CIENCIAS DE LA COMPUTACIÓN
UNIVERSIDAD DE CHILE

Privacidad en DM a.k.a. "Privacy preserving DM"

Bárbara Poblete

- De manera voluntaria y entusiasta nos hemos vuelto monitoreables
- Datos recolectados por medio de dispositivos que nos trackean
- Los entregamos muchas veces para poder conectar con otros

- Caso Target

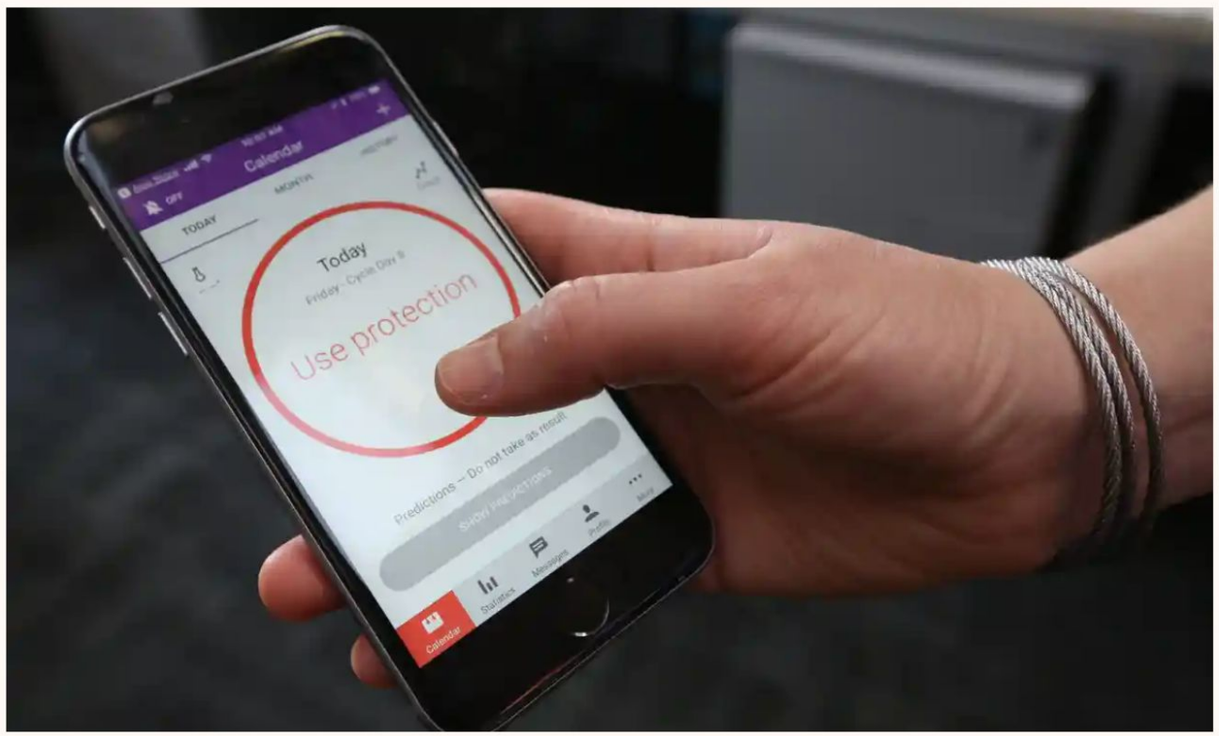
How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Kashmir Hill, FORBES STAFF

