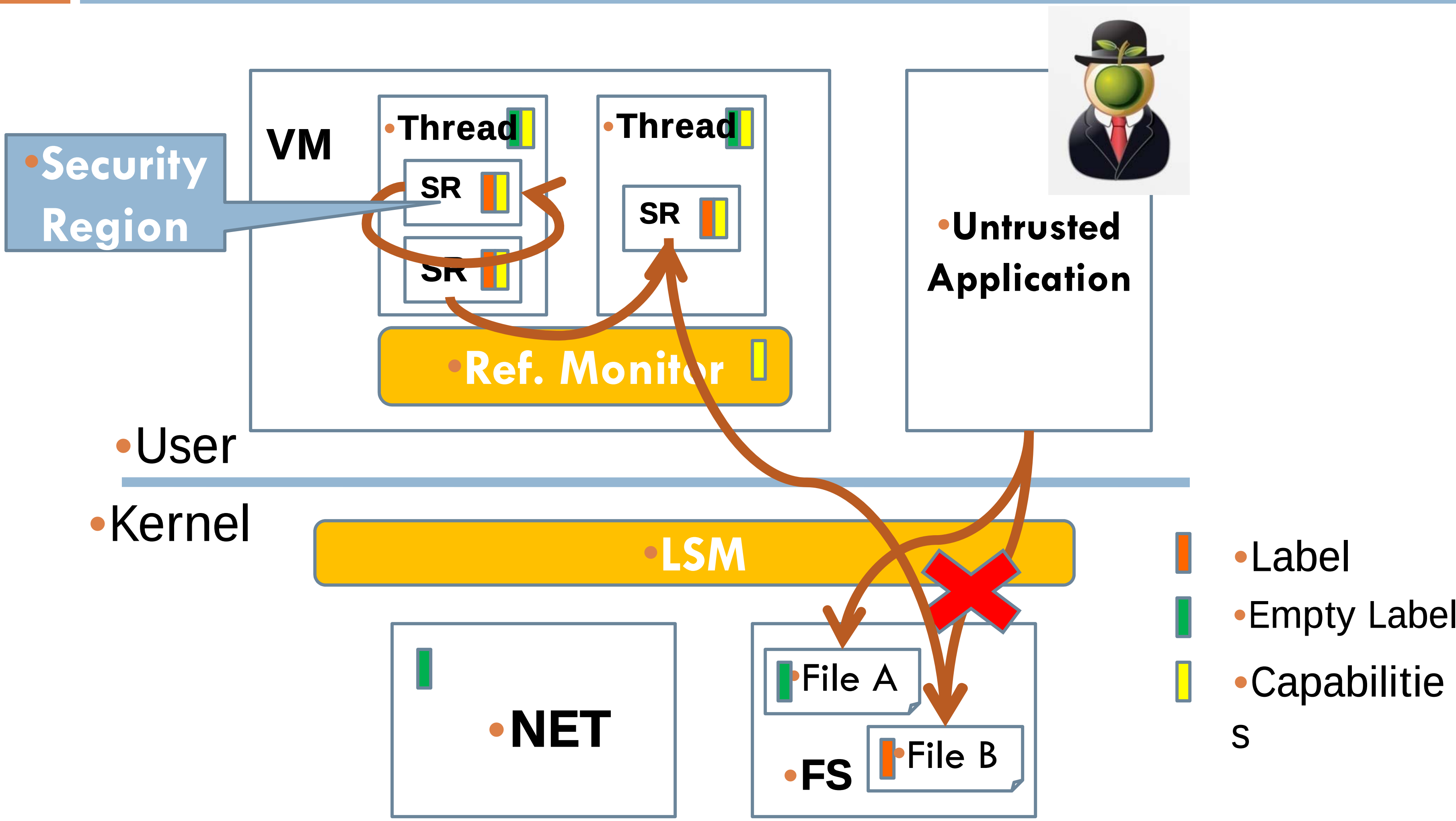


• Information flow in a lattice

User	Mon.	Tue.	Wed.
Alice	Watch game	Office work	Free
Bob	Free	Meet doctor	Free

• Calendar data structure

Security checks: example

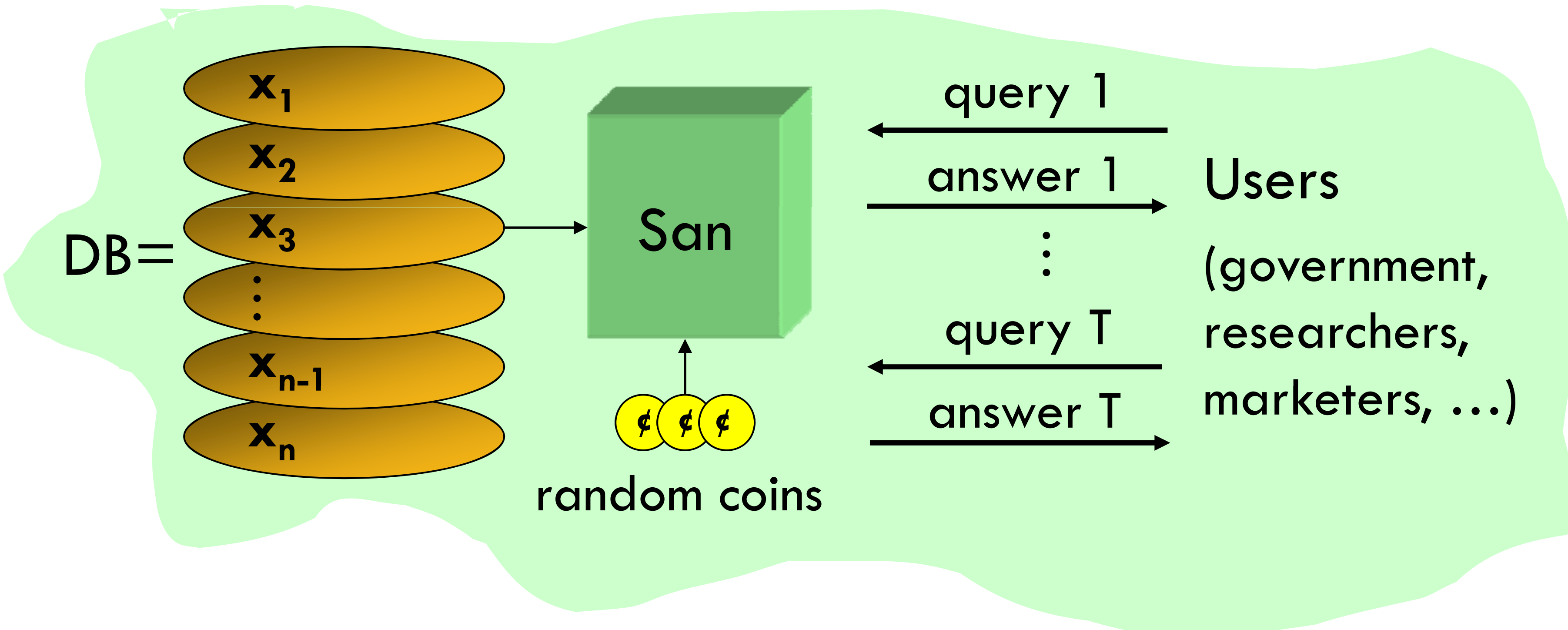


Outline

- Motivation
- Background & definitions for security
  - Cryptographic operations for security
- Netflix deanonymization attack
- Anonymity and privacy of social networks
- Just a touch of cloud computing
- Mandatory access control
- Differential privacy – interactive privacy

Potential solutions

Basic Setting



Examples of Sanitization Methods

- Input perturbation
  - Add random noise to database, release
- Summary statistics
  - Means, variances
  - Marginal totals
  - Regression coefficients
- Output perturbation
  - Summary statistics with noise
- Interactive versions of the above methods
  - Auditor decides which queries are OK, type of noise

Classical Intution for Privacy

- "If the release of statistics S makes it possible to determine the value [of private information] more accurately than is possible without access to S, a disclosure has taken place." [Dalenius 1977]
  - Privacy means that anything that can be learned about a respondent from the statistical database can be learned without access to the database
- Similar to semantic security of encryption
  - Anything about the plaintext that can be learned from a ciphertext can be learned without the ciphertext

