Network analysis and visualization
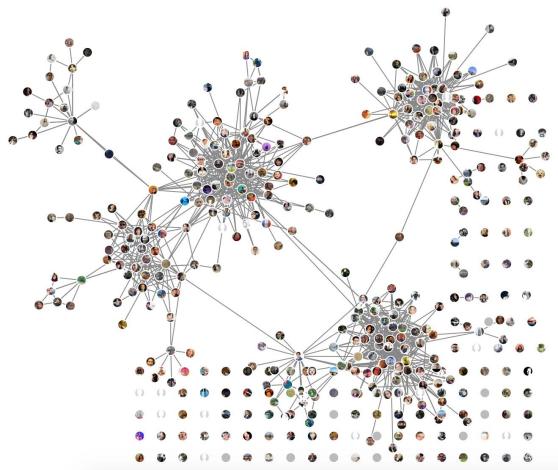
Weiai Xu (Wayne), PhD
Assistant Professor
Department of Communication, UMass-Amherst
Email: weiaixu@umass.edu
curiositybits.cc

A couple of reminders...
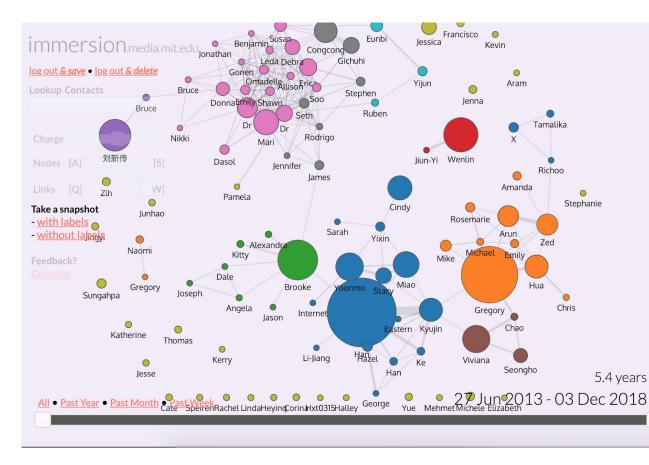
Assignment 3: Friday, March 28, 2019, 11:59 PM

In assignment 3, you will complete the set of challenges in the scripts for text mining.

Please submit the R script that shows your solution to the challenges.
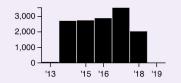
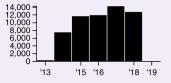# Examples of online social networks

Examples of online social networks
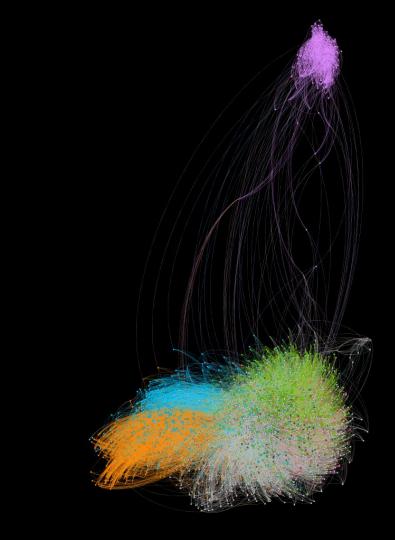
Examples of online social networks

Social networks ≠

Social networks =

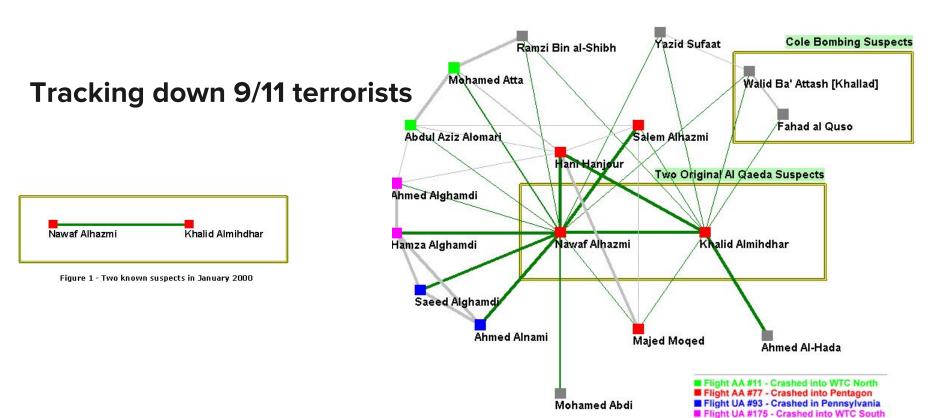$$C_b(n_i) = \frac{\sum\limits_{\substack{j \neq k \\ i \neq j,k}} \dfrac{g_{jk}(n_i)}{g_{jk}}}{(g-1)(g-2)}$$

