

Online Privacy
Ponnurangam Kumaraguru (“PK”)
Professor Giri
Department of Computer Science
Indian Institute of Information Technology, Hyderabad
Week 03
Social Media Privacy

(Refer Slide Time: 0:16)

http://mattckon.com/facebook-privacy/

27

INTERNATIONAL INSTITUTE OF
INFORMATION TECHNOLOGY
HYDERABAD

Information Revelation and Privacy in Online Social Networks
(The Facebook case)

The proceedings version. AICM Workshop on Privacy in the Electronic Society (PPES), 2005

Ralph Gross
Data Privacy Laboratory
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
rgross@cs.cmu.edu

Alessandro Acquisti
H. John Heinz
School of Public Policy and Management
Carnegie Mellon University
Pittsburgh, PA 15213
aquisti@andrew.cmu.edu

ABSTRACT

Participation in social networking sites has dramatically increased in recent years. Services such as Facebook, MySpace, or the Facebook allow millions of individuals to create online profiles and share personal information with vast networks of friends – and, often, unknown members of strangers. In this paper we study patterns of information revelation in online social networks and their privacy implications. We analyze the online behavior of more than 4,000 Carnegie Mellon University students who have joined a popular social networking site named to colleges. We evaluate the amount of information they disclose and study their usage of the site's privacy settings. We highlight potential attacks on various aspects of their privacy, and we show that only a minimal percentage of users change the highly personalized privacy preferences.

Categories and Subject Descriptors

K.4.1 [Computer and Society]: Public Policy Issues—Privacy

28

Social Media Privacy: Online Vs Offline

Now, let us look at Social Media Privacy, particularly looking at social privacy, offline and online, connecting offline and online privacy. This is one of very popular work done by Alessandro and Ralph, which is looking at information flow on social networks, and looking at the offline way that people behave also.

And please keep in mind, these are phenomenal work, like a decade or this I think, is 2004, 2005 paper. So, this is WPES, 2005. Keep that in mind, while we are thinking these kinds of work are influential. But these kinds of work, you should also keep the context in mind and when this should have been done and I am all of these studies can be repeated now I am pretty sure. If you think about it, these kinds of studies can be repeated now to study how privacy is what people think of privacy now.

(Refer Slide Time: 1:24)

The slide is titled "Background" and contains the following text:

- In 2000, 100 billion photos were shot worldwide
- In 2010, 2.5 billion photos per month were uploaded by Facebook users only
- In 2015, 1.8 billion photos uploaded everyday on Facebook, Instagram, Flickr, Snapchat, and WhatsApp
- Facebook, Microsoft, Google, Apple, have acquired / licensed products that do Face recognition

There is a red handwritten word "Anchor" written over the text. The slide also features the NPTEL logo in the top left, the International Institute of Information Technology Hyderabad logo in the top right, and a small video inset of a man in a yellow shirt in the bottom right corner.

I will go through this paper in detail a little bit to give you a sense of connecting this offline and online, what interesting questions can be asked? Again, I am assuming that you will also think of some ideas as I have said in week 1, it will be really, really nice if some of you actually take up some interesting projects, as part of this course, do some data collection, do some analysis, do some system building and see how it works.

It is not part of the evaluation. But I think it will be fun for you. In 2000, so, this is again, giving you some numbers, giving you some numbers about what the, how the status at that point was. And I added some 2015 numbers. I am sure you can add some numbers, even please, more recently. In 2000, 100 billion photographs, were shot worldwide. In 2010, 2.5 billion photos per month were uploaded on Facebook. In 2015, 1.8 billion photos were uploaded every day on Facebook, Insta, Flickr, Snapchat and WhatsApp.

Those are huge numbers, photos are getting uploaded everywhere. I am sure, from the week 1 until now, in the last 15 days, I am sure many of you would have uploaded many pictures on these platforms. I am sure you would see me only uploading pictures as part of this class itself.

Please, again if you think that there are interesting pictures that you can take, which is you listening to my lecture, share it with me, I am happy to post it on social media, again, it is part of the class. It is part of the class, studying how much information can be found and understanding this whole mechanism of sharing these pictures and looking at what is going on.

And for me, I think as if you have started following me by now, on any of these platforms, I actually upload a lot of pictures, I contribute to this 1.8 billion pictures that are uploaded on these platforms. And all companies like Facebook, Microsoft, Google, Apple, Amazon I think in you can add the list of Amazon also here.

All of them have actually acquired companies, which does face recognition because it is part of their solution. Meaning if you use products, I mean in phone unlocks now because it is actually looking at your phone, face through the camera, it is actually seeing that face recognition solution. So, that is a manufacturing of phone is actually using these products now.

(Refer Slide Time: 4:23)

NPTEL

29

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Many things are colluding

- Increasing public self-disclosures through online social networks
 - Photos
 - Location
 - Daily life details
- Improving accuracy in Face recognition
 - Thanks to Facebook / Google ...
- Cloud, ubiquitous computing
 - Number of Cloud providers
- Re-identification techniques are getting better

CS PR

30 / 52

30

And from the privacy standpoint, it is actually interesting that many things are coming together. One is uploading of pictures are increasing but it is also that public self-disclosure through online social app, we are uploading pictures but we are also giving away more information than just picture. So, you are saying where the picture was taken, who you are with, all that photos, location, daily life details, all of that is being uploaded. Like for example, I think I uploaded a picture.

When I started my first lecture recording, I am hoping that I will do the last lecture recording also by taking a picture. Improving accuracy in face recognition is increasing uploading of pictures, information with that picture is increasing, recognizing the objects in the images are also increasing.

Again thanks to lots of computer vision solutions which is helping these things happen better. The ubiquitous computing part which is only available devices everything is also increasing day by day. Re-identification techniques are also getting better, which is if you have the one that I said earlier if you have CS prof as an account on Reddit can you actually re-identify the CS prof the techniques are getting better and you have a picture in which PK is there can you identify picture PK in that picture when that picture does not, so to say, it explicitly state that PKs in that picture.

So, those are the things that you can actually get better now, in all the years the technologies have become better to our trust is. Information being uploaded, pictures being uploaded, information attached to it, the face recognition and technologies for that is improved basically cloud service providers and information that you can put it on cloud and keep it there for as many as long as you want has increased, Re-identification has also increased. So, keeping all this in mind is what they started asking a question.

