# Privacy in the age of big data

Sourendu Gupta, TIFR

HCAP Meeting
St Xavier's College, Mumbai
16 March, 2015

# Introduction


YOU SEEM MUCH NICER THAN YOUR FACEBOOK PAGE.

# Three keywords

- Privacy
- Security
- Confidentiality

# This talk is not about security

✗ This talk will not teach you how to keep your passwords safe

✗ I do not plan to talk about how to secure your mobile or tablet from snoopers

✗ I will not discuss how to take backups from your laptop and store them securely

✗ I will assume that you know that you may be criminally liable if a criminal uses your devices and accounts

# What privacy?

# Privacy is a legal right

- Right to privacy not a fundamental right

- The supreme court of India ruled that Articles 19 and 21 together imply the right to privacy (1978: Maneka Gandhi vs Union of India)

- Limited rights given to the state in the interests of maintaining law and order and security.

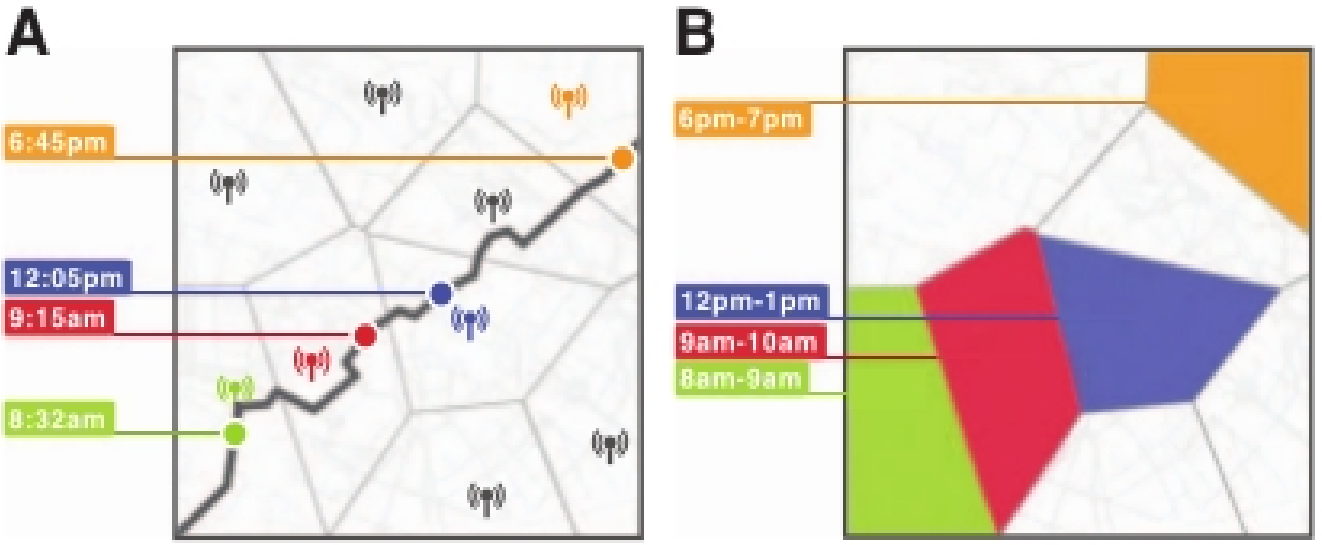
# What does Google sell?

You

# Is there anonymity in a database?

Yes, but limited

# Who called?



# Who called?

3 mobile phone calls are enough to identify a person with 90% certainty

de Montjoye et al (2012)

# Who are you?

# The netflix/IMDb hack

Narayanan Shmatikov (2006)

# Who did you vote for?

## The vote booth hack

Unauthenticated reports (2013)

# Differential privacy

Distill knowledge about you from a database which does not have you.

Dwork (2006)

# Am I unique?

Quite likely.
That is the problem.

# Structural issues


WELL, HOW ELSE AM I SUPPOSED TO KNOW WHO'S BEEN NAUGHTY OR NICE.

# Data ownership

## Do you own your personal data?

EU draft law: others can collect data on you, but may not disclose it to a third party; some provisions in Indian IT Act (2000) and amendment (2008)

What if you don't know your own data?

# Data permanence

## Can you take back something you should not have said?

EU Court ruling of 2014: the (constrained) right to be forgotten

# Data monopoly

## Can you decide who gets to store your data?

MRTP Act 1969: protection from monopolies, except those owned by the government and financial institutions

# Common sense

# Avoid single vendor

Use different vendors for mail, photos, blogs

Do not put the same information in several places

# Limit information leakage

Randomize inessential information sought by websites.

# Take control of robots

Control cookies: set your browser to work for you.

Make your choice of balance between convenience and privacy

# Care is not paranoia

> 7 billion people
> 1 billion PCs in use

If you can imagine it,
someone is doing it.

# Summary

# What this talk was about

- Privacy is a legal right in India.
- Privacy depends on security and confidentiality of your service provider.
- Massive data sets and very fast computation pose new challenges.
- Legal framework necessary; evolving