



MAX PLANCK INSTITUTE
FOR SOFTWARE SYSTEMS

The good and the bad of privacy in Social Media

Mainack Mondal

IIT KGP, Social Computing course, August 2017

Roadmap

This Lecture

How to **better protect privacy** in Online social media sites (OSMs) – the good of privacy

Next lecture

Online abuse: The ill side-effect of privacy and how to defend against the online abuse – the bad of privacy

Roadmap

This Lecture

How to **better protect privacy** in Online social media sites (OSMs)

Next lecture

Online abuse: The ill side-effect of privacy and how to defend against the online abuse

What is privacy

Warren and Brandeis – privacy as “right to be let alone”

Alan Westin’s definition of privacy

Altman’s privacy boundary theory

Sandra Petronio’s CPM theory

Palen and Dourish’s boundary regulation

Solove’s privacy taxonomy

Nissenbaum’s contextual integrity

What is privacy

Warren and Brandeis – privacy as “right to be let alone”

Alan Westin's definition of privacy

Altman's privacy boundary theory

Sandra Petronio's CPM theory

Palen and Dourish's boundary regulation

Solove's privacy taxonomy

Nissenbaum's contextual integrity

Westin's definition of privacy

*“**Privacy** is the ability for people to determine for themselves when, how, and to what extent, **information about them is communicated to others**” - A. Westin. *Privacy and Freedom*, 1970*

Four key states of privacy

Anonymity: freedom from identification

Solitude: being free from observation by others

Intimacy: limiting communication within a small group

Reserve: desire to limit information disclosure to others

The need for anonymity in OSMs

“The freedom from identification and from surveillance...”

- A. Westin.

140 CHARACTERS

Online Activists Harassed & Jailed in Arab Gulf States

In a nod to Twitter's 140-character limit, this report presents the profiles of 140 prominent Bahraini, Kuwaiti, Omani, Qatari, Saudi, and Emirati social and political rights activists and dissidents, and their struggles to resist government efforts to silence them. All 140 have faced government retaliation for exercising their right to freedom of expression, and many have been arrested, tried, and sentenced to fines or prison terms.

[Enter the site »](#)

THE VERGE TECH ▾ SCIENCE ▾ CULTURE ▾ MORE ≡

POLICY & LAW \ US & WORLD \ TECH \

Facebook, Twitter, and Instagram surveillance tool was used to arrest Baltimore protestors