(Refer Slide Time: 6:44)

The slide contains the following text:

Question they were interested in

- Can one combine publicly available online social network data with off-the-shelf face recognition technology for
- Individual re-identification
- Finding potentially, sensitive information

Handwritten red notes on the slide:

- Neutronicity
- Alcoholic
- Salary
- Location

A small video inset in the bottom right corner shows a man in a light-colored shirt speaking.

They start asking the question that can one combine publicly available information which is on social network data with offline with off the shelf face recognition technology for individual Re-identification which is you take a Facebook post which has PK in it, nobody said PK is not tagged in the picture, but you are uploading the picture and can you re-identify PK in that picture.

Individual Re-identification find potentially sensitive information also, can you identify that PK is actually in a location where he is not supposed to be or PK is with people who he not

supposed to be there or PK said that he is doing a but you have picture to show that he is doing b, all of that. All of these are very interesting questions that come up as these pictures can be used for analysis.

We have actually seen some interesting examples of these also, for example, matrimonial websites. Matrimonial websites has these interesting feature because there people say what they want people to know about them. I think you can argue for the face, social media itself like that. In matrimonial it is a little serious about getting relationships built there, so, there is a feature alcoholic.

Meaning there are features like salary, all of that gender. All this is there. But one of the features that we found interesting at some point in time was alcoholic. Lots of people who would say that I am non-alcoholic, they would put non-alcoholic for probably a better impression I guess. It is all about impression management.

When we analyse some of these matrimonial websites, we found that again, all of this is publicly available, we found that people will say non-alcoholic on a matrimonial websites, you look at their Facebook or a Twitter post, they would actually be in a pub, they would have uploaded a picture where it shows that they are actually consuming alcohol.

Interesting, which is what am I trying to do is connecting to this question, which is take up social media information, which is somebody posted a picture on Twitter in a pub or a bar, can you take that picture and use this other information in this case matrimonial website to re-identify this person saying that it is actually PK, he or she is from this location. Interesting question. Interesting connection you can make. And we found some interesting patterns after this also.

(Refer Slide Time: 9:56)

The slide features the NPTEL logo on the top left and the International Institute of Information Technology Hyderabad logo on the top right. Handwritten red notes at the top center include 'Matrimonial' with an arrow pointing to 'Alcoholic', 'Salary', and 'Location'. The main text on the slide reads: 'Goal is to Use un-identified source {Match.com, photos from Flickr, CCTVs, etc.} + identified sources {Facebook, LinkedIn, Govt. websites, etc.} Get some sensitive information of the individual {gender orientation, SSN, Aadhaar #, etc.}'. A small inset image of a man in a yellow shirt is visible in the bottom right corner of the slide.

Good, use unidentified source again the goal was online-offline. So, use unidentified source which is Matchme.com, photos from Flickr, CCTVs, etcetera. And plus identified sources like Facebook, LinkedIn, Government websites, etc.

Matchme.com is matrimonial website, any matrimonial website you can take, Flickr pictures uploaded, we are not identified there. Whereas in Facebook, you can upload. If I upload a picture, there is a high chance that I am going to be on the picture and if I tag you there is a high chance that you are also on that picture. That is the identified source on the non-identify, unidentified source.

That is the difference of these two datasets. You get some sensitive information of individual using this information itself, gender orientations, social security number, adhaar number, can you actually derive all this information from this online and offline source? So, those were the goals for the study that they had in mind. One is to put the unidentified source with the identified source. With that, can they actually find out some sensitive information of individuals, let us continue looking at the study.

(Refer Slide Time: 11:24)

The slide features a Venn diagram with two overlapping circles. The left circle is labeled 'Medical Data' and contains the following items: Ethnicity, Visit date, Diagnosis, Procedure, Medication, and Total charge. The right circle is labeled 'Voter List' and contains: Name, Address, Date registered, Party affiliation, and Date last voted. The intersection of the two circles, highlighted with a red oval, contains the items: ZIP, Birth date, and Sex. The text 'Re-identification' is positioned to the right of the diagram. In the bottom right corner, there is a small video inset showing a man in a white shirt speaking.

Figure 1 Linking to re-identify data

So, one of the concepts that we will also, as I said before, we will deal more in detail later in the semester, this whole idea for real identification, automated techniques, all that. So, this is one work that became very popular, because of which the idea of re-identification anonymity also became very, very important.

So, this was work done by LaTonya Sweeney at MIT. And what she did was she basically took medical records and voter list, put them together, and then found that just this ZIP, birth date and gender will actually be useful in re-identifying people in the US. So, this may look slightly trivial now. But this is this was the first time it was done.

So, it was actually a phenomenal piece of work in 90s, where, where she looked at putting these two data together, and then re-identifying people, and particularly she also re-identifying some politicians there with, given their finding their medical records from this medical data was actually super cool at that point in time.

(Refer Slide Time: 12:56)

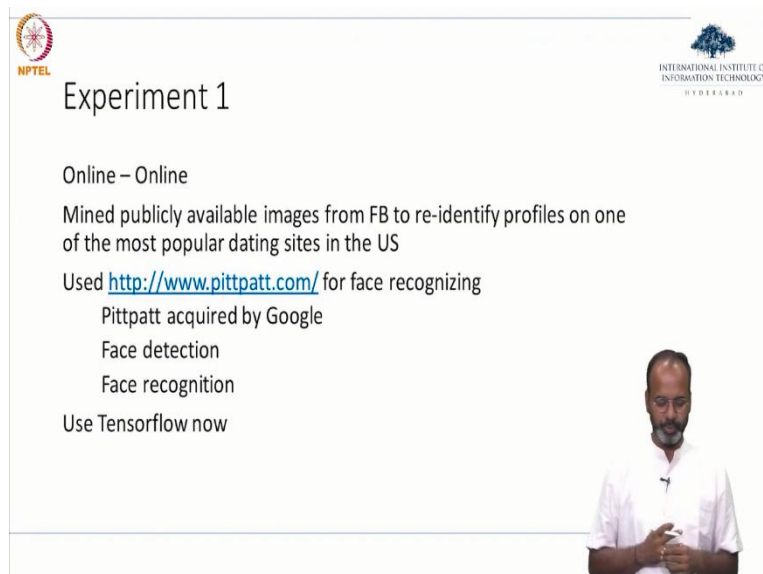


The slide features the NPTEL logo in the top left and the International Institute of Information Technology Hyderabad logo in the top right. The main title is "Techniques to Anonymize". Below the title is a list of techniques: K-anonymity, L-diversity, T-closeness, Differential Privacy, and A presenter is visible in the bottom right corner of the slide.