Problems with Classic Intuition

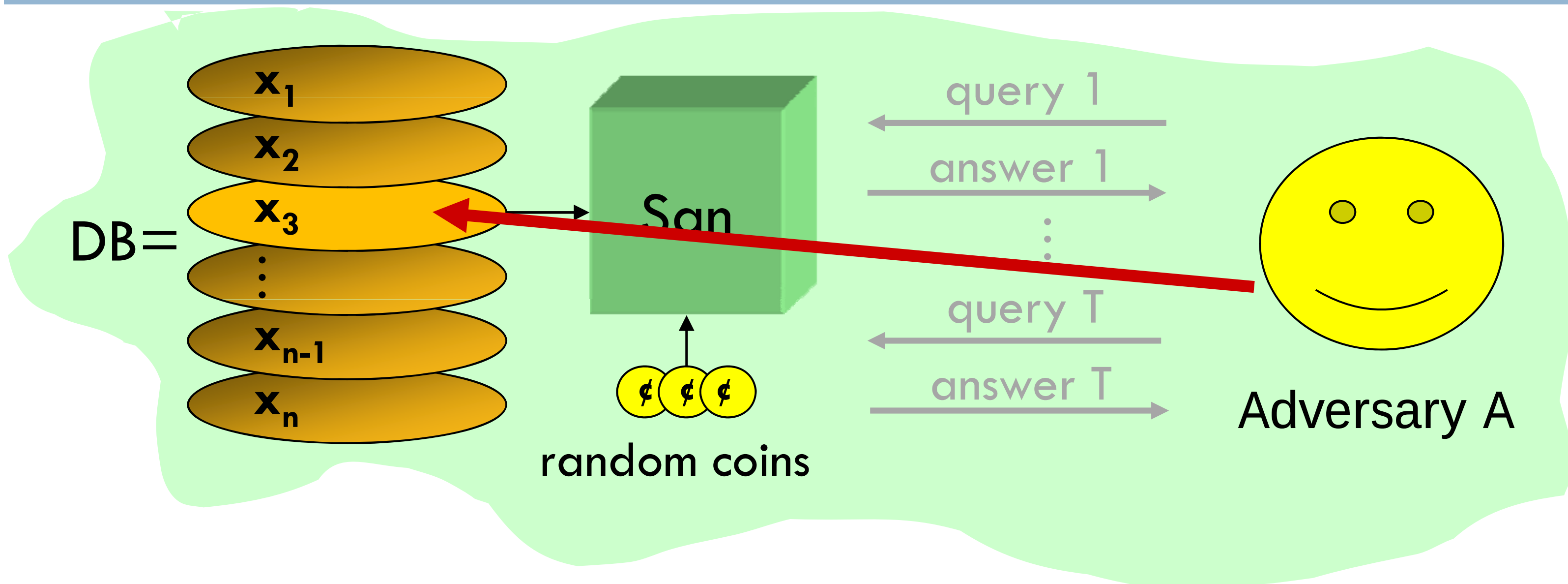
- Popular interpretation: prior and posterior views about an individual shouldn't change "too much"
  - What if my (incorrect) prior is that every UTCS graduate student has three arms?
- How much is "too much?"
  - Can't achieve cryptographically small levels of disclosure <u>and</u> keep the data useful
  - Adversarial user is <u>supposed</u> to learn unpredictable things about the database

Impossibility Result

[Dwork]

- Privacy: for some definition of "privacy breach," $\forall$ distribution on databases, $\forall$ adversaries A, $\exists$ A' such that $\Pr(A(San)=breach) - \Pr(A'()=breach) \leq \varepsilon$
  - For reasonable "breach", if San(DB) contains information about DB, then some adversary breaks this definition
- Example
  - Vitaly knows that Josh Leners is 2 inches taller than the average Russian
  - DB allows computing average height of a Russian
  - This DB breaks Josh's privacy according to this definition… even if his record is not in the database!

