Lecture 16:
CS 5306 / INFO 5306:
Crowdsourcing and Human Computation

Netflix Prize Postscript: 2007

Robust De-anonymization of Large Sparse Datasets

Arvind Narayanan and Vitaly Shmatikov

The University of Texas at Austin

Abstract

We present a new class of statistical de-anonymization attacks against high-dimensional micro-data, such as individual preferences, recommendations, transaction records and so on. Our techniques are robust to perturbation in the data and tolerate some mistakes in the adversary's background knowledge.

We apply our de-anonymization methodology to the Netflix Prize dataset, which contains anonymous movie ratings of 500,000 subscribers of Netflix, the world's largest online movie rental service. We demonstrate that an adversary who knows only a little bit about an individual subscriber can easily identify this subscriber's record in the dataset. Using the Internet Movie Database as the source of background knowledge, we successfully identified the Netflix records of known users, uncovering their apparent political preferences and other potentially sensitive information.

1 Introduction

and sparsity. Each record contains many attributes (*i.e.*, columns in a database schema), which can be viewed as dimensions. Sparsity means that for the average record, there are no "similar" records in the multi-dimensional space defined by the attributes. This sparsity is empirically well-established [7, 4, 19] and related to the "fat tail" phenomenon: individual transaction and preference records tend to include statistically rare attributes.

Our contributions. Our first contribution is a formal model for privacy breaches in anonymized micro-data (section 3). We present two definitions, one based on the probability of successful de-anonymization, the other on the amount of information recovered about the target. Unlike previous work [25], we do not assume *a priori* that the adversary's knowledge is limited to a fixed set of "quasi-identifier" attributes. Our model thus encompasses a much broader class of de-anonymization attacks than simple cross-database correlation.

Our second contribution is a very general class of de-anonymization algorithms, demonstrating the fundamental limits of privacy in public micro-data (section 4). Under very mild assumptions about the distribution from

Netflix Prize Postscript: 2009

NETFLIX

Press Releases

Current Press Releases

Netflix Awards $1 Million Netflix Prize and Announces Second $1 Million Challenge

NEW YORK, Sept. 21 /PRNewswire/ -- After almost three years and submissions by more than 40,000 teams from 186 countries, Netflix, Inc., the world's largest online movie rental service (NASDAQ: NFLX), today awarded the $1 million Netflix Prize to a team of engineers, statisticians and researchers who achieved the competition's goal of a 10 percent improvement over the accuracy of the Netflix movie recommendation system when the competition was launched in Oct. 2006. Netflix members already are benefiting from improvements Netflix Prize contestants have contributed to the recommendations system.

Moments after bestowing the $1 million prize, Netflix announced a second $1 million challenge, asking the world's computer science and machine learning communities to keep the improvements coming.

The team "BellKor's Pragmatic Chaos," the merging of three teams that had previously competed against one another in the contest, received the $1 million Netflix Prize in an award ceremony hosted here today by Netflix Co-Founder and CEO Reed Hastings and Chief Product Officer Neil Hunt.

"We had a bona fide race right to the very end," said Mr. Hastings. "Teams that had previously battled it out independently joined forces to surpass the 10 percent barrier. New submissions arrived fast and furious in the closing hours and the competition had more twists and turns than 'The Crying Game,' 'The Usual Suspects' and all the 'Bourne' movies wrapped into one."

The winning team is comprised of software and electrical engineers, statisticians and machine learning researchers from Austria, Canada, Israel and the United States. All seven team members - Bob Bell, Martin Chabbert, Michael Jahrer, Yehuda Koren, Martin Piotte, Andreas Toscher and Chris Volinsky - attended the awards ceremony. It was the first time all seven had met one another in person. How the $1

MEDIA CENTER

Press Kit
Press Releases
Management
Image Library
Upcoming Events
Investor Relations
Contact Public Relations
RSS
◀ Back to Homepage

Netflix Prize Postscript: 2009

NETFLIX

Press Releases

MEDIA CENTER

Current P...