There are many techniques that has been built on this idea called anonymity, which is to protect the data, when it is shared, and people cannot, nobody can actually re-identify people in the data set. So, the idea called K-anonymity, which is done by LaTonya, then there was idea L diversity, T closeness.

Now, differential privacy is one of the very popular techniques by which anonymity is provided. I am sure if you are using the Apple phones, if you are using tools, which as differential privacy, and these days, many tools are actually applying differential privacy in their features. We will look at this more detail later in the semester, just focusing only on what they are, how to technically achieve them.

(Refer Slide Time: 13:55)



The slide features the NPTEL logo in the top left and the International Institute of Information Technology Hyderabad logo in the top right. The main title is "Experiment 1". Below the title is the text: "Online – Online", "Mined publicly available images from FB to re-identify profiles on one of the most popular dating sites in the US", "Used <http://www.pittpatt.com/> for face recognizing", "Pittpatt acquired by Google", "Face detection", "Face recognition", and "Use Tensorflow now". A presenter is visible in the bottom right corner of the slide.

So, what was Alessandra and Ralph trying to do? Their first experiment was to trying to connect online and online. What is the goal? Mined publicly available images from Facebook to re-identify profiles on one of the most popular dating websites. If you remember, I said dating websites, it is match.com.

You can think of it as jeevansathi.com any of those kind of websites, you could actually think as one of the source. What they did? They found these images, so, they wanted to figure out image who is that in the image. So, they use this tool called pittpatt.com.

Now it is acquired by Google, but the idea is to face detection and when there is an image, if there is a face, the method would actually find out where the face is, and if there is a face, can it actually find in that who was it in the picture also, meaning now I am sure you are more skilful in terms of identifying these kinds of things. Many of them have become very popular tools, which you can just give as input as images and output as faces in the pictures.

(Refer Slide Time: 15:18)

Data

Identified

Downloaded FB profiles from one city in USA

Profiles: 277,978

Images: 274,540

Faces detected: 110,984

The data that they had was so again, remember identified and unidentified data, identified sources. The data that they had is downloaded Facebook profiles from one city in the US profiles.

They have profiles 270,000, and images 274,000, faces detected is about 110,000. So, this is the identified source because it is Facebook, because Ponnurangam Kumaraguru is uploading the pictures and it is taken from my profile.

So, you can actually identify that PK is in this picture are not. Today, this may be a little hard to do, because I think collecting the data from Facebook may be a little harder you have to do much more than what they did, during 2004 2005, when they could actually easily collect, give input as a city and get the output as a profile from that city. But as today you have to build an app get a lot more approval from the users who are using these solutions for the data to be collected. That is identified.

(Refer Slide Time: 16:27)

The image displays three sequential slides from a presentation. Each slide features the NPTEL logo in the top left and the International Institute of Information Technology Hyderabad logo in the top right. A small number '35' is visible in the top right of the first slide, and '36' is visible in the top right of the second slide. A presenter is visible in the bottom right corner of each slide.

Slide 1:

- Data**
- Un-Identified
- Downloaded profiles of one of the popular dating websites
- Pseudonyms to protect their identities
- Photos can be used to identify
- Same city was used to search
- Profiles: 5,818
- Faces detected: 4,959

Slide 2:

- Data**
- Identified
- Downloaded **FB** profiles from **one city in USA**
- Profiles: 277,978
- Images: 274,540
- Faces detected: 110,984

Slide 3:

- Data**

Downloaded FB profiles from one city in USA
Profiles: 277,978
Images: 274,540
Faces detected: 110,984

Data
Un-Identified
Downloaded profiles of one of the popular dating websites

Now looking at the unidentified data, unidentified data is its matrimonial website jeevansathi.com or matrimonial.com, I could actually create a profile in whatever name I want and now that it is unidentified, that is the data that they collected. Downloaded profiles on one of the popular dating websites pseudonymous to protect their identities, which people use I think earlier, one of the classes I said CS Prof as a pseudonymous identifier for me, which could be possible and photo can be used to identify photo from the Facebook profile, same city was used to search given that they did a search for one city in the US, they use the same city for data collection here.

Why? They wanted to make the comparison, they wanted to merge the data, so they pick the data from Hyderabad and if they pick the data from Chennai, today meaning there is high chance that you are going to find me in Chennai's data whereas if you find the Facebook profiles from Chennai, and the matrimonial website from Chennai, there are high chances that you will find me in both the data.

And the profiles that they collected were about 5818 and faces detected are about close to 5000 photos. You can clearly see that is the data that they are going to play around with photos here. Faces here.

(Refer Slide Time: 18:11)

Experiment 1: Approach

Unidentified {Dating site photos} + Identified {FB photos} --> Re-identified individual

More than 500 million pairs compared

Used only the best matching pair for each dating site picture

PittPatt produces score of -1.5 to 20

Crowd sourced to Mturkers for validating PittPatt

Likert scale, 1 - 5

At least 5 turkers for each pair

So, what was the experiment 1? Experiment 1 was connecting this unidentified source, which is the dating web sites to be identified source which is its Facebook profiles to re-identify people, I am sure you are already guessed it. Putting these two together, we should be able to find out who the CS prof in the matrimonial website is, from the Facebook profile that you have actually got mine more than 5 million compared.

So, you have to compare this 100,000 pictures with the 5000 pictures from these both sources used only the best match matching pair for each dating site picture. So, the idea is that they bothered to get the sort to say, the match should have been the maximum, that is when you can actually argue that look, these two are the same people.

And I am sure you can do many different algorithms to figure out whether this person is same or not, but they ended up making this choice as best matching pair pittpatt produces a course of score of minus 1.5 to 20. And then they also had the MTurkers. Which is Mechanical Turk platform where you actually show these images that they have found out that this is same PK from match. matrimonial.com and PK from Facebook.

