Temporal properties of social networks

Team 18

Rik Schreurs, Francois van der Ven, Loek Tonnaer

{h.c.m.scheurs, f.v.d.ven, l.m.a.tonnaer}@student.tue.nl

Data

- Edge data from three social networks, connections between users

	Facebook	YouTube	Flickr
# edges	817.037	9.375.374	33.140.017
# vertices	63.731	3.223.589	2.302.925
Size	15,5 MB	257,1 MB	901,5 MB

Programming setup

- *Apache Flink*
- Graph processing API and library: *Gelly*

Objectives

Temporal aspects

- How long does it take for a new user to achieve a certain **number of connections**?
- **Degrees of separation**, in how many steps can you go from one user to any other? How does this evolve over time?
- How **"connected"** is each of the graphs? How many connected components are there, and how does this change over time?

Results

figures/edges.png

figures/std-dev.png

figures/FacebookConnectedComponents.png

figures/YouTubeConnectedComponents.png

figures/FlickrConnectedComponents.png

Results

figures/FacebookDegreesOfSeperation.png

Conclusions

- Conclusion 1
- Conclusion 2
- The number of connected components for the Facebook dataset decreases over time. A possible explanation for this is that groups of friends get bigger. YouTube subscribers follow a similar pattern. As channels get bigger, they become central **hubs** and they connect people to each other. The Flickr dataset does not show interesting characteristics regarding connected components.