...eases

...ent

...ary

...Events

...lations

...ublic Relations

...omepage

Netflix Award...

NEW YORK, Sep...
teams from 186 ...
today awarded t...
achieved the co...
recommendation...
benefiting from i...

Moments after b...
world's compute...

The team "BellK...
one another in t...
by Netflix Co-Fou...

The new data set, providing more than 100 million data points, will include, among other things, information about renters' ages, genders, ZIP codes, genre ratings and previously chosen movies. As with the first Netflix Prize, all data provided is anonymous and cannot be associated with a specific Netflix member.

"We had a bona ...ide race right to the very end," said Mr. Hastings. "Teams that had previously battled it out independently joined forces to surpass the 10 percent barrier. New submissions arrived fast and furious in the closing hours and the competition had more twists and turns than 'The Crying Game,' 'The Usual Suspects' and all the 'Bourne' movies wrapped into one."

The winning team is comprised of software and electrical engineers, statisticians and machine learning researchers from Austria, Canada, Israel and the United States. All seven team members - Bob Bell, Martin Chabbert, Michael Jahrer, Yehuda Koren, Martin Piotte, Andreas Toscher and Chris Volinsky - attended the awards ceremony. It was the first time all seven had met one another in person. How the $1

Netflix Prize Postscript: 2009

Scott A. Kamber
David A. Stampley
KamberEdelson, LLC
11 Broadway, 22nd Floor.
New York, NY. 10004
Telephone: (212) 920-3072
skamber@kamberedelson.com
dstampley@kamberedelson.com

Joseph H. Malley
Law Office of Joseph H. Malley
1045 North Zang Blvd
Dallas, TX 75208
Telephone: (214) 943-6100
malleylaw@gmail.com

David Parisi (SBN 162248)
Suzanne Havens Beckman (SBN 188814)
Parisi & Havens LLP
15233 Valleyheart Drive
Sherman Oaks, California 91403
Telephone: (818) 990-1299
dcparisi@parisihavens.com
shavens@parisihavens.com

Attorneys for Plaintiffs

IN THE UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF CALIFORNIA
SAN JOSE DIVISION

JANE DOE, individually; NELLY VALDEZ-MARQUEZ, ANTHONY SINOPOLI, PAUL NAVARRO, individually and on behalf of a class of similarly situated individuals

Plaintiffs,

v.

NETFLIX, INC., a Delaware Corporation, and DOES 1 THROUGH 50, inclusive,

Defendants.

C09 05903 JW PVT

JURY DEMAND

CLASS ACTION COMPLAINT FOR:

1) Video Privacy Protection Act, 18 U.S.C. § 2710
2) Video Privacy Protection Act, 18 U.S.C. § 2710
3) California Consumers Legal Remedies Act, Civil Code § 1750
4) California Customer Records Act, Civil Code § 1798.80
5) California Unfair Competition Law, Business and Professions Code § 17200
6) California False Advertising Law, Business and Professions Code § 17500
7) Unjust Enrichment
8) Public Disclosure of Private Facts

FILED

2009 DEC 17 A 10: 16

RICHARD W. WIEKING
CLERK, U.S. DISTRICT COURT
N.D. CAL. SAN JOSE

ADR

E-filing

Netflix Prize Postscript: 2009

I.   NATURE OF THE ACTION

1.   On October 2, 2006, Netflix perpetrated the largest voluntary privacy breach to date, disclosing sensitive and personal indentifying consumer information. The information was not compromised by malicious intruders. Rather, it was given away to the world freely, and with fanfare, as part of a contest intended to benefit its trusted custodian, Netflix.

2.   This right to privacy does not appear to be significant to Netflix. This lawsuit is brought as a class action by and on behalf of similarly situated Netflix subscribers, qualified by the class definition and class period, whose privacy was violated by the actions of Netflix, Inc., ("Netflix") pursuant to their contest, "Netflix Prize." Jane Doe, a lesbian, who does not want her sexuality nor interests in gay and lesbian themed films broadcast to the world, seeks ano-nymity in this action. Paul Navarro files this action to prevent Netflix from going through with its announced intentions to make additional disclosures of personal identifying information in-cluding, but not limited to, users' video renting history and rating habits.

