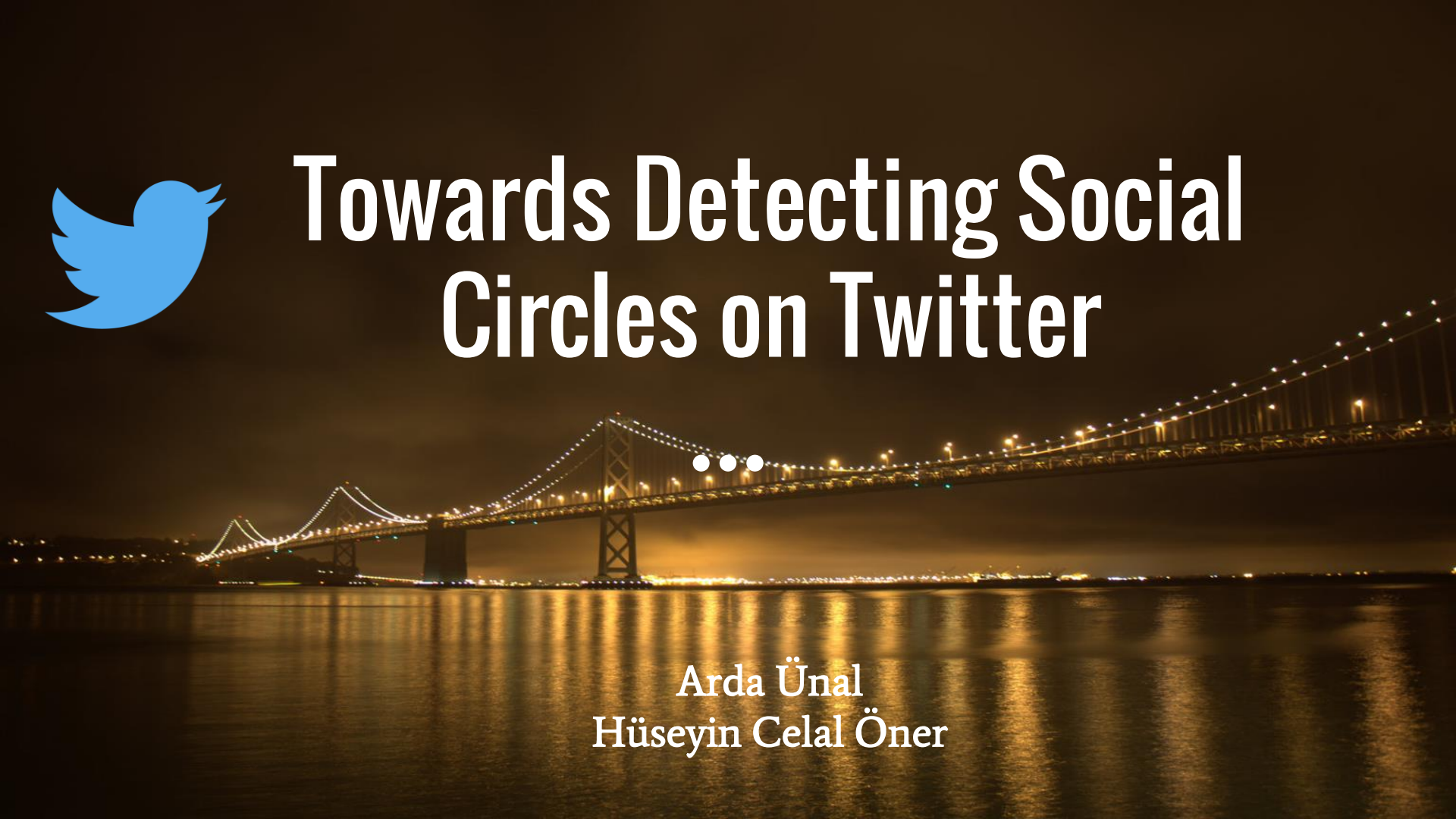


Towards Detecting Social Circles on Twitter

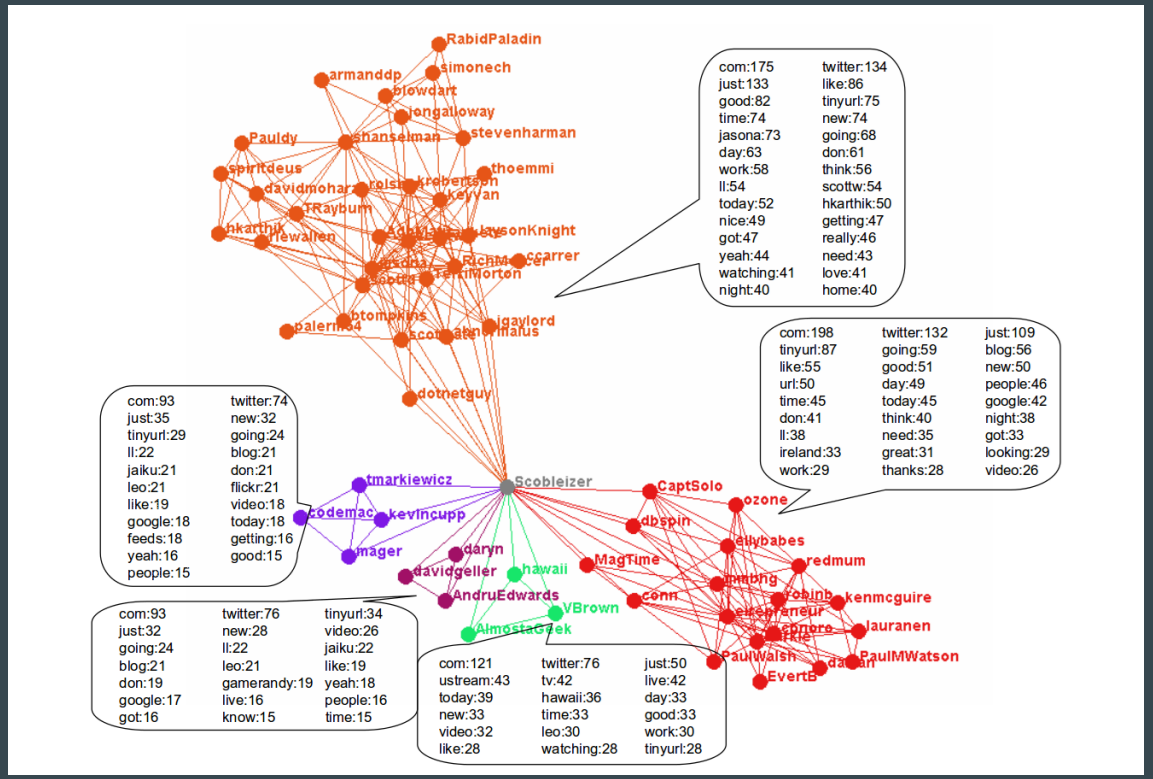
Arda Ünal
Hüseyin Celal Öner

Outline

- Problem Definition
- Applied Methodology
- Data Set
- Evaluation Metric