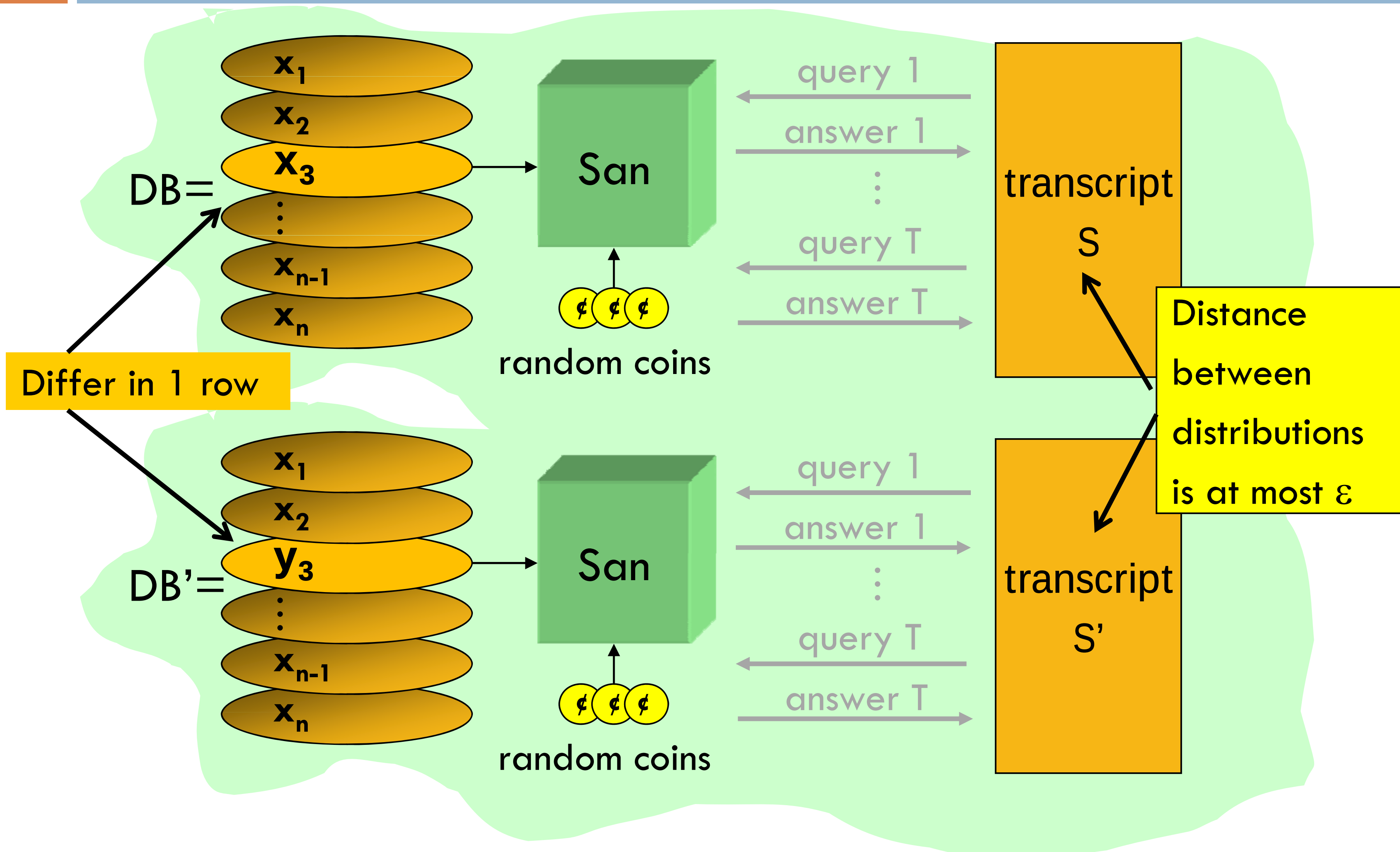
Differential Privacy (1)