3.   Netflix knowingly authorized, directed, ratified, approved, acquiesced, or par-ticipated in the disclosure to third parties of the sensitive information and/or personal identify-ing information derived from the activity of the Netflix subscribers' online electronic commu-nications, when they accessed the Netflix website to rent and rate videos.

4.   Netflix is an "Electronic Communication Service Provider" to its subscribers and knowingly disclosed to third parties the contents of Netflix's subscribers' communications, including but not limited to, subscribers' rental and rating videos information, while in elec-

Netflix Prize Postscript: 2010

UNITED STATES OF AMERICA
FEDERAL TRADE COMMISSION
WASHINGTON, D.C. 20580

Maneesha Mithal
Associate Director
Division of Privacy & Identity Protection

Direct Dial: 202.326.2771
Fax : 202.326.3062
E-mail: mmithal@ftc.gov

March 12, 2010

BY E-MAIL & FEDERAL EXPRESS

Reed Freeman
Morrison & Foerster LLP
2000 Pennsylvania Ave., NW
Washington, DC 20006

Dear Mr. Freeman:

On October 13, 2009, staff from the FTC's Division of Privacy and Identity Protection contacted your client, Netflix, Inc. ("Netflix"), regarding the privacy implications of Netflix's planned release of customer movie viewing data in connection with the company's efforts to improve its movie recommendation algorithm. Specifically, staff expressed concern that, despite Netflix's efforts to "anonymize" the customer data prior to its release, it would be possible to re-identify specific customers and thereby associate them with their movie viewing histories and preferences.

Staff's concerns about Netflix's planned release stemmed from research published after the

Netflix Prize Postscript: 2010

2006 AOL Data Release

- August 4, 2006
- "anonymized" data for 20 million US-based searches for 658,000 users in March-May 2006 (1/3% of all queries, 1.5% of monthly users)
- Sample searches:
  - Names
  - SSNs
  - "how to kill oneself by natural gas"
  - "child porno"
  - "date rape" "is it normal to cook your friend breakfast after he rapes you" "the morning after being raped" "sexual assault"

2006 AOL Data Release

- NY Times article: "A Face Is Exposed for AOL Searcher No. 4417749"
  - Searches included:
    - "numb fingers", "60 single men", "dog that urinates on everything", "school supplies for Iraq children", "safest place to live", "the best season to visit Italy", "termites", "tea for good health", "mature living", "hand tremors", "nicotine effects on the body", "dry mouth", "bipolar"
    - "landscapers in Lilburn, Ga", "homes sold in shadow lake subdivision gwinnett county Georgia", searches for multiple people with last name Arnold
  - Identified Thelma Arnold, age 62 (with her permission)

2006 AOL Data Release

- Data removed August 7, 2006
- CTO and two others lost their jobs
- AOL Spokesperson:

    This was a screw up, and we're angry and upset about it. It was an innocent enough attempt to reach out to the academic community with new research tools, but it was obviously not appropriately vetted, and if it had been, it would have been stopped in an instant.

    Although there was no personally-identifiable data linked to these accounts, we're absolutely not defending this. It was a mistake, and we apologize. We've launched an internal investigation into what happened, and we are taking steps to ensure that this type of thing never happens again.

- Copies still on the Internet
- Other discoveries…

2006 AOL Data Release

UNITED STATES OF AMERICA
FEDERAL TRADE COMMISSION

_____
)
In the Matter of   )
)
AOL LCC,   )   DOCKET NO.
   a majority-owned subsidiary of   )
   TIME WARNER INC.   )
)
_____   )

REQUEST FOR INVESTIGATION AND
COMPLAINT FOR INJUNCTIVE RELIEF

The Electronic Frontier Foundation ("EFF"), having reason to believe that AOL LLC ("AOL") has violated the Federal Trade Commission Act, and that investigation and injunctive relief is in the public interest, alleges that AOL committed unfair and deceptive trade practices by intentionally and publicly disclosing Internet search histories of more than half a million AOL users. Press reports, the analysis of commentators, and EFF's own research show that these data include sensitive, personal information that can be linked to individuals.