- Results

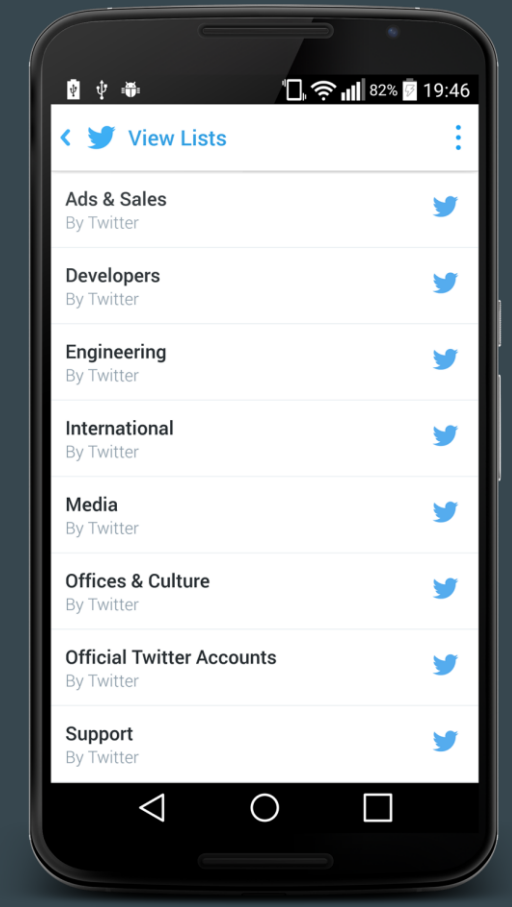
How Do People Use Twitter?

- Many Twitter users follows friends, family, celebrities, sport teams or even brands.
- Barack Obama approximately follows 636,000 users on Twitter.