They showed these two pictures let us take I do an exercise with all the class students in the class. I get you to do, actually say click on it saying okay, is this the same person or no? This is the same person or no? I will show you 50 pictures, and I get to tagged.

And similarly, I will show the same 50 pictures to 55 other people, all of them will actually tagged and accordingly, I can actually make a choice that whether it is same person or not,

this method is actually very common annotation, and getting real people to actually do this task.

Which will help the results that you have more confidence in it, now that people are actually doing and it is not just the algorithm that is doing it, people have confirmed it, what the algorithm form and they did a Likert scale of 1 to 5, which is matching, whether it is matching fully, whether it is matching partially or whether it is not matching, that is the results that they got.

So, why is this better, because, see otherwise, if you were to do go and do all the 500 million by the M Turkers, it is going to, it is going to take a lot of time, it is going to take a lot of money, instead, run it by an algorithm, reduce that number to a smaller size, and then actually show it to users. By that way, you are actually reducing the time, reducing the cost and probably making the results better also, at least 5 turkers for each pair.

So, this is what I said, I would take 50 images and get 5 of you to actually annotate the same picture, therefore I have more confidence again, if 5 people agree that it is the same person, or it is not the same person that is actually chances that it is not the same person. It is not all 5 people agree, let us take if 3 people agree that it is the same person, then just making my confidence level, higher is what all of this is about getting more people to annotate, getting actually people to annotate also, in addition to the algorithms that are found.

(Refer Slide Time: 22:06)

The slide displays the following text: "Experiment 1: Results", "Highly likely matches: 6.3%", "Highly likely + Likely matches: 10.5%", and "1 on 10 from the dating site can be identified". There are red handwritten annotations: a bracket above the percentages, a circled "1-5" next to them, and a circled "10.5%". The slide also features the NPTEL logo, the International Institute of Information Technology Hyderabad logo, and a small video inset of a man in a white shirt.

Experiment 1 results, what did they find? They found that highly likely matches is about 6.3 percent. When users enter it was actually matched, it was about 6.3 percent, highly likely plus

likely this is big, this is happening because of the 5 point Likert scale 1 to 5, highly likely, likely matches neutral, highly unlikely and then highly unlikely matches.

So, this is 10.5 percent, which means 1 in 10, from the dating website was actually Re-identified. Just imagine, just from one city, data was collected, and they could actually identify 1 on 10 percent of the people re-identification. So, therefore, it is actually many as you think about it, think about ways by which you can actually do this in India, also try some small experiments that do it.

(Refer Slide Time: 23:18)

The slide is titled "Reactions?" and contains the question "What can you do better if you were the attacker?". It features handwritten red text and a diagram. The text includes "Penguin CS pivot" and a diagram showing a horizontal line with a vertical line extending upwards from the center and another vertical line extending downwards from the right end. Below the left vertical line is "100%", and below the right vertical line is "0%", which is circled. A small inset video shows a man in a white shirt.

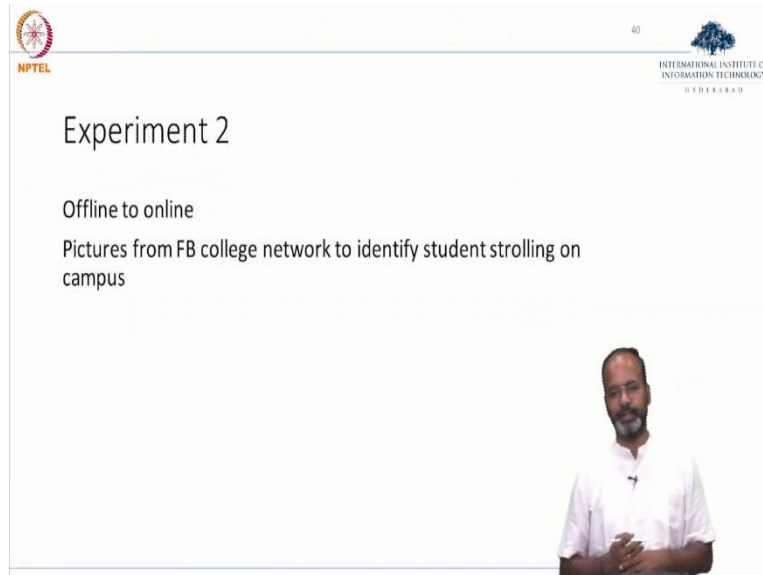
So, if you think about the methods that we just saw, which is just taking images, putting them together, figuring out whether they are the same person through an algorithm and then using an Mturker to confirm it, mean if you are the attacker, attacker in terms of actually finding either in one side, can you actually do things that it will not be re-identified. So, that I as a user, I speak here, I want to put the pictures on matrimonial website I am not on Facebook, but I do not want people to actually connect them together.

Versus on the other side, I want to put pictures together and I want people to re-identify that it is the same person. So, that is the spectrum that generally comes into discussion bar, where a user completely wants anonymity which is, I do not want to be re identified.

This let us take as 100 percent anonymity and this is 0 percent anonymity, just like, just connect me I am okay, even though if I have a handle says Punnurangam, I am CS Prof. I would like you to I would like an algorithm to actually figure out that it is actually the same

PK that is here. All decisions, all technologies, all platforms that we build, keeping this anonymity in mind will fall into this spectrum.

(Refer Slide Time: 25:00)



The slide features the NPTEL logo in the top left corner and the International Institute of Information Technology Hyderabad logo in the top right corner. The main content is centered and includes the title 'Experiment 2', the subtitle 'Offline to online', and the description 'Pictures from FB college network to identify student strolling on campus'. A small video inset in the bottom right corner shows a man in a white shirt speaking.

They did experiment 2, what they try, they tried offline to online, which is collect data somewhere in the offline mode and then link it to the data that is provided that is available in the online mode, pictures from Facebook college network to identify students trawling on the campus, so they took one campus, they collected data from that particular campus Facebook profiles.

Again, this may not be possible now, because of all the restrictions on Facebook. But, but during those days it was possible, which is find a search particular campus. And then you can get all the profiles from that particular campus.

