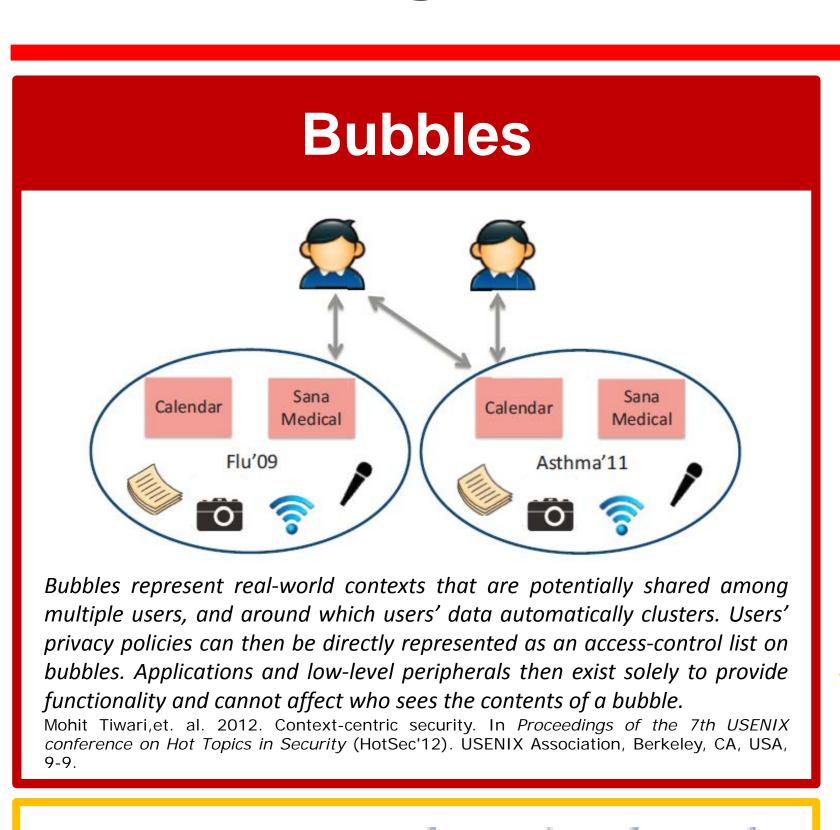
Cloud Storage Service Organization Study Automatic creation of Bubbles from user data