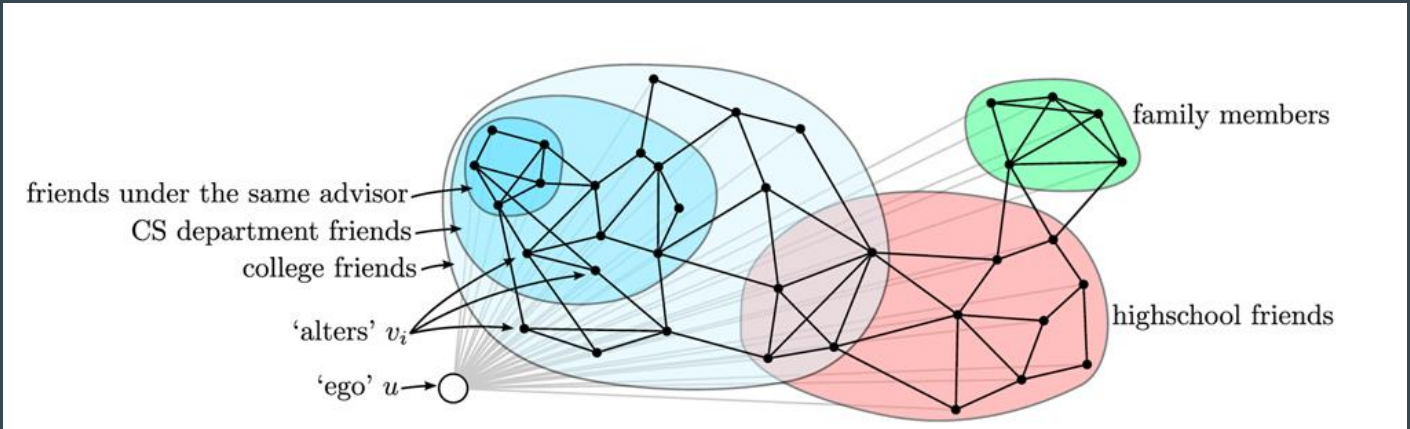
Problem Definition

- The personal social network of many Twitter users is big and untidy.
- In 2009, Twitter released a feature, “lists”
  - Designed to provide a better way to organize information on Twitter.
- Using lists to organize information on Twitter is a tedious task



Project Goal

- Automatically infer users' social circles using the set of connections between users and their social network profiles.



Applied Methodology

1. Smooth out inconsistencies in the basic adjacency matrix so that it gives a better notion of the degree of connectedness of various node. (Exponential Adjacency Matrix)
2. Apply clustering. (Spectral Clustering)
3. Eliminate some circles found by the clustering algorithm if their density is smaller than a threshold
4. Add more people to the circles if they have number of friends over a threshold given in that circle

Data Set

- High quality hand-labeled data from major social networking sites (Facebook, Google+, Twitter)
- All data are available on snap.stanford.edu/data/

	# of ego-networks	# of social circles	# of nodes	# of edges
Twitter	1000	4,869	81,306	1,768,149

Evaluation Metric

Results are evaluated using the edit distance between the predicted social circles and the ground truth. Each of the following operations costs one edit:

- Add a user to an existing circle
- Create a circle with one user
- Remove a user from a circle
- Delete a circle with one user

Evaluation Metric: Example

Your solution contains circles with user ID's => circle1: 3 1 2 circle2: 2 3 circle3: 5 4 6