The services have revoked access after an ACLU report

by Russell Brandom | @russellbrandom | Oct 11, 2016, 1:42pm EDT

How is anonymity enforced in (some) OSMs

Anonymous social networks are emerging

Whisper, secret.ly, Yik Yak

No unique usernames or profile information



The need for intimacy and solitude in OSMs

Information should reach only the intended audience



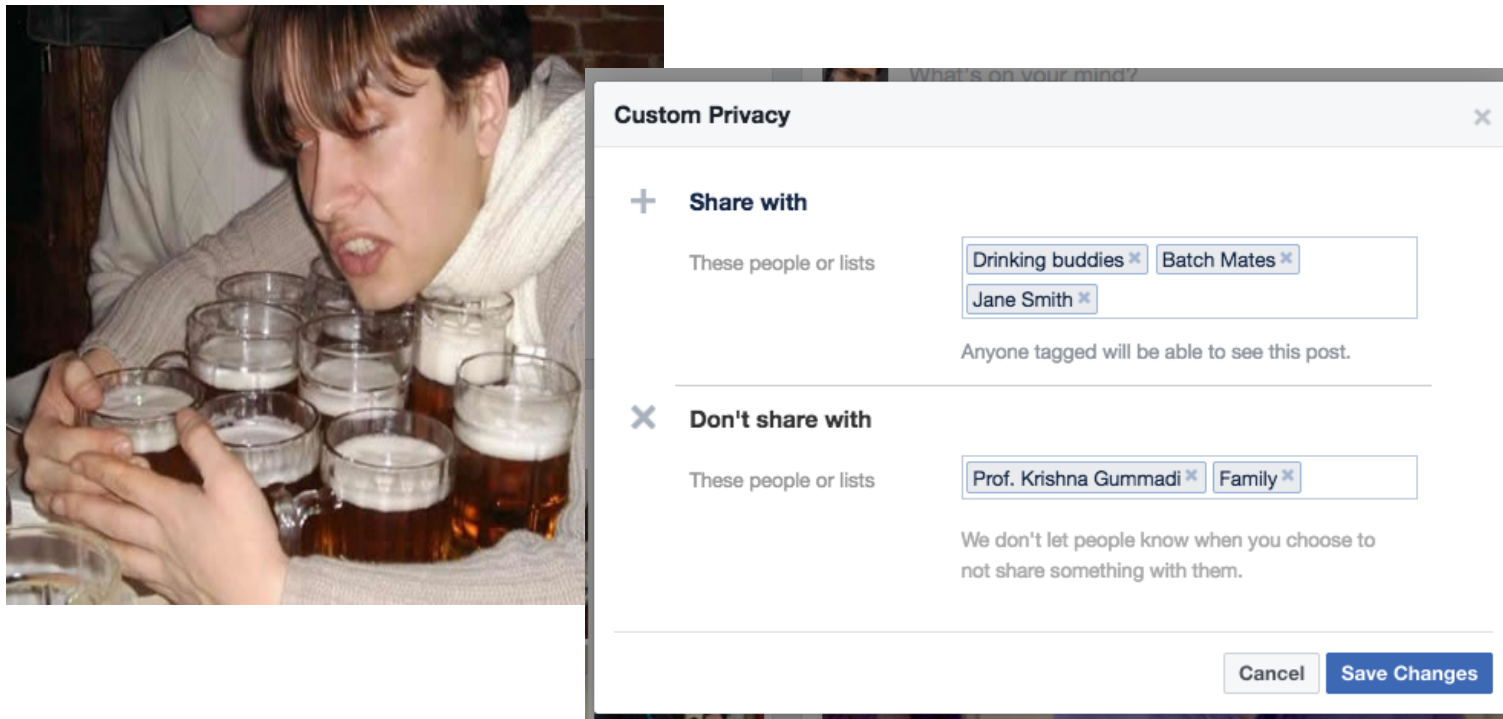
Clearly the rant was meant for close friends, not the boss

Easy way out: users need to **restrict access** to their content

How do users set restrictive access control?

Users specify **Social Access Control Lists** (SACLs)

SACLs: Share with a subset of friends



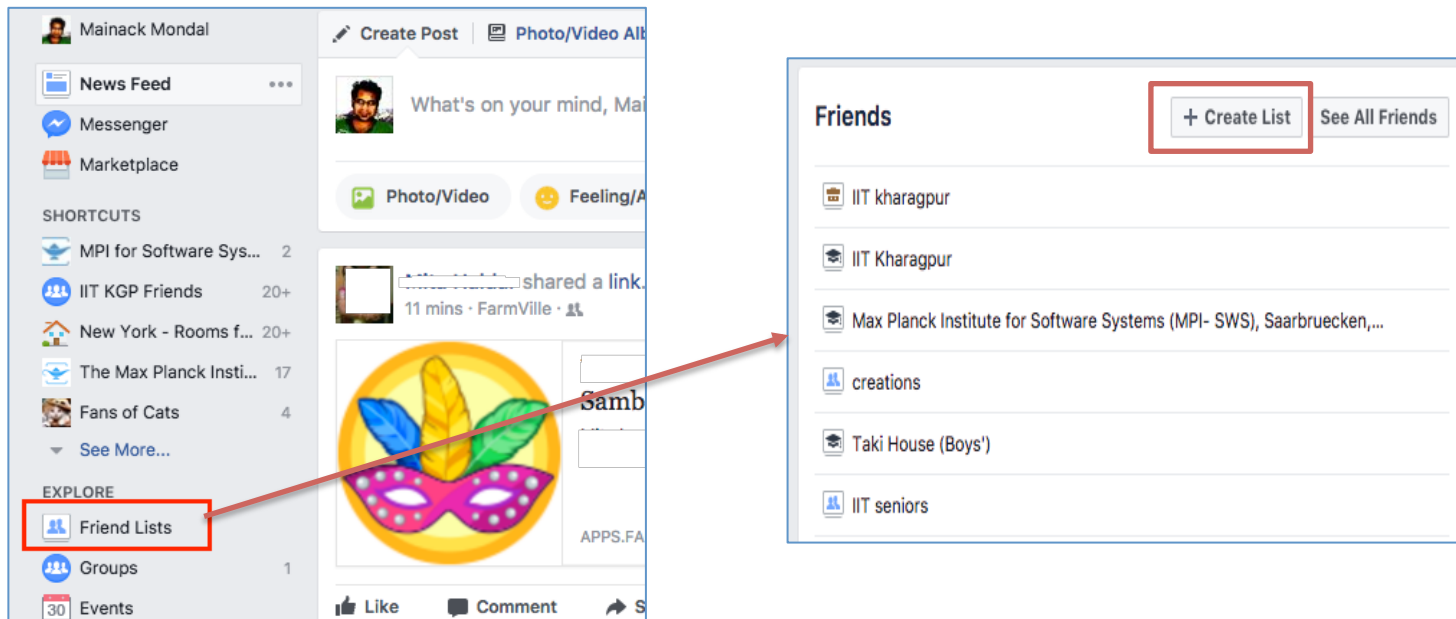
Each time users have to **specify** these SACLs **manually!**