(Refer Slide Time: 25:48)

NPTEL

41

INTERNATIONAL INSTITUTE OF
INFORMATION TECHNOLOGY
HYDRABAD

Experiment 2: Data

Webcam to take 3 pics per participant
Collected over 2 days
Facebook data for the university
Profiles: 25,051
Images: 26,262
Faces detected: 114,745

So, they had, what for the offline thing, what they did was, they just set up a small webcam on the campus and then for 2 days, and they just took pictures of people walking around on campus, they took 3 pictures of a particular participants, of course, it was all done, IRB approved everything.

So, the use of consent and into doing the study, and 3 pictures were taken about. So, let us assume that I am a student of this campus, I walk around, and then the study administrator stopped me and say would you be going to participate in the study?

If so, they would take 3 pictures of mine, and then use it for the data, use it for the analysis. Facebook data for the university that they collected profiles is about 25,000, images is about 26,000. And about 11,000 faces detected. It looks like they had a lot of faces and images, 26,000 images with 114,000 faces. That is the data.

(Refer Slide Time: 26:59)



The slide features the NPTEL logo on the top left and the International Institute of Information Technology Hyderabad logo on the top right. The slide number '42' is centered below the logos. The main content is a list of five steps describing the experimental process. A small video inset in the bottom right corner shows a man in a white shirt speaking.

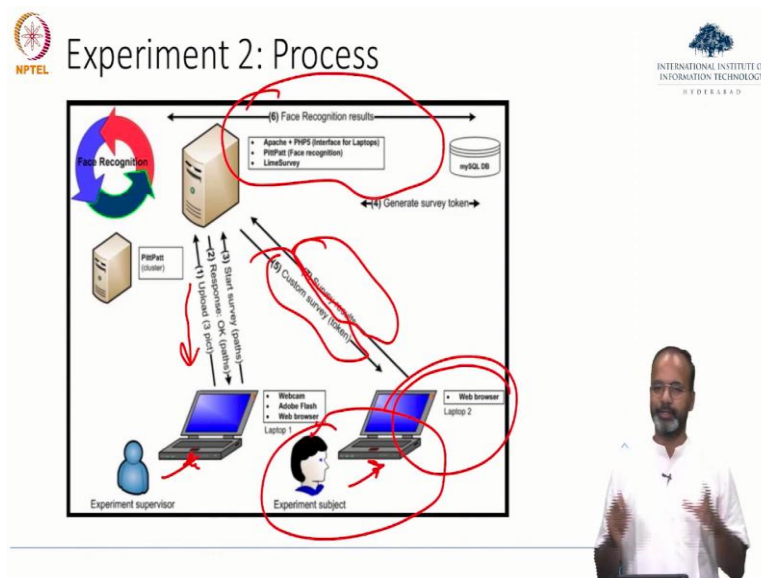
Experiment 2: Process

- Pictures taken of individuals walking in campus
- Asked to fill online survey
- Pictures matched from cloud while they are filling survey
- Last page of the survey with options of their pictures
- Asked to select the pics which matched closely, produced by the recognizer

Experiment 2 pictures taken of individuals walking in the campus asked to fill in online survey. So, while pictures of the 3 people were taken, and then the participants were asked to fill a survey, pictures match from the cloud, which is when they are filling the survey, there was matching down between the pictures that were taken to the pictures that they are already connected from Facebook.

I will show an image where everything is in probably like an architecture diagram, which will walk you through the process, pictures match from the cloud while they were filling the survey and last page of the survey, were the options of the pictures, which is by the time they finished the survey, they could actually do the comparison and bring it back I will show it in the last page saying, look, these are the pictures that we got from Facebook. Comparing the pictures that you just showed, that we just took keep yours, asked to select the pictures which match closely produced by the recognizer.

(Refer Slide Time: 28:07)



So, that is the flow of the experiment, which is starting here, upload 3 pictures, upload 3 pictures of the user and compare it with the server.

And then when compared with the server, bring it back, generate the survey token and survey meaning the user is filling the survey by that time they would compare and then by that comparison.

So, that is the architecture diagram that I was referring to, what they do is they get the user to take a picture of the user, they upload it from the users, meaning they take the picture of the user upload it to the server, do the comparison, face recognition results are showing up here.

And then they are filling the survey customer custom survey that they are filling, the results for the survey is given back. So, this is the experiment subject, this is the supervisor so essentially the supervisor is the one who is taking the picture uploading the picture getting the comparison and if you see this, this is me doing the study.

So, essentially, it is simple, taking my pictures, uploading it to a server, doing the comparison, asking me to fill the survey and then showing the pictures again to me.

(Refer Slide Time: 29:41)

NPTEL

INTERNATIONAL INSTITUTE OF
INFORMATION TECHNOLOGY
HYDRABAD

Experiment 2: Results

- 98 participants
- All students had FB accounts
- 38.18% of participants were matched with correct FB profile
- Including a participant who mentioned that he did not have a picture on FB (Shadow profiles!)
- Average computation less than 3 seconds

45

Experiment 2: Results

98 participants all students had Facebook account, 38 percent of the participants were matched with the correct Facebook profile, which is the correct message is coming from, I am being asked to mark saying, is this me from the picture.

That is the last part of the study, including a participant who mentioned that he did not have a picture on Facebook. So, that is the idea of shadow profiles, which I think in one of the videos that I asked us to see, you must have heard about a shadow profile, which is the social media platforms having information about users when they are not even on the platform.

That is the shadow profile, which is let us take some of you are not on Instagram and but, Insta actually has information about you, when you show up in their platform, they would actually use that for recommendations, use that for other purposes.

That is that idea is called Shadow profile. In this case, there was one participant in the study who said that he has never uploaded a picture on Facebook. But he was actually there was a picture which the researchers found that he was in the picture, average computation, this is just to show the computational side of it 3 seconds.

(Refer Slide Time: 31:46)

NPTEL Experiment 2: Results

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY HYDRABAD

Experiment 3

SOCIAL SECURITY

So, this is one mean a kind of anonymized picture. The idea is that they took this is the picture. This is the kind of picture that they would actually they were actually taking in the study, asking a participant, and they use this picture to find out this person was actually in a group picture from Facebook.

This is the study picture. While the study 3 pictures that they took, and it is this is the picture from Facebook that they found off the user. So, that is the kind of input and output that that was there in the study.