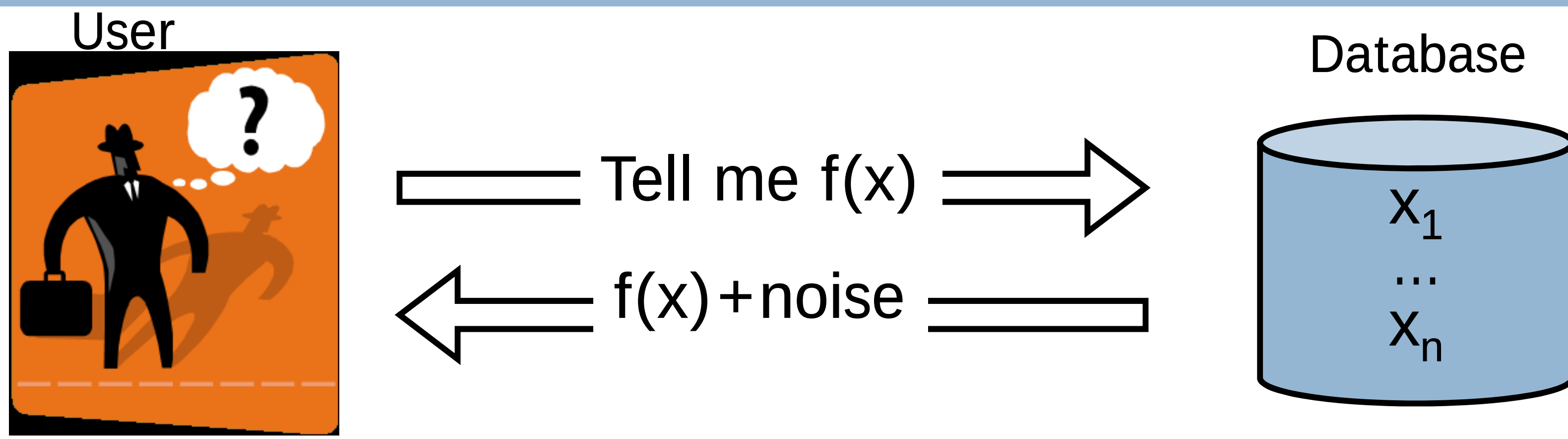
- Example with Russians and Josh Leners
  - Adversary learns Josh's height even if he is not in the database
- Intuition: "Whatever is learned would be learned regardless of whether or not Josh participates"
  - Dual: Whatever is already known, situation won't get worse

Indistinguishability

slide 74



Diff. Privacy in Output Perturbation



- Intuition: f(x) can be released accurately when f is insensitive to individual entries $x_1, \ldots x_n$
- Global sensitivity $GS_f = \max_{\text{neighbors } x,x'} ||f(x) - f(x')||_1$
  - Example: $GS_{average} = 1/n$ for sets of bits