A mathematical view of relationships and social structure

Questions that can be answered through network analysis

- How to start/stop rumors?
- How to track terrorists?
- How to start or destroy a movement?
- How to divide a nation?
- How to find jobs?
- How to find an effective way to promote a product?

…………

Tracking down 9/11 terrorists

Figure 1 - Two known suspects in January 2000

Cole Bombing Suspects

Two Original Al Qaeda Suspects

Flight AA #11 - Crashed into WTC North
Flight AA #77 - Crashed into Pentagon
Flight UA #93 - Crashed in Pennsylvania
Flight UA #175 - Crashed into WTC South
Other Associates of Hijackers

Copyright © 2001, Valdis Krebs

Figure 2 - All nodes within 1 step [direct link] of original suspects

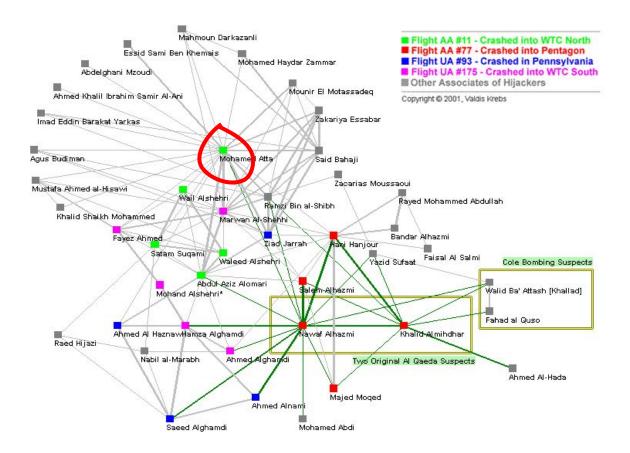# Mapping connections of their connections

Figure 3 - All Nodes within 2 steps / degrees of original suspects

Mapping your offline social network

On the left: list 10 people you last talked to

In a column to the right:
- For each, what is your relationship with them:
- How do you know them?
    - Friend, family, colleague, classmate, acquaintance...

In the third column: who introduce you to that person

XXXX	family	xxx
xxxxx	friend	xxx
xxxx	colleague	xx
xxx	colleague	xxx

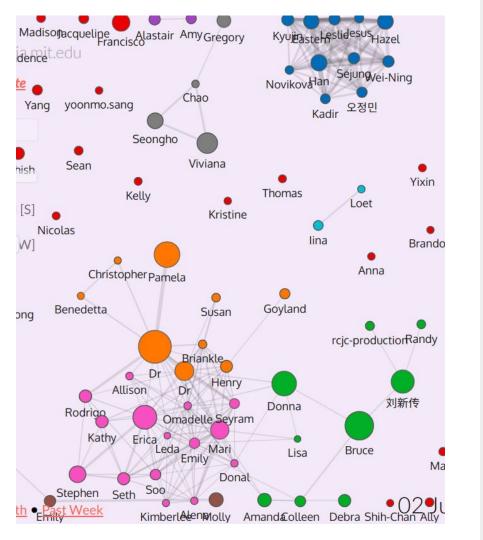# Mapping your offline social network

Flip the paper and on the other side:

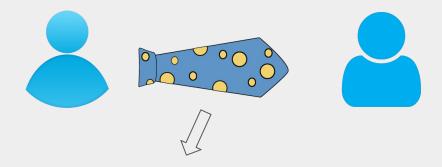Write out the names so that they were spread out all over the paper, like a big circle.