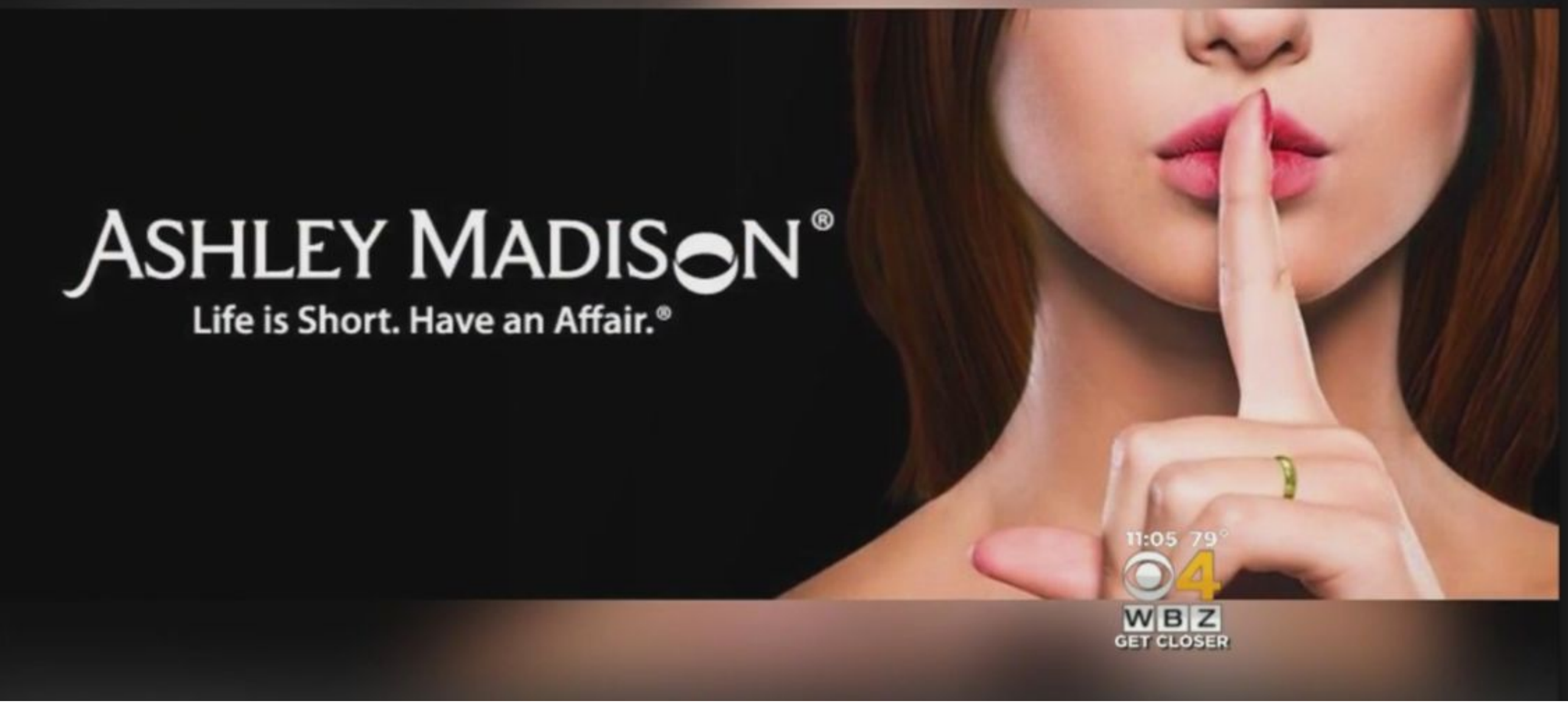
Welcome to The Not-So Private Parts where technology & privacy collide **FULL BIO**

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. Target TGT +0.13%, for example, has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.

Women's health apps are again raising concerns of privacy as a new study finds some are sharing information without consent



▲ Privacy International has released its findings that some menstruation apps have been sharing users' information with Facebook. Photograph: Nishat Ahmed/Associated Press



- Caso Ashley Madison

Ashley Madison data breach

From Wikipedia, the free encyclopedia

In July 2015, a group calling itself "The Impact Team" stole the user data of Ashley Madison, a commercial website billed as enabling extramarital affairs. The group copied personal information about the site's user base and threatened to release users' names and personally identifying information if Ashley Madison was not immediately shut down. On 18 and 20 August, the group leaked more than 25 gigabytes of company data, including user details.