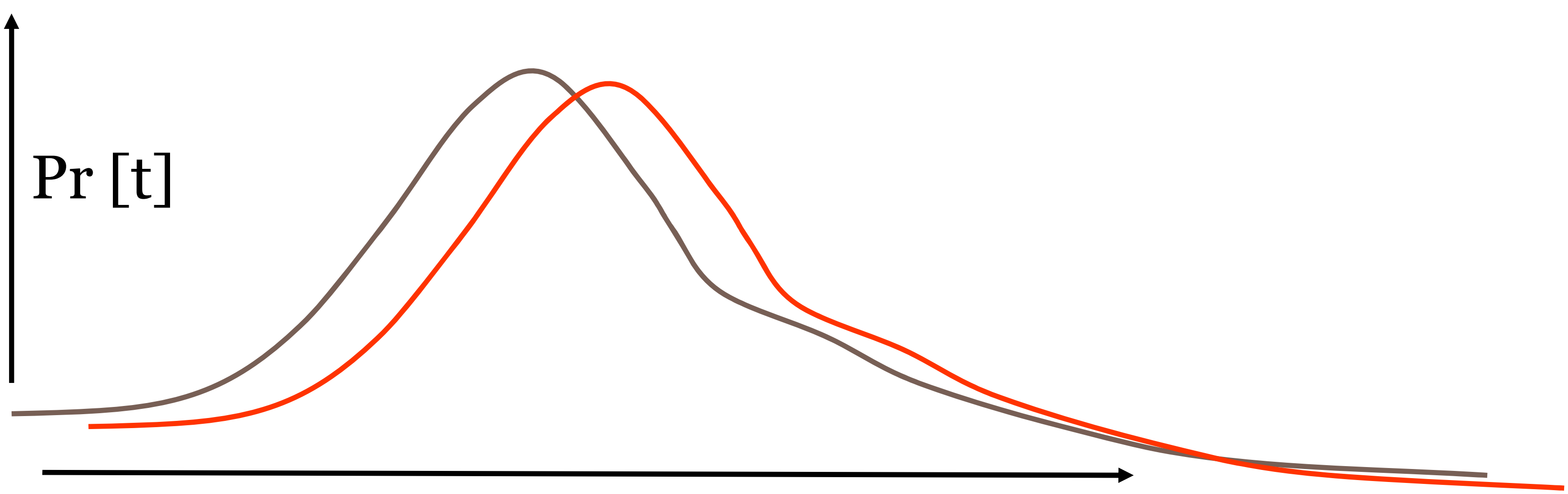
Lipschitz constant of f

- Theorem: $f(x) + Lap(GS_f / \varepsilon)$ is ε-indistinguishable
  - Noise generated from Laplace distribution

Differential Privacy: Summary

- K gives ε-differential privacy if for all values of DB and Me and all transcripts t:

$$\frac{\Pr[\ \mathcal{K}(DB - Me) = t]}{\Pr[\ \mathcal{K}(DB + Me) = t]} \leq e^{\varepsilon} \approx 1 \pm \varepsilon$$



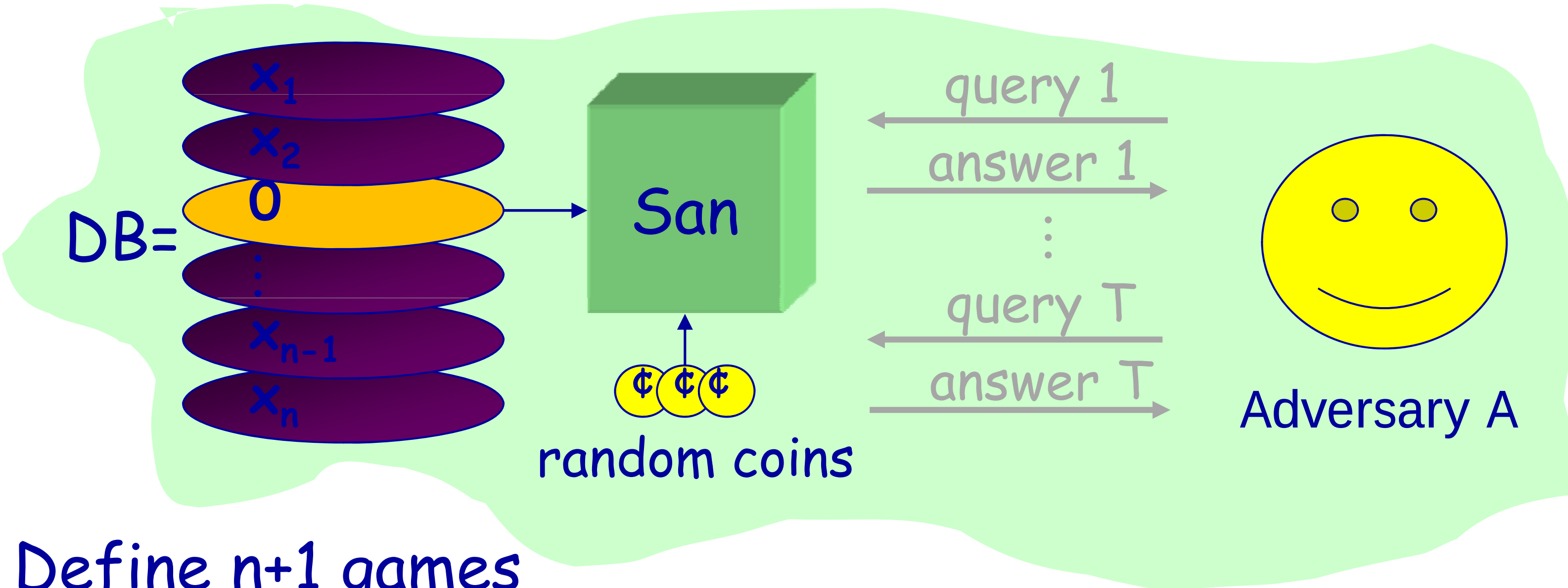
Please teach the mindset of debugging

- Contrary to assignments, programs are rarely finished
  - Specifications are unclear
  - Specifications change
- Students view getting a program right
  - Write code
  - Compile it
  - Does it work in 1 case? If yes, then done, else s9tep 1
- Debugging != Debugger

Thank you & Thanks for your work!