Friendlists: Helping users to choose subsets of friends

Facebook allows a user to pre-create groups of friends

These groups are called friendlists

These friendlists can be used to create SACs

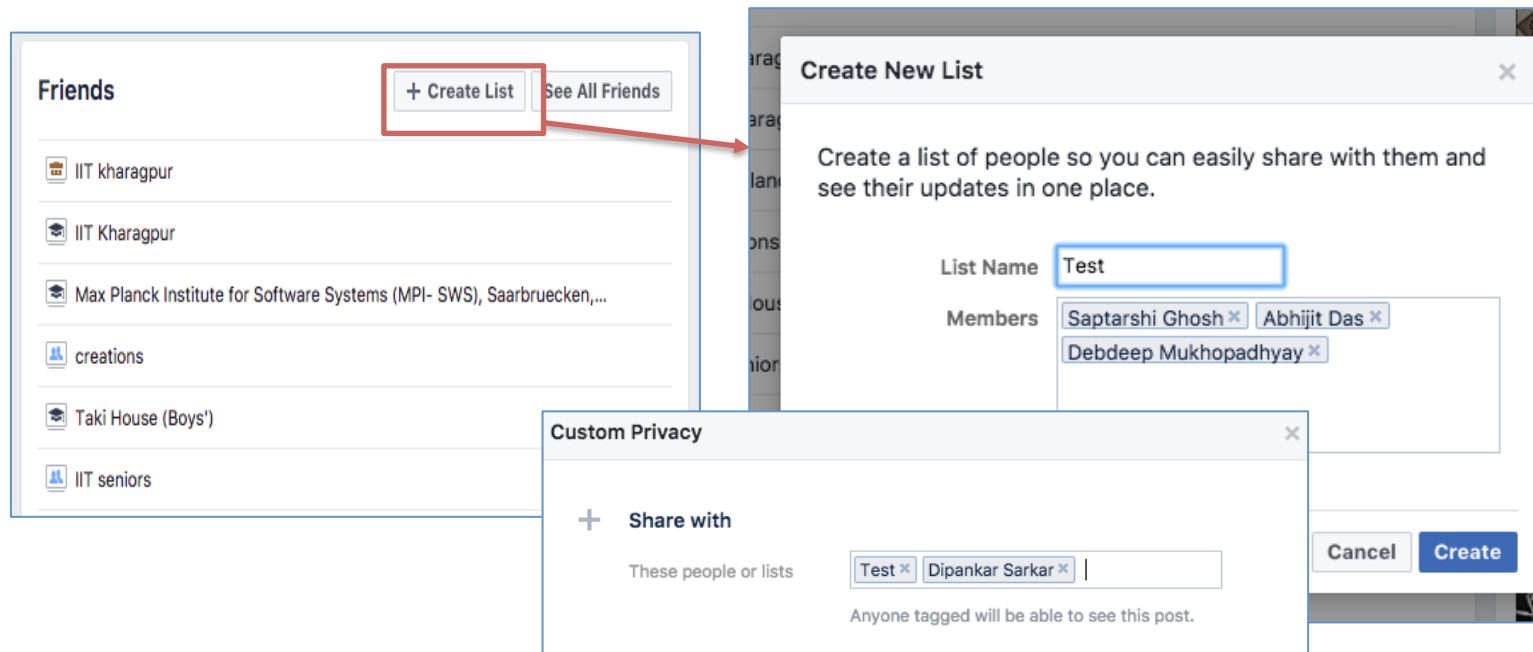


Friendlists: Helping users to choose subsets of friends

Facebook allows a user to pre-create groups of friends

These groups are called “friendlists”

These friendlists can be used to create SACs



Friendlist limitations

User required to

create friendlists and enumerate members

maintain friendlists as relationships change

maintain friendlists as new relationships form

Create New List

Create a list of people so you can easily share with them and see their updates in one place.