Draw a line from 1 person to another if person 1 knows person 2

Mapping your offline social network

- Does everyone know everyone in the network?
- Does your network consist of two or more separated groups?
- Who has the most connections in your network?

The basis of a connection

Based on friendship, financial transactions, email correspondences, Twitter following, retweeting, etc.

Edges from online activities

So people (internet users) are interconnected through instant messaging, retweets, Twitter following, Facebook friendships, Facebook tagging, emails, website hyperlinking and etc.

What does a tie/edge imply?

1. User A follows user B
2. User A retweets user B
3. User mentions user B
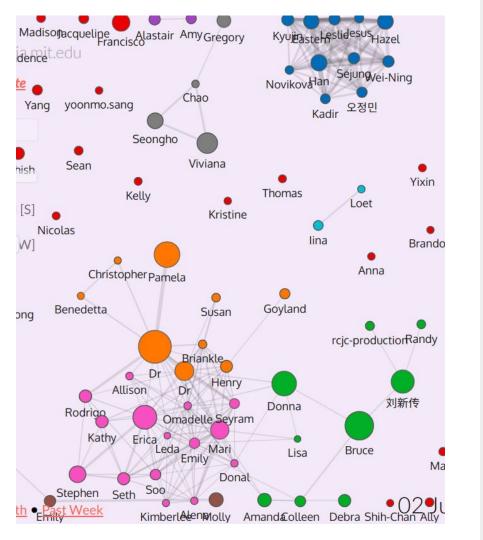
Implication of edges
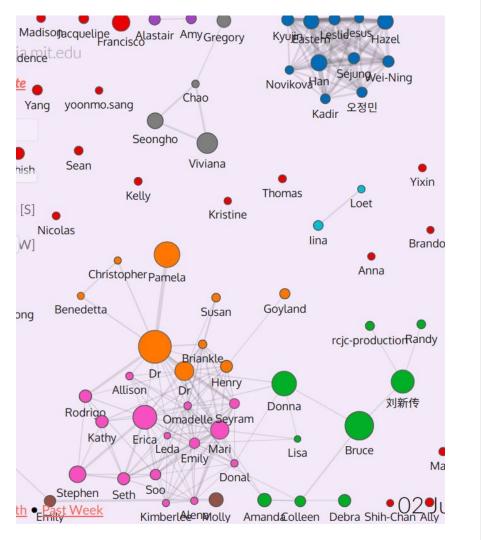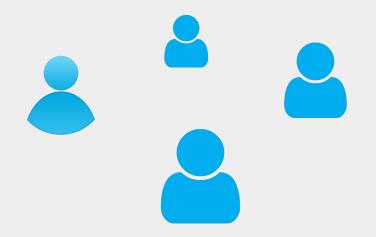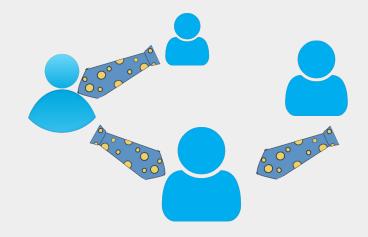
Nodes (also called vertices)

People in a network

Node attributes

Edges (or ties/links)

A connection between two people

Edge attributes: weight

Where network properties may reveal insight

Position in a network matters
The size and shape of a network matters

Location, location, location

Influence

Where a person is located in a network indicates how much influence that person has over other people in the network

Look at your network, can you find the influencer in your network?

But, what kind of influence are we talking about?