Because of the site's policy of not deleting users' personal information – including real names, home addresses, search history and credit card transaction records – many users feared being publicly shamed.[1]

This article is part of a series on
Computer hacking

History
Phreaking · Cryptovirology

Hacker ethic
Hacker Manifesto · Black hat · Grey hat · White hat

Conferences
Black Hat Briefings · DEF CON · Chaos Communication Congress

Computer crime
Crimeware · List of computer criminals · Script kiddie

Hacking tools
Vulnerability · Exploit · Payload

Malware
Rootkit · Backdoor · Trojan horse · Virus · Worm · Spyware · Botnet · Keystroke logging · Antivirus software · Firewall · HIDS

Computer security
Application security · Network security

Groups
Hacker group

V · T · E

Timeline [edit]

The Impact Team announced the attack on 15 July 2015 and threatened to expose the identities of Ashley Madison's users if its parent company, Avid Life Media, did not shut down Ashley Madison and its sister site, "Established Men".[2]

On 20 July 2015, the website put up three statements under its "Media" section addressing the breach. The website's normally busy Twitter account fell silent apart from posting the press statements.[3] One statement read:

> "At this time, we have been able to secure our sites, and close the unauthorized access points. We are working with law enforcement agencies, which are investigating this criminal act. Any and all parties responsible for this act of cyber-terrorism will be held responsible. Using the Digital Millennium Copyright Act (DMCA), our team has now successfully removed the posts related to this incident as well as all Personally Identifiable Information (PII) about our users published online."[4] The site also offered to waive the account deletion charge.

Although Ashley Madison denied reports that a mass release of customer records occurred on 21 July,[5] over 60 gigabytes worth of data was confirmed to be valid on 18 August.[6] The information was released on BitTorrent in the form of a 10 gigabyte compressed archive and the link to it was posted on a dark web site only accessible via the anonymity network Tor.[7] The data was cryptographically signed[8] with a PGP key. In its message, the group blamed Avid Life Media, accusing the company of deceptive practices: "We have explained the fraud, deceit, and stupidity of ALM and their members. Now everyone gets to see their data ... Too bad for ALM, you promised secrecy but didn't deliver."[9]

In response, Avid Life Media released a statement that the company was working with authorities to investigate, and said the hackers were not "hacktivists" but criminals.[10] A second, larger, data dump occurred on 20 August 2015, the largest file of which comprised 12.7 gigabytes of corporate emails, including those of Noel Biderman, the CEO of Avid Life Media.[11]

- Emails, sms, social media, historial de búsqueda web, uso de sitios web y cruce de datos entre sitios

- Ubicación de celulares, historial de compras de tarjetas de crédito, wish lists, productos vistos/comentados, tarjetas de fidelización,

- Camaras (tuya, de otros, en la calle, reconocimiento facial) , micrófonos, Google glass, Fitbits, etiquetas GPS en fotos

- E-readers, uso de streaming de video, MOOCs,

- Lectores de patentes, uso de pasaporte, radio-frequency identification (RFID) readers, imágenes satelitales

Tormenta perfecta

- Espacio en disco a bajo costo
- Acceso a la red rápido y fácil
- Nuevos tipos de dispositivos de tracking
- Deseo y voluntad de las personas a llevarlos
- Necesidad humana de conexión con otros
- Desinterés en la privacidad

Información Personal

- Es toda información asociada a un individuo que incluye información que lo identifica (**identifying**) y que no (**non-identifying**)

- La información es dinero

Información Personal

- identifying: me permiten identificar la persona (rut, tarjeta de crédito, nombre, email...)
- non-identifying: edad, nivel educacional, sexo, historial criminal, etc...
- ¿Existe otra información privada?

¿Por qué es tema la privacidad en DM?

When do Data Mining Results Violate Privacy?*

Murat Kantarcıoğlu
Purdue University
Computer Sciences
250 N University St
West Lafayette, IN
47907-2066
kanmurat@cs.purdue.edu

Jiashun Jin
Purdue University
Statistics
150 N University St
West Lafayette, IN
47907-2067
jinj@stat.purdue.edu

Chris Clifton
Purdue University
Computer Sciences
250 N University St
West Lafayette, IN
47907-2066
clifton@cs.purdue.edu

ABSTRACT

Privacy-preserving data mining has concentrated on obtaining valid results when the input data is private. An extreme example is Secure Multiparty Computation-based methods, where only the results are revealed. However, this still leaves a potential privacy breach: Do the results themselves violate privacy? This paper explores this issue, developing a framework under which this question can be addressed. Metrics are proposed, along with analysis that those metrics are consistent in the face of apparent problems.