And the correct solution is => circleA: 4 5 circleB: 1 3 2 4

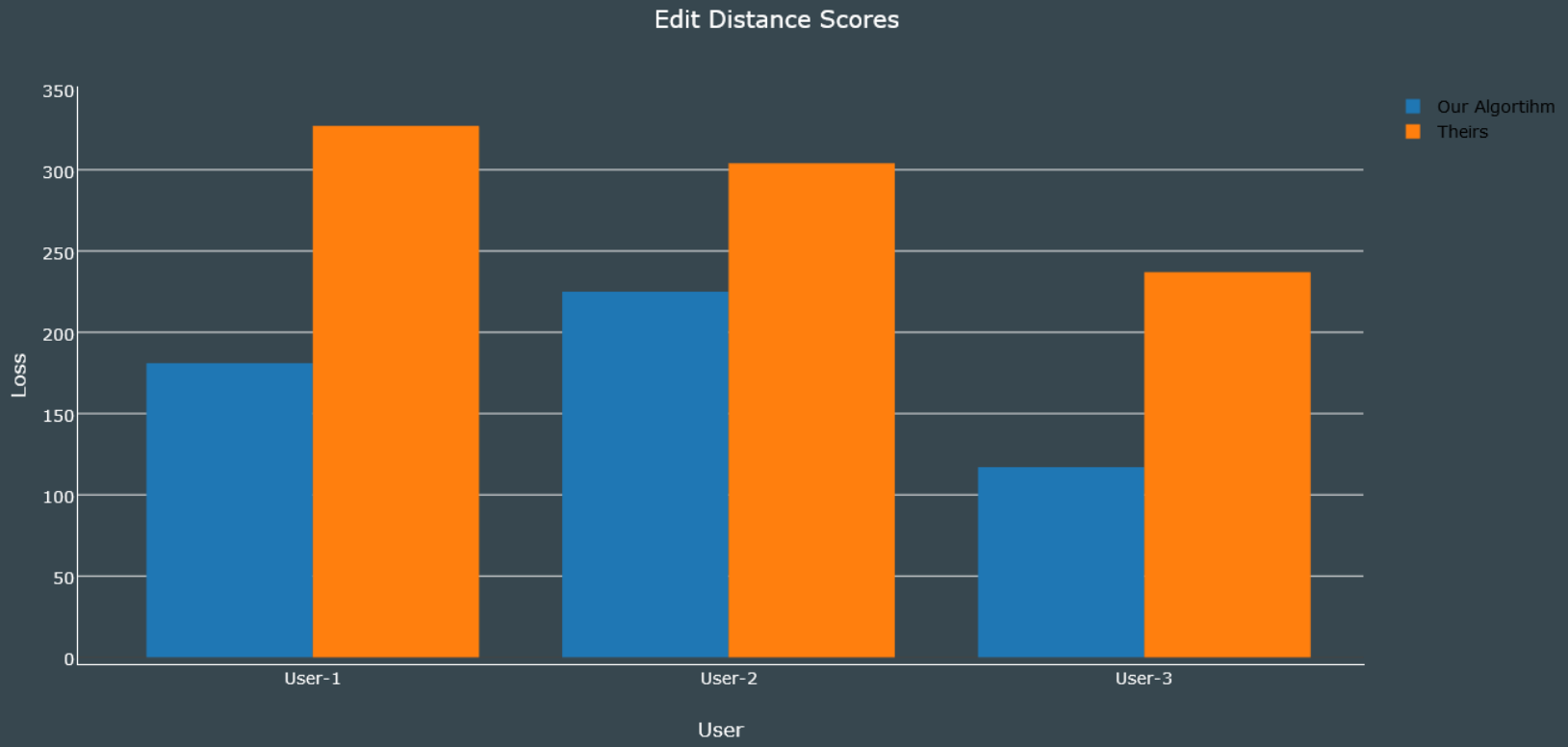
Then the minimum number of required edits would be following 4 edits:

- Add user 4 to circle1
- Delete user 2 from circle2
- Delete user 3 from circle2
- Delete user 6 from circle3

Results

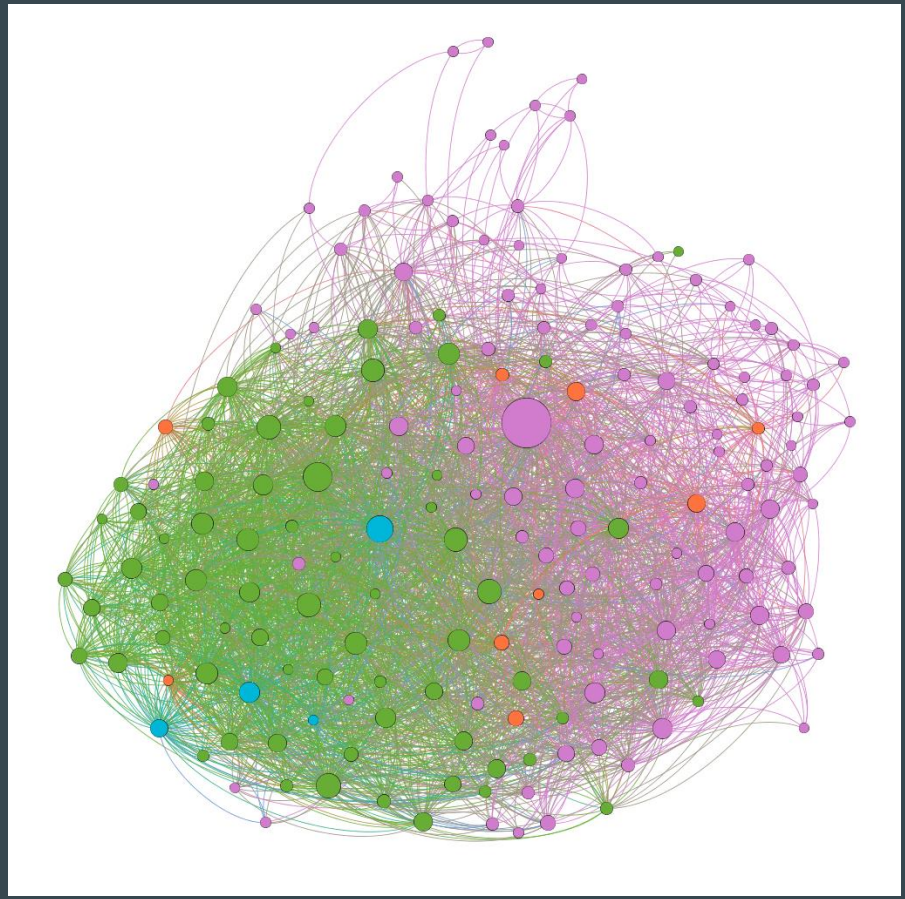
- Found the implementation for the “Learning to discover social circles in ego networks” which is published in *Neural Information Processing Systems (NIPS)* on http://cseweb.ucsd.edu/~jmcauley/
- Our implementation outperforms their implementation based on Kaggle’s evaluation metric mentioned in the previous slide.