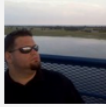
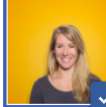
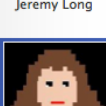
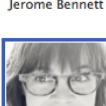
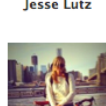
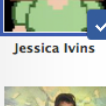
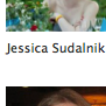
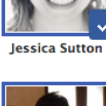
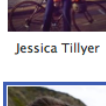
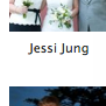
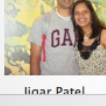
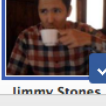
List Name:

Members:

☐ Bimal Viswanath Saarbrücken

Edit My Work Friends

Friends

 Jennifer Bethea	 Jennifer Brennan	 Jennifer Rose Gaunt	 Jenny Pfof	 Jen Xiques
 Jeremy Long	 Jermal M Smith	 Jerome Bennett	 Jesse Lutz	 Jessica Gavit
 Jessica Ivins	 Jessica Sudalnik	 Jessica Sutton	 Jessica Tillyer	 Jessi Jung
 Jinar Patel	 Jill Whitesell	 Jimmy Stones	 Jim Schachterle	 Jin Cronov

Add and remove people from this list by clicking on their names.

Can we **automatically manage** friend groups?

Automatically manage friend groups

Idea: Can we automatically infer friend groups

Leveraging the network to find friend groups

Leveraging profile attributes to find friend groups

Leveraging activity to find friend groups

Automatically detecting friend groups

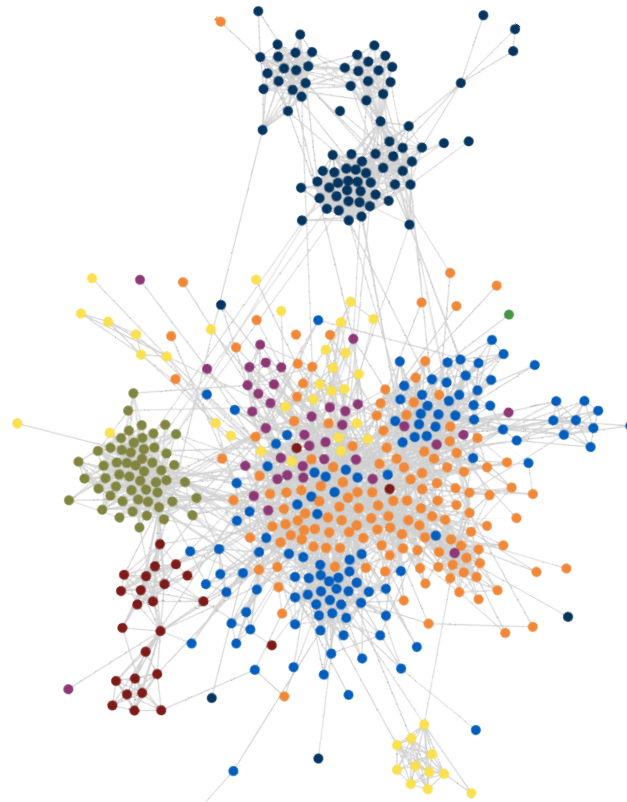
Idea: Can we automatically infer friend groups

Leveraging the network to find friend groups

Leveraging profile attributes to find friend groups

Leveraging activity to find friend groups

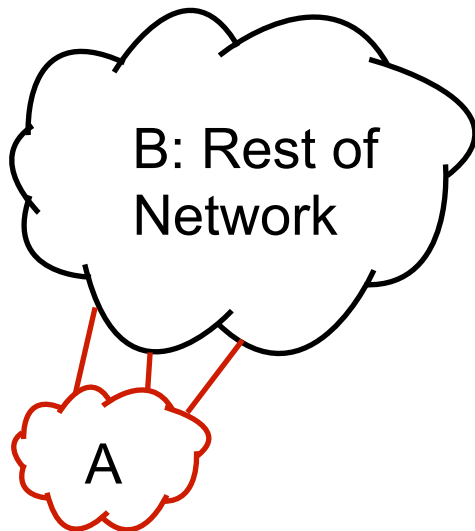
Can we leverage the network?