Differential Privacy (2)

slide 78

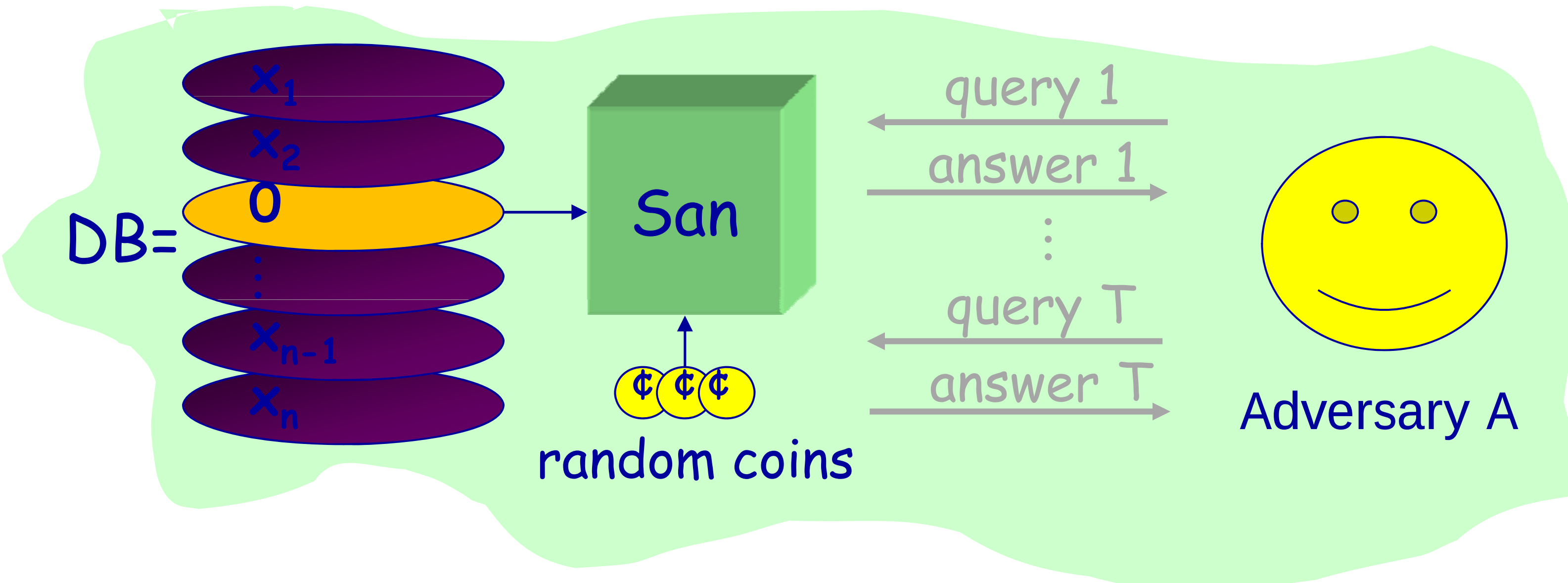


- Define n+1 games
  - Game 0: Adv. interacts with San(DB)
  - Game i: Adv. interacts with San(DB_{-i}); $DB_{-i} = (x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_n)$
- Given S and prior p() on DB, define n+1 posterior distrib's

$$p_i(DB|S) = p(DB|S \text{ in Game } i) = \frac{p(San(DB_{-i}) = S) \times p(DB)}{p(S \text{ in Game } i)}$$

Differential Privacy (3)

slide 79



Definition: San is safe if

∀ prior distributions p(¢) on DB,

∀ transcripts S, ∀ i =1,…,n

StatDiff(p_0(¢|S) , p_i(¢|S)) ≤ ε