Categories and Subject Descriptors

H.2.8 [**Database Management**]: Database Applications—*Data mining*; H.2.7 [**Database Management**]: Database Administration—*Security, integrity, and protection*

General Terms

Security

Keywords

Privacy, Inference

for two parties, it has been shown in [10] that this is also feasible for many parties (e.g., rather than providing "noisy" survey results as in [3], individuals provide encrypted survey results that can be used to generate the resulting data mining model.) This is discussed further in Section 4.

However, though these provably secure approaches reveal nothing but the resulting data mining model, they still leave a privacy question open: Do the resulting data mining models inherently violate privacy?

This paper presents a start on methods and metrics for evaluating the privacy impact of data mining models. While the methods are preliminary, they provide a cross-section of what needs to be done, and a demonstration of techniques to analyze privacy impact. Work in privacy-preserving data mining has shown how to build models when the training data is kept from view; the full impact of privacy-preserving data mining will only be realized when we can guarantee that the resulting models do not violate privacy.

To make this clear, we present a "medical diagnosis" scenario. Suppose we want to create a "medical diagnosis" model for public use: a classifier that predicts the likelihood of an individual getting a terminal illness. Most individuals would consider the classifier output to be sensitive – for ex-

¿Qué podemos hacer al respecto?

Anonimización en Big Data

- Datos recolectados, utilizados por el recolector (uso directo? anonimizado? se venden?)

- ¿Si se anonimizan para compartir datos para fines científicos?