Influence

- Pass new information
- Introduce new opportunity
- Resource mobilization
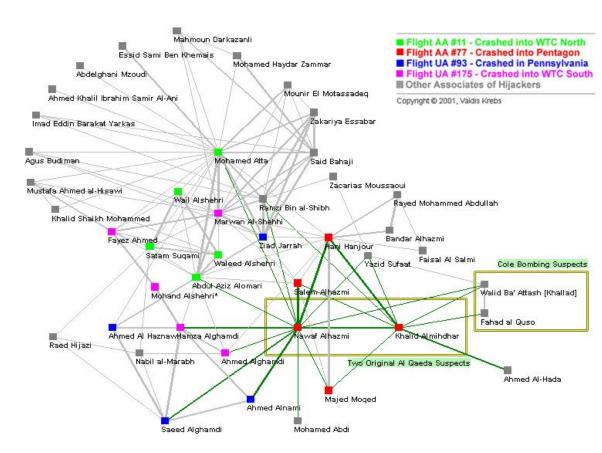- Social capital

Figure 3 - All Nodes within 2 steps / degrees of original suspects

Social capital

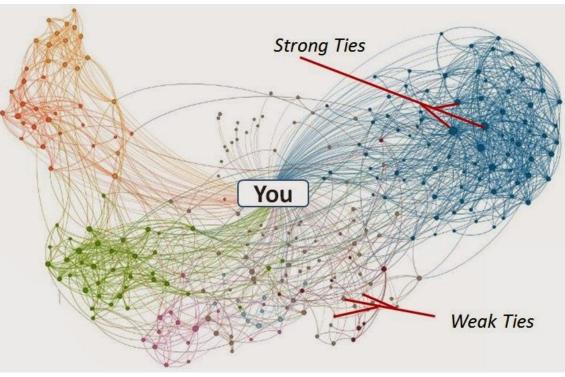Tangible and intangible resources embedded in your social relationships

- Financial support
- Social support
- Career opportunity
- Trust
- ....

Social capital

Strong ties and *weak ties*

Strong ties: someone you know very well or share many similarities

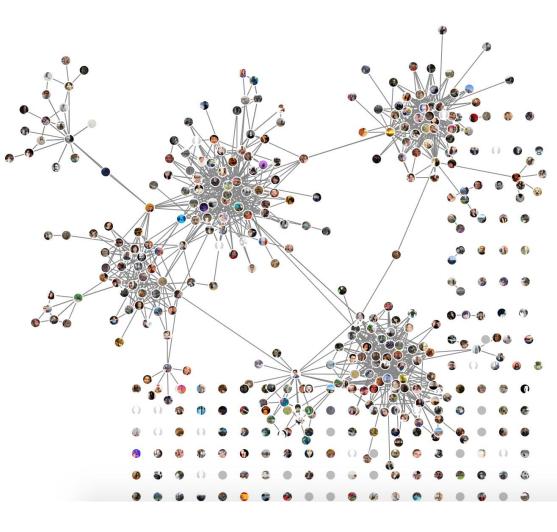**Weak ties:** acquaintances who are likely different from you

Social capital

Who are the strong ties and *weak ties* in your network and what sort of resource can each type of nodes provide?

Where is influence located?

Central position
and/or play a _bridge_ role

Such position can be quantified
using a network metric

Social objects

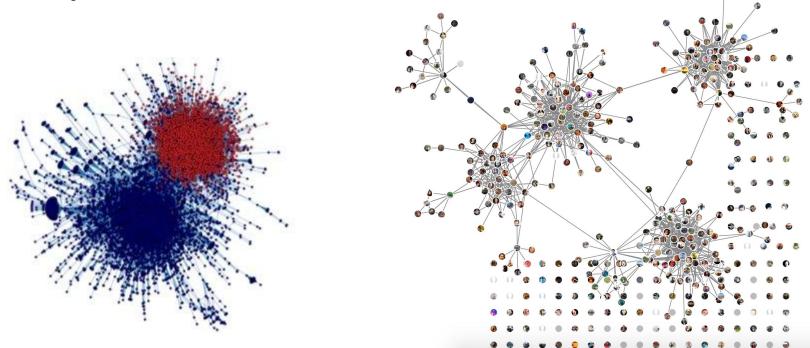*An object through which we socialize*

Homophily: *Birds of a feather flock together*

Some common social objects: proximity, family association, ethnicity, religion, political values, hobbies, etc.

Social objects

Clusterings in a network reflects how people are united or separated by social objects.