In support of its complaint, EFF alleges as follows.

1. The Electronic Frontier Foundation is a 501(c)(3) nonprofit organization founded in 1990 to protect civil liberties in the digital age. Based in San Francisco, CA, EFF is a membership-supported organization that litigates and educates the public on issues such as free expression, freedom of the press, fair use of copyrighted works, anonymity, security, and privacy as they relate to computing and the Internet.

2. AOL is a Delaware corporation and a majority-owned subsidiary of Time Warner Inc. AOL maintains its principal place of business at 22000 AOL Way, Dulles, VA 20166. Time Warner Inc. is a Delaware corporation and maintains its principal place of business at One Time Warner Center, New York, NY 10019. AOL describes itself to the public as "a Web portal that provides a variety of custom content on top of linking you to relevant information available on the Internet."[1] AOL also provides Internet connectivity options and specialized client software to its registered users.[2] AOL's mission statement says that the company is dedicated "to the simple premise

User 17556639

17556639 how to kill your wife
17556639 how to kill your wife
17556639 wife killer
17556639 how to kill a wife
17556639 poop
17556639 dead people
17556639 pictures of dead people
17556639 killed people
17556639 dead pictures
17556639 dead pictures
17556639 dead pictures
17556639 murder photo

17556639 steak and cheese
17556639 photo of death
17556639 photo of death
17556639 death
17556639 dead people photos
17556639 photo of dead people
17556639 www.murderdpeople.com
17556639 decapatated photos
17556639 decapatated photos
17556639 car crashes3
17556639 car crashes3
17556639 car crash photo

User 672368

- "you're pregnant he doesn't want the baby"

- ...

- "foods to eat when pregnant"

- ...

- "abortion clinics charlotte nc"

- ...

- "can christians be forgiven for abortion"

Others

- "i hate my job"

- "why am i so ugly"

- "i hurt when i think too much i love roadtrips i hate my weight i fear being alone for the rest of my life"

- User 927

Types of Crowdsourcing

- Overt
  - Collecting (Amazon Reviews)
  - Labor Markets (Amazon Mechanical Turk)
  - Collaborative Decisions (Prediction Markets)
  - Collaborative Creation (Wikipedia)
  - Smartest in the Crowd (Contests)
  - Games with a Purpose
- Covert / Crowd Mining
  - Web page linkage, search logs, social media, collaborative filtering

Types of Crowdsourcing

- Overt
  - Collecting (Amazon Reviews)
  - Labor Markets (Amazon Mechanical Turk)
  - Collaborative Decisions (Prediction Markets)
  - Collaborative Creation (Wikipedia)
  - Smartest in the Crowd (Contests)
  - Games with a Purpose
- Covert / Crowd Mining
  - Web page linkage, search logs, social media, collaborative filtering

Types of Crowdsourcing

- Overt
  - Collecting (Amazon Reviews)
  - Labor Markets (Amazon Mechanical Turk)
  - Collaborative Decisions (Prediction Markets)
  - Collaborative Creation (Wikipedia)
  - Smartest in the Crowd (Contests)
  - Games with a Purpose
- Covert / Crowd Mining
  - Web page linkage, search logs, social media, collaborative filtering
- Dark side of crowdsourcing and human intelligence
- Collective intelligence in animals

Types of Crowdsourcing

- Overt
  - Collecting (Amazon Reviews)
  - Labor Markets (Amazon Mechanical Turk)
  - Collaborative Decisions (Prediction Markets)
  - Collaborative Creation (Wikipedia)
  - Smartest in the Crowd (Contests)
  - Games with a Purpose
- Covert / Crowd Mining
  - Web page linkage, search logs, social media, collaborative filtering
- Dark side of crowdsourcing and human intelligence
- Collective intelligence in animals