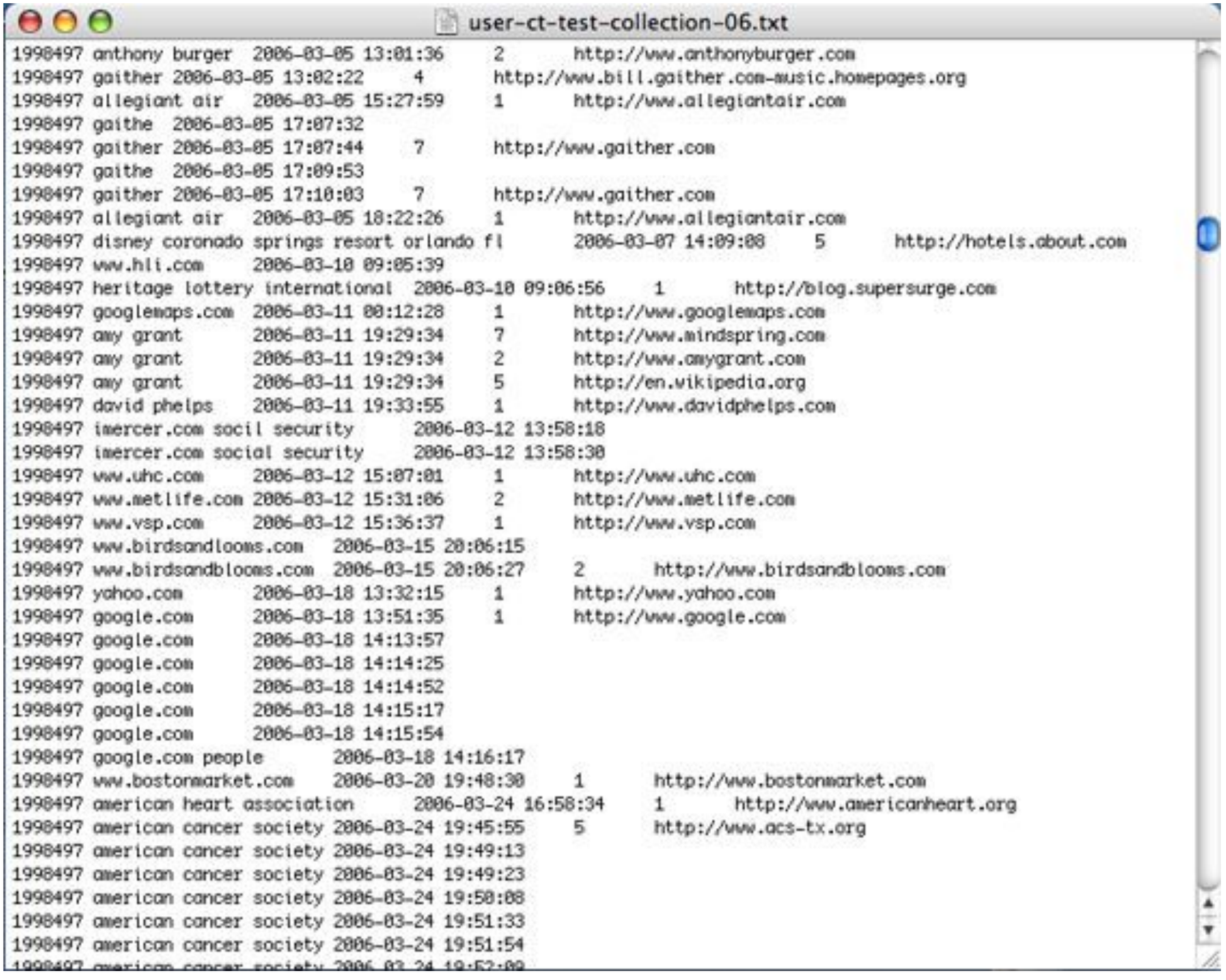
- "Right to be Forgotten" en UE

- Caso Gov. William Weld in Massachusetts

> At the time GIC released the data, William Weld, then Governor of Massachusetts, assured the public that GIC had protected patient privacy by deleting identifiers. In response, then-graduate student Sweeney started hunting for the Governor's hospital records in the GIC data. She knew that Governor Weld resided in Cambridge, Massachusetts, a city of 54,000 residents and seven ZIP codes. For twenty dollars, she purchased the complete voter rolls from the city of Cambridge, a database containing, among other things, the name, address, ZIP code, birth date, and sex of every voter. By combining this data with the GIC records, Sweeney found Governor Weld with ease. Only six people in Cambridge shared his birth date, only three of them men, and of them, only he lived in his ZIP code. In a theatrical flourish, Dr. Sweeney sent the Governor's health records (which included diagnoses and prescriptions) to his office.

Boom! But it was only an early mile marker in Sweeney's career; in 2000, she showed that 87 percent of all Americans could be uniquely identified using only three bits of information: ZIP code, birthdate, and sex.