Results


Edit Distance Scores
Our Algortihm
Theirs
Loss
350
300
250
200
150
100
50
0
User-1
User-2
User-3
User

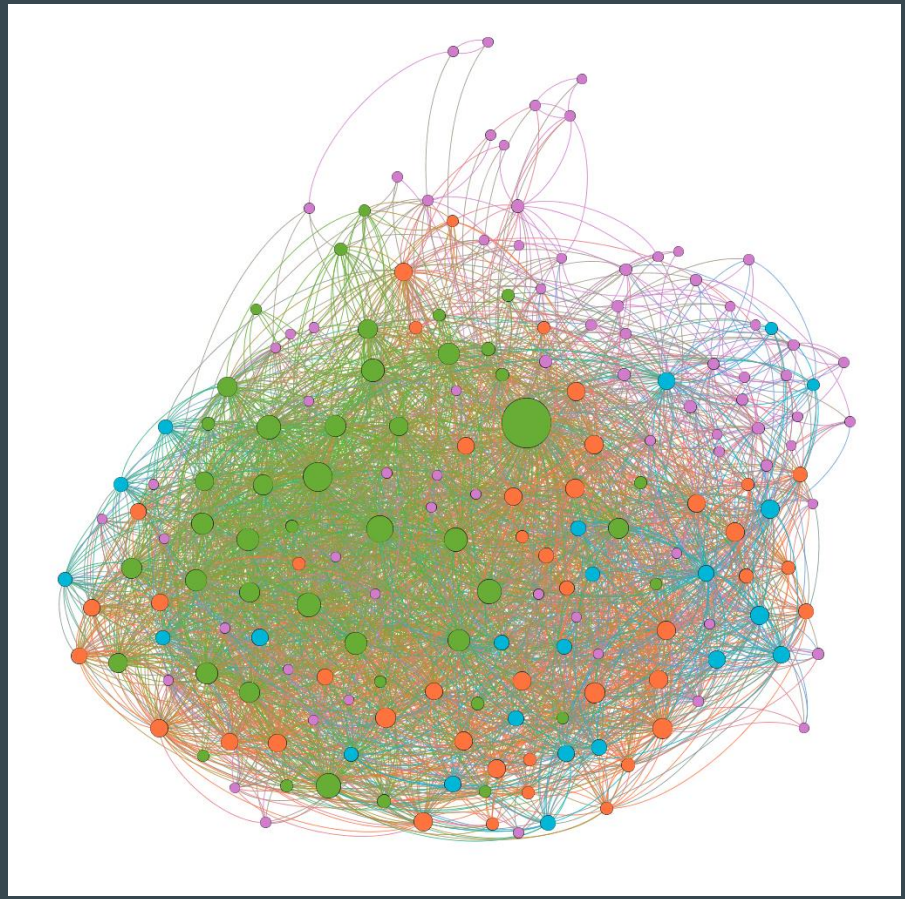
Sample Graph

- Colors indicate social circles obtained from ground truth
- Pink nodes do not belong to any circle.



Sample Graph

- Colors indicate social circles obtained from our clustering technique



References

- Java, X. Song, T. Finin and B. Tseng, "Why we twitter", Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis - WebKDD/SNA-KDD '07, 2007.
- Julian Mcauley and Jure Leskovec. 2014. Discovering social circles in ego networks. ACM Trans. Knowl. Discov. Data 8, 1, Article 4 (February 2014), 28 pages. DOI=http://dx.doi.org/10.1145/2556612
- "Soon to Launch: Lists | Twitter Blogs", Blog.twitter.com, 2009. [Online]. Available: https://blog.twitter.com/2009/soon-to-launch-lists. [Accessed: 28- Mar- 2016].


Thanks for listening