Games with a Purpose (GWAPs)

- Create a game that people play because they like the game
- As a side-effect of the game, perhaps without the players' awareness, the players are helping solve problems for the game designers

GWAPs: ESPGame

- Defunct
- Videos on YouTube

GWAPs: FoldIt

- 2010
  - "Predicting protein structures with a multiplayer online game" Seth Cooper, Firas Khatib, Adrien Treuille, Janos Barbero, Jeehyung Lee, Michael Beenen, Andrew Leaver-Fay, David Baker and Foldit players, *Nature 44*p . 756-760 (05 August 2010).

- 2011
  - Structure of Mason-Pfizer monkey virus (M-PMV) retroviral protease (AIDS-causing monkey virus)
  - Unsolved for 15 years
  - Solved in 10 days
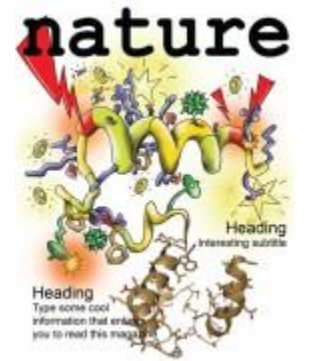
GWAPs: EteRNA

- Solve puzzles related to RNA folding
- 100,000 registered players
- 2014:
  - "RNA design rules from a massive open laboratory"
    Lee, Jeehyung; Kladwang, Wipapat; Lee, Minjae; Cantu, Daniel;
    Azizyan, Martin; Kim, Hanjoo; Limpaecher, Alex; Yoon, Sungroh;
    Treuille, Adrien; Das, Rhiju; EteRNA participants. *PNAS*: Jan 17, 2014
- Feb 2016:
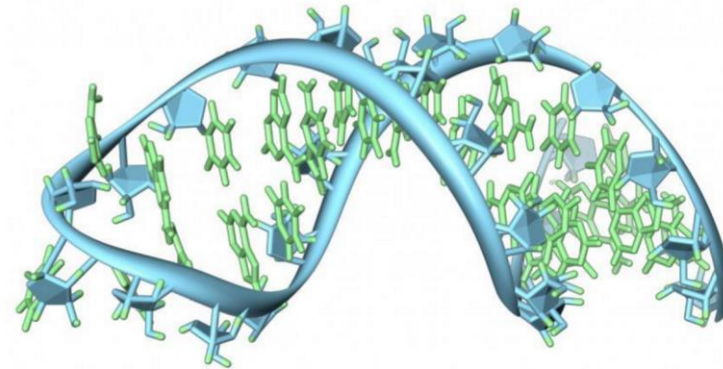 - Beat supercomputers on RNA design challenge

GWAP: EteRNA

Artist's diagram of an RNA molecule.   Vossman/Wikimedia Commons

For RNA paper based on a computer game, authorship creates an identity crisis

By John Bohannon | Feb. 17, 2016 , 5:30 PM

A journal published a paper today that reveals a set of folding constraints in the design of RNA

GWAPs: Principles

- ESPGame:
  - Independence: Two unrelated mutually unknown players
  - Agreement: Same tag
  - Shared information: All they get is shared image

  - Computes labels on images

  - "Output agreement" design pattern

GWAPs: Principles

- TagATune
  - Two players are each given a song, maybe same, maybe not
  - Exchange tags for their songs
  - Decide if it's the same song

  - Computes tags for songs

  - "Input agreement" design pattern

GWAPs: Principles

- Peekabom:
  - One player given image and a secret word, must slowly unveil pieces of the image as clues for the word
  - Second player gets image unveiled bit by bit, must guess word as quickly as possible

  - Computes salient aspects of image for a word

  - "Inversion problem" design pattern