Communities: Users **more densely connected** than overall graph

Idea: Can we automatically infer friend groups by finding communities?

Can we detect graph communities?



Community detection is well studied
A type of **unsupervised learning**

Define a **metric** for community strength
compares internal & external connectivity
e.g., **Conductance C**: Fraction of A's links within A

$$C = \frac{e_{AA}}{e_{AA} + e_{AB}}$$

Design graph algorithms optimizing for the metric

Do friend groups form communities?

Social Network	# nodes	# links	# groups
Facebook	81,916	1,823,331	154,795
Flickr	1,714,351	15,540,671	103,631
YouTube	1,138,380	2,989,999	30,087
Orkut	3,072,441	117,073,438	8,730,857

Gathered data from four popular social networks

Analyzed connectivity of user-defined friend groups

Results: Connectivity of group members

Internal connectivity: very high

compared to random groups of similar size

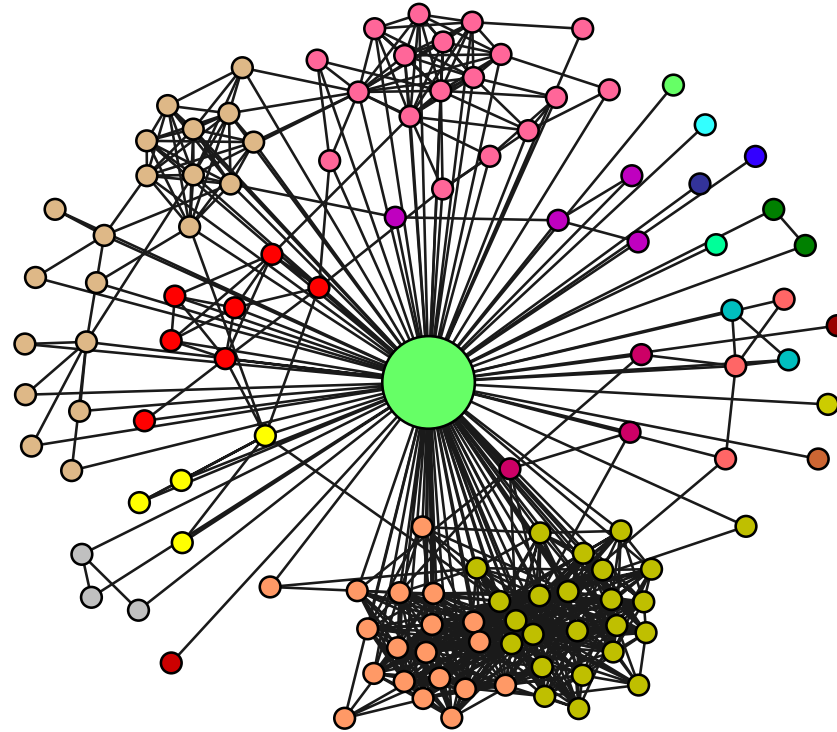
External connectivity: very high

indistinguishable from random groups

Users tend to be members of multiple groups
forming a densely inter-connected network core