Caso AOL


user-ct-test-collection-06.txt
1998497 anthony burger 2006-03-05 13:01:36 2 http://www.anthonyburger.com
1998497 gaither 2006-03-05 13:02:22 4 http://www.bill.gaither.com-music.homepages.org
1998497 allegiant air 2006-03-05 15:27:59 1 http://www.allegiantair.com
1998497 gaithe 2006-03-05 17:07:32
1998497 gaither 2006-03-05 17:07:44 7 http://www.gaither.com
1998497 gaithe 2006-03-05 17:09:53
1998497 gaither 2006-03-05 17:10:03 7 http://www.gaither.com
1998497 allegiant air 2006-03-05 18:22:26 1 http://www.allegiantair.com
1998497 disney coronado springs resort orlando fl 2006-03-07 14:09:08 5 http://hotels.about.com
1998497 ww.hli.com 2006-03-10 09:05:39
1998497 heritage lottery international 2006-03-10 09:06:56 1 http://blog.supersurge.com
1998497 googlemaps.com 2006-03-11 00:12:28 1 http://www.googlemaps.com
1998497 amy grant 2006-03-11 19:29:34 7 http://www.mindspring.com
1998497 amy grant 2006-03-11 19:29:34 2 http://www.amygrant.com
1998497 amy grant 2006-03-11 19:29:34 5 http://en.wikipedia.org
1998497 david phelps 2006-03-11 19:33:55 1 http://www.davidphelps.com
1998497 imercer.com socil security 2006-03-12 13:58:18
1998497 imercer.com social security 2006-03-12 13:58:30
1998497 ww.uhc.com 2006-03-12 15:07:01 1 http://www.uhc.com
1998497 ww.metlife.com 2006-03-12 15:31:06 2 http://www.metlife.com
1998497 ww.vsp.com 2006-03-12 15:36:37 1 http://www.vsp.com
1998497 ww.birdsandlooms.com 2006-03-15 20:06:15
1998497 ww.birdsandblooms.com 2006-03-15 20:06:27 2 http://www.birdsandblooms.com
1998497 yahoo.com 2006-03-18 13:32:15 1 http://www.yahoo.com
1998497 google.com 2006-03-18 13:51:35 1 http://www.google.com
1998497 google.com 2006-03-18 14:13:57
1998497 google.com 2006-03-18 14:14:25
1998497 google.com 2006-03-18 14:14:52
1998497 google.com 2006-03-18 14:15:17
1998497 google.com 2006-03-18 14:15:54
1998497 google.com people 2006-03-18 14:16:17
1998497 ww.bostonmarket.com 2006-03-20 19:48:30 1 http://www.bostonmarket.com
1998497 american heart association 2006-03-24 16:58:34 1 http://www.americanheart.org
1998497 american cancer society 2006-03-24 19:45:55 5 http://www.acs-tx.org
1998497 american cancer society 2006-03-24 19:49:13
1998497 american cancer society 2006-03-24 19:49:23
1998497 american cancer society 2006-03-24 19:50:08
1998497 american cancer society 2006-03-24 19:51:33
1998497 american cancer society 2006-03-24 19:51:54

AOL

AO Da

Poste

0 SHARES

The utter stupidity of this is staggering. AOL has released very private data about its users without their permission. While the AOL username has been changed to a random ID number, the abilitiy to analyze all searches by a single user will often lead people to easily determine who the user is, and what they are up to. The data includes personal names, addresses, social security numbers and everything else someone might type into a search box.

The most serious problem is the fact that many people often search on their own name, or those of their friends and family, to see what information is available about them on the net. Combine these ego searches with porn queries and you have a serious embarrassment. Combine them with "buy ecstasy" and you have evidence of a crime. Combine it with an address, social security number, etc., and you have an identity theft waiting to happen. The possibilities are endless.

Yet A

Furth
is at le
data c
but it
for anyone that wants to use (or abuse) it.

Update: Sometime around 7 pm PST on Sunday, the AOL site referred to below was taken down. The direct link to the data is still live. A cached copy of the page is here.

AOL must have missed the uproar over the DOJ's demand for "anonymized" search data last year that caused all sorts of pain for Microsoft and Google. That's the only way to

t Story

1985

OVERVIEW
AOL is a global digital media and technology company that offers video, mobile, and advertising technology and platform solutions, all within an open ecosystem, to consumers, advertisers, publishers, and subscribers worldwide. The company's brand group offers original digital print and video content

Her searches are a catalog of intentions, curiosity, anxieties and quotidian questions. There was the day in May, for example, when she typed in “termites,” then “tea for good health” then “mature living,” all within a few hours.

Her queries mirror millions of those captured in AOL’s database, which reveal the concerns of expectant mothers, cancer patients, college students and music lovers. User No. 2178 searches for “foods to avoid when breast feeding.” No. 3482401 seeks guidance on “calorie counting.” No. 3483689 searches for the songs “Time After Time” and “Wind Beneath My Wings.”

At times, the searches appear to betray intimate emotions and personal dilemmas. No. 3505202 asks about “depression and medical leave.” No. 7268042 types “fear that spouse contemplating cheating.”

Erik S. Lesser for The New York Times

Thelma Arnold's identity was betrayed by AOL records of her Web searches, like ones for her dog, Dudley, who clearly has a problem.

subdivision gwinnett county georgia.”