(Refer Slide Time: 31:52)

NPTEL

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY HYDRABAD

Experiment 3

Predicted SSN from public data

Faces / FB data + Public data → SSN

27% of subjects' first 5 SSN digits identified with four attempts – starting from their faces

Predicted sensitive information like SSN

SOCIAL SECURITY

YOUR NAME HERE

credit card issuers—including those specifically targeting individuals with poor credit (21); wireless carriers; or instant lending services (22). These services require information such as applicants' names, dates of birth, and SSNs to screen credit or service applications, thus offering an attacker a means to verify variations of predicted SSNs.

- sending mass spear phishing emails (23) based on social engineering (24). Such emails would include the target's first 5 or 6 SSN digits to elicit a revelation of the remaining digits;
- the SSA's own SSN Verification Service (www.ssa.gov/employer/ssn.htm) and the Department of Homeland Security's E-Verify system (www.uscis.gov/e-verify), 2 antifraud initiatives that allow employers to verify large numbers of employees' SSNs at a time. They could be abused if an attacker succeeded in impersonating companies' representatives or self-employed individuals.

Although defense mechanisms to detect repeated abuses are in place at those services (for instance, the SSNS tracks incorrect attempts at verifying SSNs, and financial institutions blacklist (for various days or months) IP addresses originating 3 or more failed logins or transactions (25)), "botnets" of compromised computers (26) allow attackers to test—cheaply and covertly—vast numbers of variations of targets' SSNs, strategically distributing simultaneous attempts across services, compromised machines, and target accounts. A rational attacker would focus on SSNs issued in states and years with higher prediction accuracies, taking advantage of the lack of a centralized, real-time system for the notification of hits and flags on credit account requests (27), as well as of the fact that, unlike traditional passwords, SSNs cannot be blacklisted after failed attempts, nor changed to avoid future fraud (28).

Consider, for instance, an attacker who rented a small hotnet (10,000 IP addresses) to apply for credit cards impersonating 18-year-old West Virginia-born U.S. residents (whose state and dates of birth he has obtained from commercial databases). Assuming that an IP address gets blacklisted by an online credit card issuer after 3 incorrect attempts, that the criminal distributes his or her attacks across 20 issuers and can find birth data for 50% of the potential targets, and that inquiries with the correct first 7 of 8 digits are sufficient for a CRA to contact with

necessary for the precautions is, itself, widely available: SSN predictions do not require knowledge of someone's birth zipcode but just his or her state and date of birth. Whereas SSNs are becoming harder to purchase in the open market (8) and less available in public documents (33), mass amounts of birth data for U.S. residents can be obtained or inferred—often for free or at negligible per unit prices—from multiple sources. They include data brokers (such as www.peoplefinders.com, which sells access to birth data and personal addresses for "almost every adult in the United States"); voter registration lists (for most states); online free people searches (such as www.zabasearch.com); as well as social networking sites: Our estimates indicate that at least 10 million U.S. residents make publicly available or inferable their birthdate information on their online profiles. An attacker may not even need birth data: The rise of synthetic identity theft (where fake names are combined with real SSNs and birthdates) suggests that a correspondence between birthdate and SSN can be sufficient to pass the screening of CRAs, even when names or addresses do not match those in the credit reports (21, 22). Our results show that such correspondence is inferable even without knowledge of the target's name.

These aspects are further discussed in ref. 34. There, we present an illustrative application of the prediction algorithm in which we infer alive individuals' SSNs based on public information we mined from a social networking site. To illustrate the actual threat of combining public records to infer sensitive information, we used DMF data as the analysis set to extract the most-frequent ANGNs and the SN regression coefficients for the range of states and birthdays corresponding to the alive individuals' birth data. We extracted the birth data from the public profiles of 621 students at a North American university. We then interpolated our sample's birth data with the patterns estimated from DMF records, and then predicted the former's SSNs. We verified the accuracy of our predictions against the subjects' actual SSN data (from the University Enrollment services), using a secure, IRB-approved protocol that disclosed to us only aggregate prediction accuracy statistics. We found that at parity of year and state of birth (and SN assignment), the test based on online social network data and the DMF test produced comparable results: we accurately predicted with a single attempt the first 5 digits for 6.3% of our sample.



Ralph and Alessandro did not stop here. They wanted to actually use this to do something more interesting, which is on the social security numbers. So, the TED talk that I asked you to watch last week, which is by Alessandro should have, should have given you an idea of what social security number study that Alessandro has been doing for quite some time in trying to re identify people their social security numbers.

What did they do, they predicted social security number from public data. This is another paper by itself, but it is an extension of the work that Ralph Alessandro did as in the other 3 part study. The faces are Facebook data, plus the public data, this is what they used, they had all this information, Faces, Facebook data that they had collected just now on the experiment 2.

And then the public data they could actually collect from anywhere on the internet, and then predict a social security number. So, this is for those of you who do not know what a social security number is, it is digits, numbers that are given to US citizens, like our adhaar number.

So, it is very, very meaning or I think our adhaar idea UID idea spun out of many of these national ID numbers across the world one after the social security number in the US. 27 percent of subjects first 5 SSN digits identified with 4 attempts starting from their face, which is from so the simple idea is that from the faces they were able to identify some digits of the social security number.

Why some digits interesting is because the way that the social security number is given the for example, if you see the 3 digit here, 2 digit here and then the 4 digit, this digit itself, the way it is placed itself was actually giving you some information, which is where I was born,

which city I was born, and during which time I would have I was born all that was embedded in this number, which is what was making it slightly easier and which is what, if you were able to find out, then you would actually re identify social security number of people.

Predicted sensitive information like social media number, they actually use this for predicting.