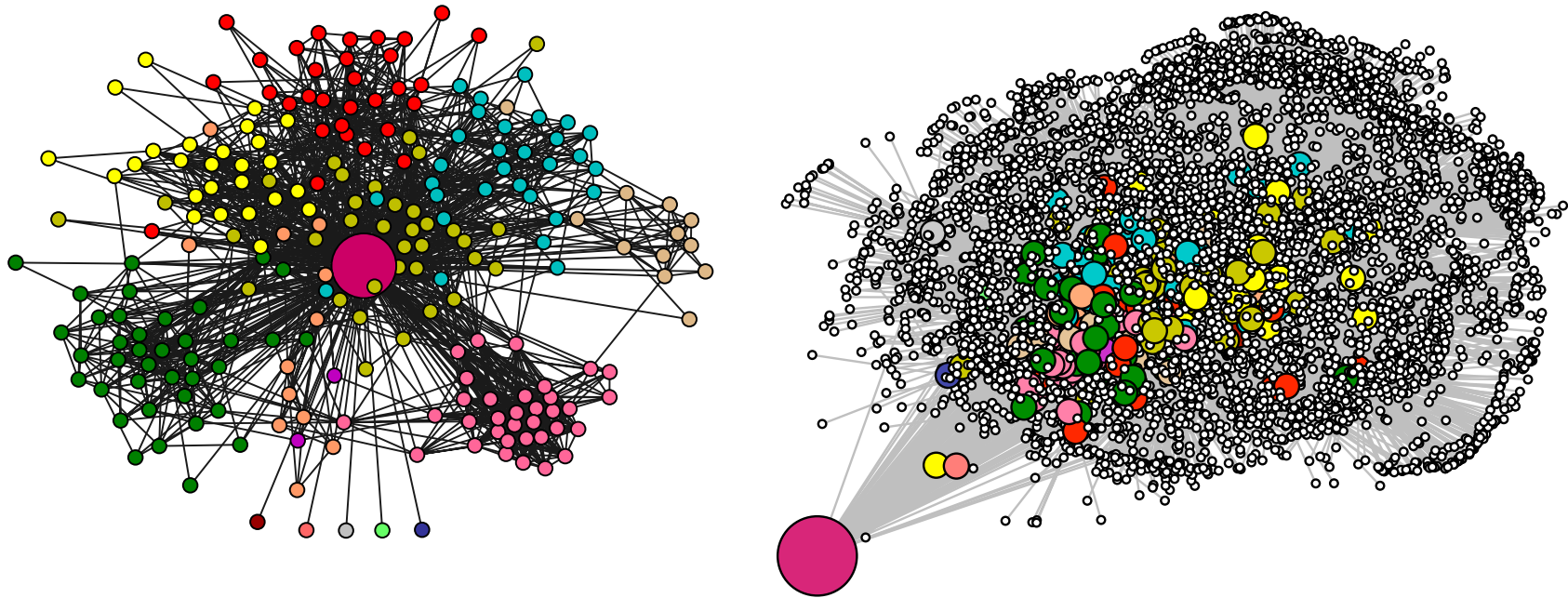
Friend groups **do not form strong clusters**
impossibly hard to detect groups in the dense core

An insight



When we ignore the links outside local friends' network
strong clusters exist within the **1-hop network**
these clusters correspond to meaningful friend groups

Why?



1-hop excludes members' links to rest of the network

less data is more useful!

So, we can **indeed automatically group 1-hop friends!**

Friendlist Manager

Built as a Facebook application [Discontinued now]

Can access user's 1-hop network, modify friendlists

Key idea: Leverage structure of user's 1-hop social network

Challenges:

Creating new friendlists of user's friends

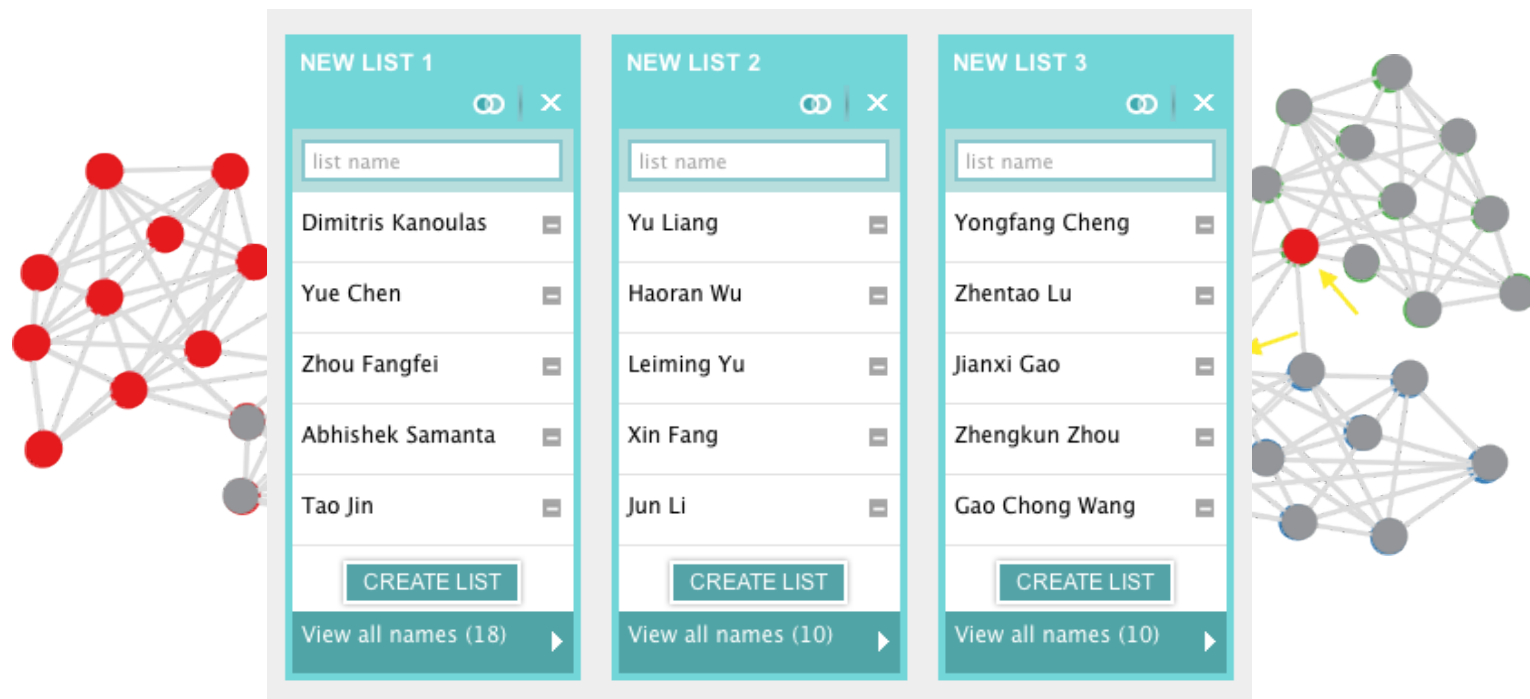
Updating the existing friendlists over time