It did not take much investigating to follow that data trail to Thelma Arnold, a 62-year-old widow who lives in Lilburn, Ga., frequently researches her friends’ medical ailments and loves her three dogs. “Those are my searches,” she said, after a reporter read part of the list to her.

La anonimización no ayuda mucho

- actualmente se pueden cruzar muchísimas bases de datos diferentes (desconocidas de antemano)
- este cruce puede resultar en identificar personas

Netflix Prize

- <user, movie, date of grade, grade>

RYAN SINGEL SECURITY 12.17.09 4:29 PM

NETFLIX SPILLED YOUR *BROKEBACK MOUNTAIN* SECRET, LAWSUIT CLAIMS

[M]ovie and rating data contains information of a more highly personal and sensitive nature. The member's movie data exposes a Netflix member's personal interest and/or struggles with various highly personal issues, including sexuality, mental illness, recovery from alcoholism, and victimization from incest, physical abuse, domestic violence, adultery, and rape.

The Plaintiffs' and class members' movie data and ratings, which were released without authorization or consent, have now become a permanent, public record on the Internet, free to be manipulated and exposed at the whim of those who have the Database.

That's why the lesbian mom joined the lawsuit as a Jane Doe, according to the complaint, since she believes that "were her sexual orientation public knowledge, it would negatively affect her ability to pursue her livelihood and support her family and would hinder her and her children's ability to live peaceful lives."

The contest ended this summer when two different teams passed the 10 percent improvement mark, with the prize money going to a team led by AT&T researchers.

¿Qué se puede publicar entonces?

- ¿Qué sets de datos?
- ¿Cómo deben ser tratados antes?
- ¿Hay consideraciones con los resultados de un estudio de DM?

k-anonymity

- Técnica para anonimizar datos (Sweeney 2002)
- "Given person-specific field-structured data, produce a release of the data with scientific guarantees that the individuals who are the subjects of the data cannot be re-identified while the data remain practically useful."

k-anonymity

Name	Age	Gender	State of domicile	Religion	Disease
Ramsha	29	Female	Tamil Nadu	Hindu	Cancer
Yadu	24	Female	Kerala	Hindu	Viral infection
Salima	28	Female	Tamil Nadu	Muslim	TB
sunny	27	Male	Karnataka	Parsi	No illness
Joan	24	Female	Kerala	Christian	Heart-related
Bahuksana	23	Male	Karnataka	Buddhist	TB
Rambha	19	Male	Kerala	Hindu	Cancer
Kishor	29	Male	Karnataka	Hindu	Heart-related
Johnson	17	Male	Kerala	Christian	Heart-related
John	19	Male	Kerala	Christian	Viral infection

Supresión y generalización

Name	Age	Gender	State of domicile	Religion	Disease
*	20 < Age ≤ 30	Female	Tamil Nadu	*	Cancer
*	20 < Age ≤ 30	Female	Kerala	*	Viral infection
*	20 < Age ≤ 30	Female	Tamil Nadu	*	TB
*	20 < Age ≤ 30	Male	Karnataka	*	No illness
*	20 < Age ≤ 30	Female	Kerala	*	Heart-related
*	20 < Age ≤ 30	Male	Karnataka	*	TB
*	Age ≤ 20	Male	Kerala	*	Cancer
*	20 < Age ≤ 30	Male	Karnataka	*	Heart-related
*	Age ≤ 20	Male	Kerala	*	Heart-related
*	Age ≤ 20	Male	Kerala	*	Viral infection

“I Know What You Did Last Summer” — Query Logs and User Privacy

Rosie Jones Ravi Kumar Bo Pang Andrew Tomkins
Yahoo! Research, 701 First Ave, Sunnyvale, CA 94089.
{jonesr,ravikumar,bopang,atomkins}@yahoo-inc.com

ABSTRACT

We investigate the subtle cues to user identity that may be exploited in attacks on the privacy of users in web search query logs. We study the application of simple classifiers to map a sequence of queries into the gender, age, and location of the user issuing the queries. We then show how these classifiers may be carefully combined at multiple granularities to map a sequence of queries into a set of candidate users that is 300-600 times smaller than random chance would allow. We show that this approach remains accurate even after removing personally identifiable information such as names/numbers or limiting the size of the query log.