(Refer Slide Time: 34:45)

Predicting Social Security numbers from public data
Alessandro Acquisti¹ and Ralph Gross
Carnegie Mellon University, Pittsburgh, PA 15213
Communicated by Stephen E. Fienberg, Carnegie Mellon University, Pittsburgh, PA, May 5, 2009 (received for review January 18, 2009)

Information about an individual's place and date of birth can be exploited to predict his or her Social Security number (SSN). Using only publicly available information, we observed a correlation between individuals' SSNs and their birth data and found that for younger cohorts the correlation allows statistical inference of private SSNs. The inferences are made possible by the public availability of the Social Security Administration's Death Master File and the widespread accessibility of personal information from multiple sources, such as data brokers or profiles on social networking sites. Our results highlight the unexpected privacy consequences of the complex interactions among multiple data sources in modern information economies and quantify privacy risks associated with information revelation in public forums.

identity theft | online social networks | privacy | statistical reidentification

In modern information economies, sensitive personal data hide in plain sight amid transactions that rely on their privacy yet require their unimpeded circulation. Such is the case with Social Security numbers in the United States. Created as identifiers for accounts tracking individual earnings (1), they have turned into sensitive authentication devices (2), becoming one of the pieces of information most often sought by identity thieves. The Social Security Administration (SSA), which issues them, has urged individuals to keep SSNs confidential (3), coordinating with legislators to reduce their public exposure (4).⁵ After embarrassing breaches, private sector entities also have attempted to strengthen the protection of their customers' and employees' data (7).⁸ However, this has not been enough to prevent the SSA from releasing information about the process through which ANs, GNs, and SSNs are issued (1). ANs are currently assigned based on the zipcode of the mailing address provided in the SSN application form [RM00201.030] (1). Low-population states and certain U.S. possessions are allocated 1 AN each, whereas other states are allocated sets of ANs (for instance, an individual applying from a zipcode within New York state may be assigned any of 87 possible first 3 SSN digits). Within each SSA area, GNs are assigned in a precise but nonconsecutive order between 01 and 99 [RM00201.030] (1). Both the sets of ANs assigned to different states and the sequence of GNs are publicly available (see www.socialsecurity.gov/employers/stateweb.htm and www.ssa.gov/history/ssn/gscand.html). Finally, within each GN, SSNs are assigned "consecutively from 0001 through 9999" (13) (see also [RM00201.030], ref. 1).

The existence of such patterns is well known (14), and has been used to catch impostors posing with invalid or unlikely SSNs (15). However, outside the SSA, the understanding of these patterns was confined to the awareness of the possible ANs allocated to a certain state and the GNs issued in a certain year or years. Based on such limited knowledge, SSN inferences described in the literature would start from known SSNs and predict, based on their digits, possible states and ranges of years when those SSNs could have been issued (15). We conjectured, however, that the functional relationship between the digits of an SSN and the location and time of its application could be reversed, allowing the inference of all of the digits of unknown SSNs starting from their consecutive state and day of application. Empirical observation of SSA policies—particularly the Enumeration of Birth (EBA) statute, which started collecting birth records in 1908 (16)—drew the connection

Predicting SSN

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDRABAD

Predicting Social Security numbers from public data
Alessandro Acquisti¹ and Ralph Gross
Carnegie Mellon University, Pittsburgh, PA 15213
Communicated by Stephen E. Fienberg, Carnegie Mellon University, Pittsburgh, PA, May 5, 2009 (received for review January 18, 2009)

Information about an individual's place and date of birth can be exploited to predict his or her Social Security number (SSN). Using only publicly available information, we observed a correlation between individuals' SSNs and their birth data and found that for younger cohorts the correlation allows statistical inference of private SSNs. The inferences are made possible by the public availability of the Social Security Administration's Death Master File and the widespread accessibility of personal information from multiple sources, such as data brokers or profiles on social networking sites. Our results highlight the unexpected privacy consequences of the complex interactions among multiple data sources in modern information economies and quantify privacy risks associated with information revelation in public forums.

identity theft | online social networks | privacy | statistical reidentification

In modern information economies, sensitive personal data hide in plain sight amid transactions that rely on their privacy yet require their unimpeded circulation. Such is the case with Social Security numbers in the United States. Created as identifiers for accounts tracking individual earnings (1), they have turned into sensitive authentication devices (2), becoming one of the pieces of information most often sought by identity thieves. The Social Security Administration (SSA), which issues them, has urged individuals to keep SSNs confidential (3), coordinating with legislators to reduce their public exposure (4).⁵ After embarrassing breaches, private sector entities also have attempted to strengthen the protection of their customers' and employees' data (7).⁸ However, this has not been enough to prevent the SSA from releasing information about the process through which ANs, GNs, and SSNs are issued (1). ANs are currently assigned based on the zipcode of the mailing address provided in the SSN application form [RM00201.030] (1). Low-population states and certain U.S. possessions are allocated 1 AN each, whereas other states are allocated sets of ANs (for instance, an individual applying from a zipcode within New York state may be assigned any of 87 possible first 3 SSN digits). Within each SSA area, GNs are assigned in a precise but nonconsecutive order between 01 and 99 [RM00201.030] (1). Both the sets of ANs assigned to different states and the sequence of GNs are publicly available (see www.socialsecurity.gov/employers/stateweb.htm and www.ssa.gov/history/ssn/gscand.html). Finally, within each GN, SSNs are assigned "consecutively from 0001 through 9999" (13) (see also [RM00201.030], ref. 1).

The existence of such patterns is well known (14), and has been used to catch impostors posing with invalid or unlikely SSNs (15). However, outside the SSA, the understanding of these patterns was confined to the awareness of the possible ANs allocated to a certain state and the GNs issued in a certain year or years. Based on such limited knowledge, SSN inferences described in the literature would start from known SSNs and predict, based on their digits, possible states and ranges of years when those SSNs could have been issued (15). We conjectured, however, that the functional relationship between the digits of an SSN and the location and time of its application could be reversed, allowing the inference of all of the digits of unknown SSNs starting from their consecutive

Predicting SSN

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDRABAD

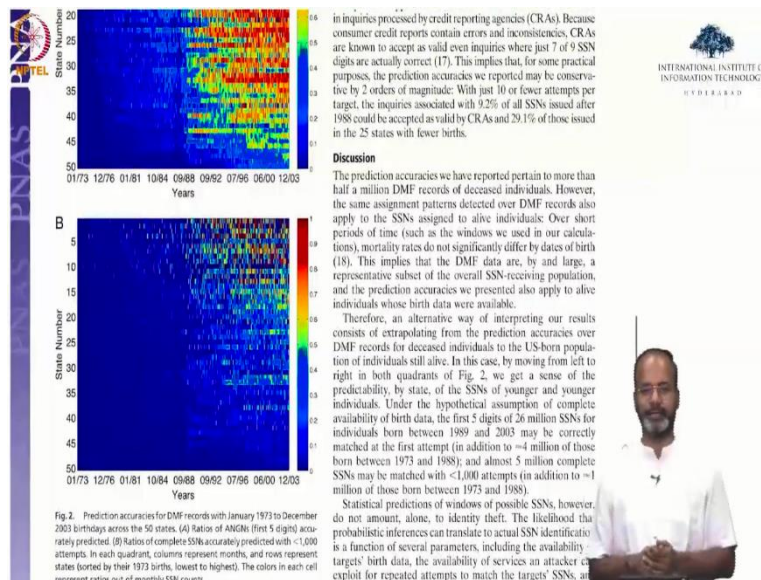


Fig. 2. Prediction accuracies for DMF records with January 1973 to December 2003 birthdays across the 50 states. (A) Ratios of AN:GN (first 5 digits) accurately predicted. (B) Ratios of complete SSNs accurately predicted with $\leq 1,000$ attempts. In each quadrant, columns represent months, and rows represent states (sorted by their 1973 births, lowest to highest). The colors in each cell represent ratios out of monthly SSN counts.