Network analysis

Attributes of nodes and all sorts of network characteristics can be expressed in numbers for comparison

Two levels of metrics

Node-level: different centrality scores, a cluster a node belongs to
Network level: size, density, reciprocity, centralization, clustering, etc.
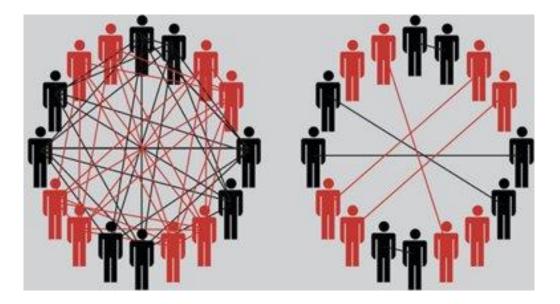
Network size

Size does matter

Number of nodes (i.e., people) in your network.

- How many people do you know
- How many people does an org have
- How many users tweeted a particular hashtags

Network density

The proportion of present edges from all possible edges in the network. A network with 100% density means everyone knows everyone else.

Is your network dense?
What sort of benefits does a dense network provide?
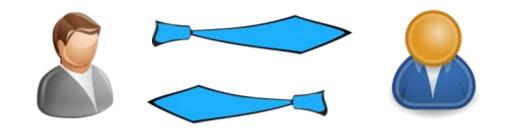What sort of benefits does a sparse network provide?

Reciprocity

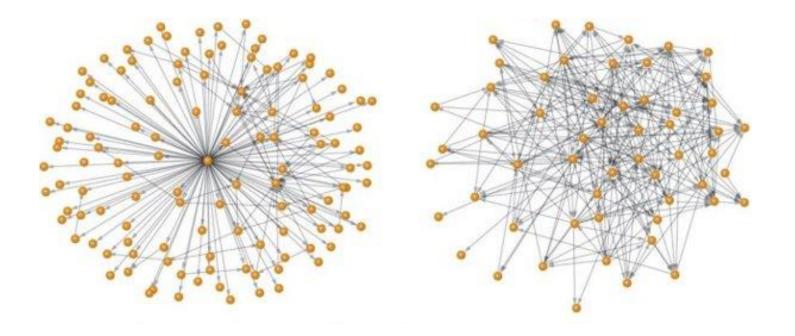the proportion of reciprocated ties

A tie can be by default
mutual or *directed*.

Can you find an edge in
your network that is NOT
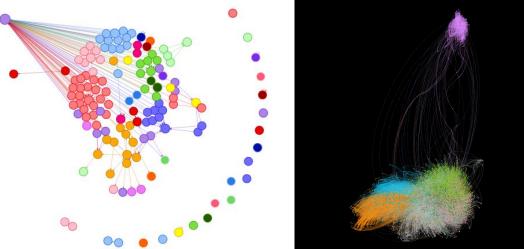reciprocated (A knows B
but B does not know A).

Centralization

Can you think of examples of centralized network and decentralized ones?

Clustering

The technique used to cluster users is called *community detection*. There are different community detection algorithms.

People in a cluster are more interconnected than with people outside of the cluster

Centrality

Centrality is a set of metrics to quantify a person's central position in a network

In-degree, *out-degree*, *betweenness centrality*, *PageRank*, etc.
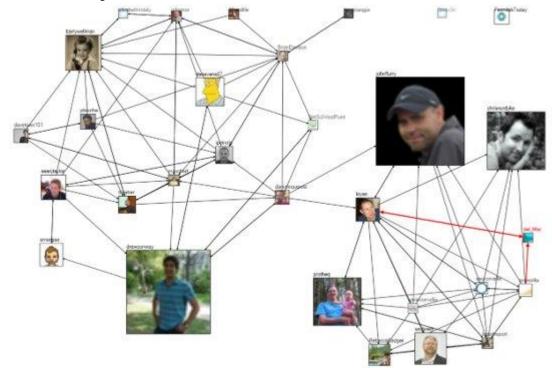
Centrality

In-degree and out-degree

Centrality

Betweenness centrality

Practice

http://rpubs.com/cosmopolitanvan/networks_497db