How to create new friendlists?

Idea: Use community detection to suggest friendlists

Create "seed" friendlists via a global algorithm [JSTAT'08]

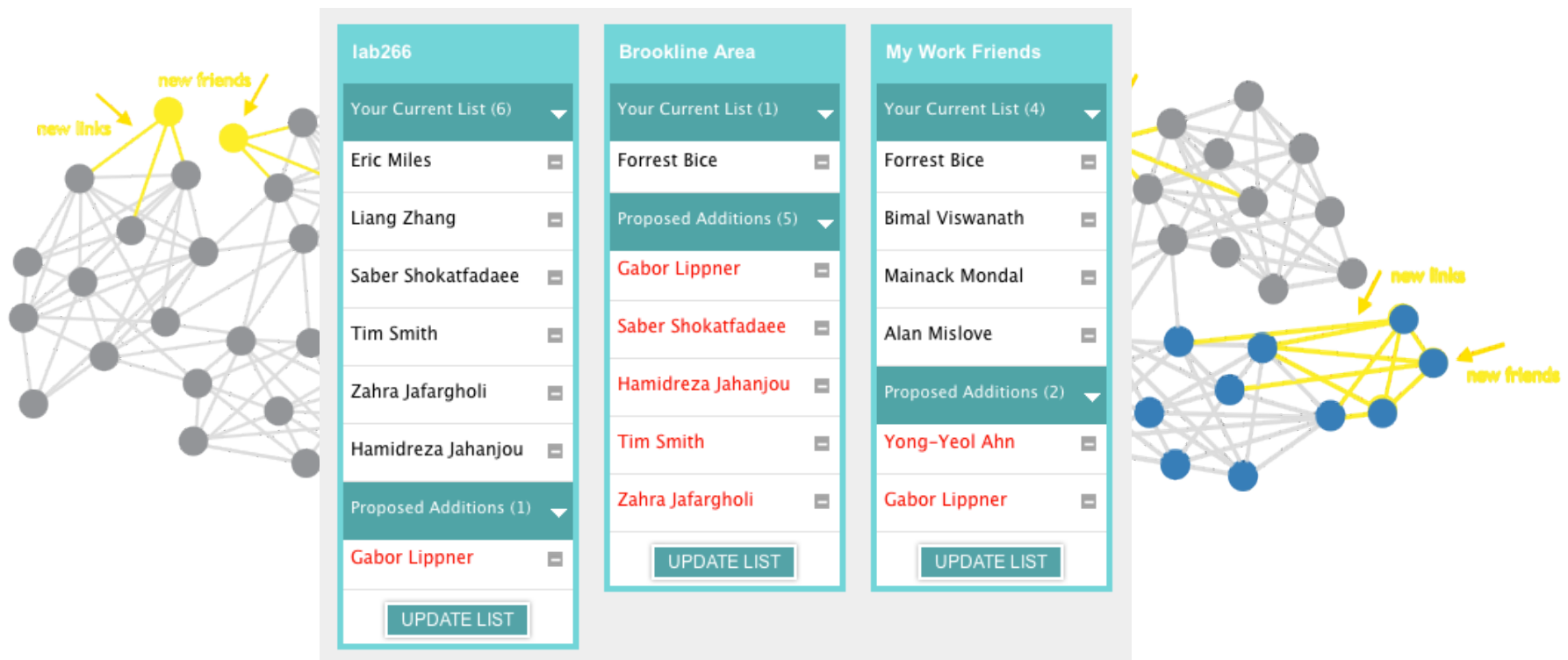
Expand "seeds" via a local algorithm [WSDM'10]