in inquiries processed by credit reporting agencies (CRAs). Because consumer credit reports contain errors and inconsistencies, CRAs are known to accept as valid even inquiries where just 7 of 9 SSN digits are actually correct (17). This implies that, for some practical purposes, the prediction accuracies we reported may be conservative by 2 orders of magnitude: With just 10 or fewer attempts per target, the inquiries associated with 9.2% of all SSNs issued after 1988 could be accepted as valid by CRAs and 29.1% of those issued in the 25 states with fewer births.

Discussion

The prediction accuracies we have reported pertain to more than half a million DMF records of deceased individuals. However, the same assignment patterns detected over DMF records also apply to the SSNs assigned to alive individuals: Over short periods of time (such as the windows we used in our calculations), mortality rates do not significantly differ by dates of birth (18). This implies that the DMF data are, by and large, a representative subset of the overall SSN-receiving population, and the prediction accuracies we presented also apply to alive individuals whose birth data were available.

Therefore, an alternative way of interpreting our results consists of extrapolating from the prediction accuracies over DMF records for deceased individuals to the US-born population of individuals still alive. In this case, by moving from left to right in both quadrants of Fig. 2, we get a sense of the predictability, by state, of the SSNs of younger and younger individuals. Under the hypothetical assumption of complete availability of birth data, the first 5 digits of 26 million SSNs for individuals born between 1989 and 2003 may be correctly matched at the first attempt (in addition to ~ 4 million of those born between 1973 and 1988); and almost 5 million complete SSNs may be matched with $\leq 1,000$ attempts (in addition to ~ 1 million of those born between 1973 and 1988).

Statistical predictions of windows of possible SSNs, however, do not amount, alone, to identity theft. The likelihood that probabilistic inferences can translate to actual SSN identification is a function of several parameters, including the availability of birth data, the availability of services an attacker can exploit for repeated attempts to match the targets' SSNs, and

So, here is a paper, we would not get into details of full details of this paper. But here is the paper, which actually looked at this, which is predicting social security numbers from public data.

They went and argued that look, we could actually figure out social security numbers because of various reasons, which is when the number could have been given, given the number, you could actually predict over this number might have been given in this period, and therefore this person would have been born around this period, all these kinds of predictions they were trying to do.

And with the data that was publicly only available, that is the cool part about it, that they used only the publicly available information to predict the social security number. So, yeah, please

take a look at the paper if you are interested. And if there is any questions, feel free to post it on the mailing list, I will be happy to take it there. So, that was about using social data for predicting some sensitive information.

(Refer Slide Time: 35:55)

Activity Mandatory
Edward Snowden, Permanent Record

<https://youtu.be/PArFP7ZJrtg>

So, as I said, every week I will get you to watch some videos, TED Talks, documentaries, all that for this week. So I will try and mark it as mandatory some and I will try mark some as optional. Because I think for documentaries on Netflix, and all I do not know whether to make it mandatory, but I am, but I am sure you as all of you may have a Netflix or Prime Video account. But try and watch it even if it is optional.

So, this one is actually an interview. Trevor Noah and Edward Snowden I do not know how many of you have read the book permanent record. But this interview, give you an insight this is a very recent interview also would give you a good insight on how Edward has been thinking about this idea called mass surveillance?

And what are the things that he thinks would help in setting my what are the complications that is going on in terms of whistleblowing within the government that he did?

(Refer Slide Time: 37:23)

<https://youtu.be/Pa1FP7Z1t8>

NPTEL

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDRABAD

Activity Optional

- Watch Snowden documentary; please watch it again, if you have watched it before 😊
- Submit your thoughts on
 - Privacy concerns that the documentary highlights
 - Things you learned (small or big) that you never knew about
 - Things you can do to help broader audience be aware of the topic discussed in the documentary

49

This is an optional one, just because it is, I think this documentary is on Netflix. But this will build on meaning if you watch the interview, and this documentary will give you a good sense of what happened. And what is the current state of Edward, in terms of his activities.

So, this is a documentary which talks about what he did within the government and whistleblowing all of that. What do I want you to submit when you have watched both these videos, a privacy concerns that the documentary highlights or the view video highlights? Things that you learnt small or big, that you never knew about? There must be something that is showing up in these videos, which that you did not know before.

I think he would talk about mass surveillance, he would talk about legal rights that citizens have in terms of protecting yourself, if you were to become a whistle-blower, things you can do to help broader audience be aware. So, one of the ideas for me to get you to watch these videos is that meaning I think these will give you a good sense of what the problem is.

But I think we should also think about what are the methods to tell others about this problem, can you actually think of awareness creation of these topics to others, because I think more people understand these kind of ideas, mass surveillance, privacy, anonymity all that I think the general usage, digital literacy would increase and in the process, I think our all of our experiences on using these tools will also increase, which is actually pretty low in India to start with.

(Refer Slide Time: 40:42)

NPTEL

51

INTERNATIONAL INSTITUTE OF
INFORMATION TECHNOLOGY
HYDRABAD

pk.profgiri

Ponnurangam.kumaraguru

/in/ponguru

ponguru

pk.guru@iiit.ac.in

Thank you
for attending
the class!!!

So, that was week 3 and please again use the mailing list effectively to talk about the topics. Any questions about the lectures feel free to ask, we would also do this online session again to just make you feel comfortable with the topic and also bring out some of the other topics that you may be interested in. Thanks again for watching the lectures. See you soon on week 4.