We also present a new attack in which a real-world acquaintance of a user attempts to identify that user in a large query log, using personal information. We show that combinations of small pieces of information about terms a user would probably search for can be highly effective in identifying the sessions of that user.

We conclude that known schemes to release even heavily scrubbed query logs that contain session information have significant privacy risks.

Categories and Subject Descriptors: H.3.3 **[Information Storage and Retrieval]**: Information Search and Retrieval

bilities; this is the goal of this paper. We initiate the study of subtle cues to user identity that exist as vulnerabilities in web search query logs, which may be exploited in attacks on the privacy of users.

Privacy attack models. We begin with a characterization of two key forms of attack against which a query log privacy scheme must be resilient. The first is a *trace attack*, in which an attacker studies a privacy-enhanced version of a sequence of searches (*trace*) made by a particular user, and attempts to discover information about that user. In our study of this type of attack, we draw upon the framework of k-anonymity [13], and study the extent to which information about gender, age, and location may be guessed from queries in a web search search log, and how effectively uncertain information along these dimensions will allow us to identify a small number of users containing the true user who generated the trace.

The second is a *person attack*, in which an unscrupulous agent attempts to discover that traces in a search engine log correspond to a particular known user. This is possible if some personal/background information of the user is accessible by the agent. We show that the risks of this form of attack are significant: even if the logs have been scrubbed by removing information about names and places, a few pieces of independent information may quickly shatter the set

Intentémoslo nuevamente:

- l-Diversity
- t-Closeness

Intentémoslo nuevamente:

- Differential privacy

  (differentially private ML)

  http://www.kdnuggets.com/2015/01/differential-privacy-data-mining-compatible.html

Differential privacy

- Differential privacy ensures the privacy of people but not of populations.

- adding noise to any answer returned by the database

- This may seem counterintuitive. We typically mine data in the hope of seeing through noise. Why would we deliberately add it? The hope with differential privacy is that the amount of noise added should be large enough to conceal the effects of individuals, but small enough that it does not seriously impact the usefulness of the answer.

Aldeen *et al. SpringerPlus (2015) 4:694*
DOI 10.1186/s40064-015-1481-x

SpringerPlus
a SpringerOpen Journal

REVIEW Open Access

A comprehensive review on privacy preserving data mining

Yousra Abdul Alsahib S. Aldeen[1,2*], Mazleena Salleh[1] and Mohammad Abdur Razzaque[1]

*Correspondence:
yohrmz_8@yahoo.com
[1] Faculty of Computing, University Technology Malaysia, UTM, 81310 UTM Skudai, Johor, Malaysia
Full list of author information is available at the end of the article

Abstract

Preservation of privacy in data mining has emerged as an absolute prerequisite for exchanging confidential information in terms of data analysis, validation, and publishing. Ever-escalating internet phishing posed severe threat on widespread propagation of sensitive information over the web. Conversely, the dubious feelings and contentions mediated unwillingness of various information providers towards the reliability protection of data from disclosure often results utter rejection in data sharing or incorrect information sharing. This article provides a panoramic overview on new perspective and systematic interpretation of a list published literatures via their meticulous organization in subcategories. The fundamental notions of the existing privacy preserving data mining methods, their merits, and shortcomings are presented. The current privacy preserving data mining techniques are classified based on distortion, association rule, hide association rule, taxonomy, clustering, associative classification, outsourced data mining, distributed, and k-anonymity, where their notable advantages and disadvantages are emphasized. This careful scrutiny reveals the past development, present research challenges, future trends, the gaps and weaknesses. Further significant enhancements for more robust privacy protection and preservation are affirmed to be mandatory.

Keywords: Privacy preserving, Data mining, Distortion, Association, Classification, Clustering, Outsourcing, K-anonymity

¿Conclusiones?

dcc

CIENCIAS DE LA COMPUTACIÓN
UNIVERSIDAD DE CHILE

www.dcc.uchile.cl

/ DCCUCHILE