How to expand existing friendlists?

Idea: Locate missing members using a local algorithm
[WSDM'10]

Update each list using the new local social network



What about selecting friends-of-friends?

Unfortunately, community detection does not work well over 2-hop friends

But, we can detect missing members of a friendlist from friends-of-friends

E.g., co-workers missing from 1-hop friends

Using a local cluster detection algorithm

Automatically detecting friend groups

Idea: Can we automatically infer friend groups

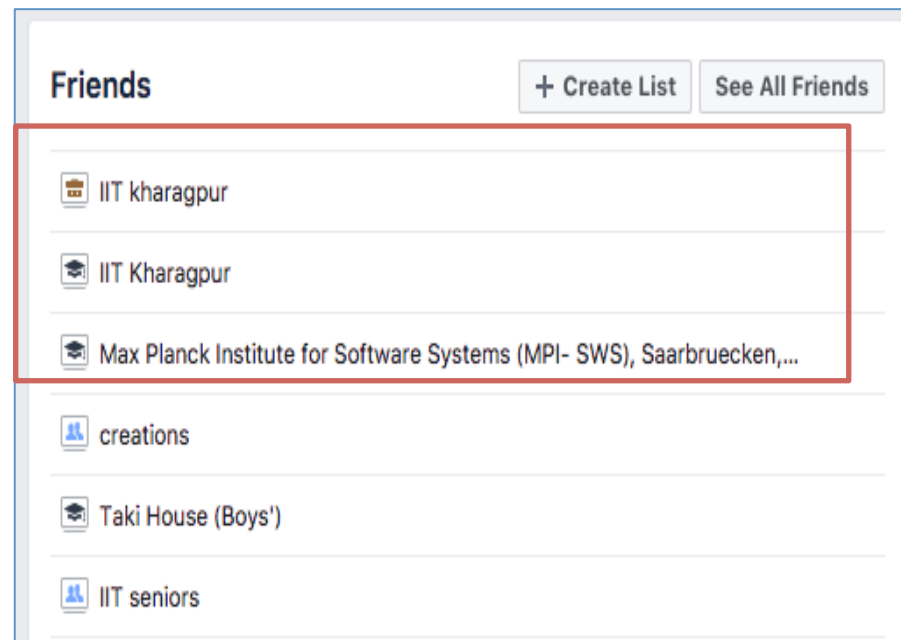
Leveraging the network to find friend groups

Leveraging profile attributes to find friend groups

Leveraging activity to find friend groups

Simply group users who have same profile attribute

OSMs leverage the user reported attributes to do so
So called Facebook smart lists



Automatically detecting friend groups

Idea: Can we automatically infer friend groups

Leveraging the network to find friend groups

Leveraging profile attributes to find friend groups

Leveraging activity to find friend groups

How to leverage activity in OSMs for friend grouping?

We can measure activity for each of a user's friends
How many times liked, commented, tagged

Then we need to group or cluster users based on their number of activities

- friends with “Similar” number of activities should be in same cluster
- friends with “Different” number of activities should be in different cluster

Classic problem solved by **k-means** clustering

What did we learn about privacy management today?

What is privacy

Westin's definition

How can users manage privacy

Many possible mechanisms, SACL is one common mechanism

How to help users manage privacy

Provide them meaningful friend groupings

We explored graph structure, profile attribute and activity based group detection

Read more:

<https://people.mpi-sws.org/~mainack/publications/soups2014-final27.pdf>

<https://people.mpi-sws.org/~mainack/publications/Friendlist-WWW-Demo.pdf>

BUT, Privacy (and anonymity) are Double edged swords

Abuse of privacy and anonymity in OSMs

Hate speech

Abusive language

Cyber bullying

Trolling ...



How to deal with them? Next lecture...