Casen Hunger², Yehuda Katz¹, Michelle Mazurek¹, Ch. Papamanthou¹, Mohit Tiwari²

¹ University of Maryland / ² University of Texas at Austin



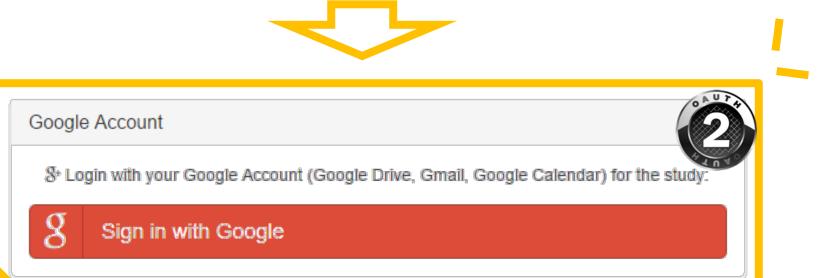


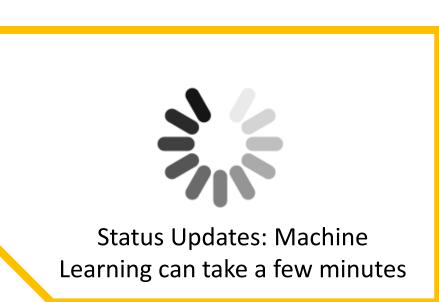


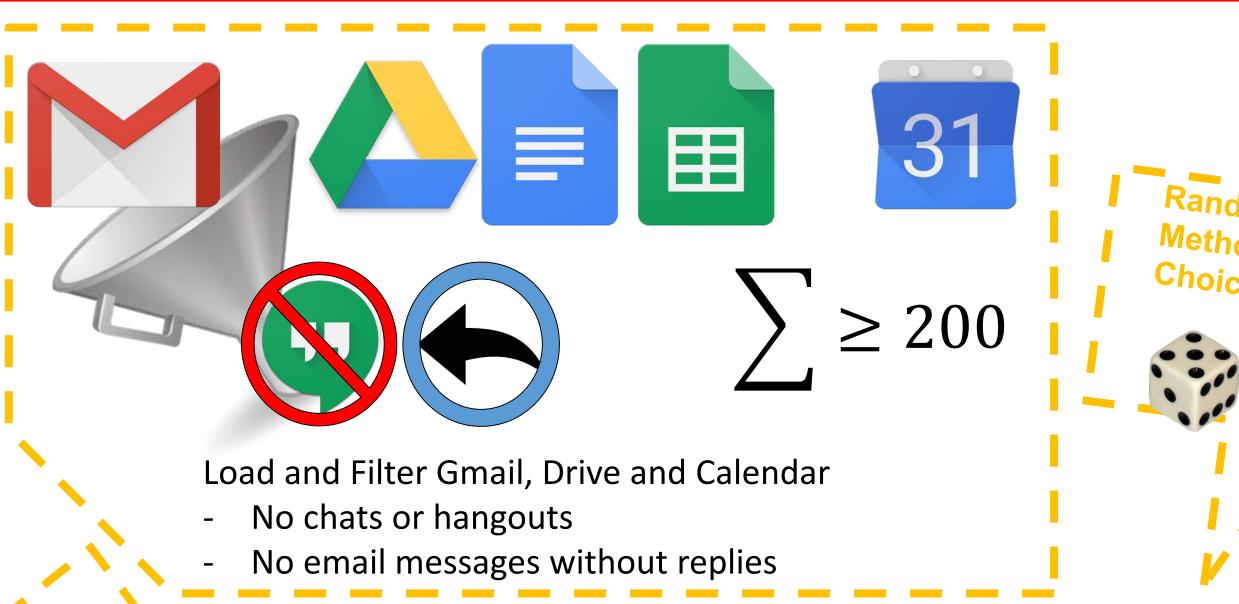
- Google Account at least •
- 3 months old
- At least 600 emails and

Demographic Info

20 files in Google Drive Fluent in English Live in USA







Clustering Algorithm #1 Email/File Subject Frequency Analysis

Clustering Algorithm #2 Bi-Gram Frequency KMeans Clustering of Email Subject

Clustering Algorithm #3 Bi-Gram Frequency KMeans Clustering of Email Body

> Clustering Algorithm #4 TF-IDF of Email Body

Bubble Rating

- 1. Choose one or more categories that this Bubble fits in
- 2. Choose how each person is related to you
- 2. For the whole group, rate how much you agree:
 - Contents fit together well
 - All on the same topic
 - All shared with the same people
 - Important content is missing from this group
 - Is the missing content older that the shown content? (We don't pull full mailbox or drive, just the minimum to make some number-N Bubbles in reverse chronological order)
 - If I were manually creating groups, this is a group I would create

Example Relations

- Me
- 2. Work or School Colleague
- Someone I Supervise
- A Boss, Teacher, or Supervisor
- . A Close Friend
- A Romantic Partner
- A Family Member I Am Close With
- A Family Member I Am Not Close With
- 9. A Colleague in a Non-Work Activity
- 10. Someone I Don\'t Know in Real Life
- 11. A mailing list or not a specific person

Generated Bubble #1

Generated Bubble #2

Generated Bubble #3

Generated Bubble #4

Generated Bubble #5

Randomly Generated Control Bubble

Rate each Bubble on Accuracy (see inset box)

Example categories

- Archive or long-term storage of content that I expect to look at months Publicizing an event or promoting myself, such as an ad for a musical or years later
- Event planning, such as coordinating a group meeting or trip
- Household management, such as handling budget or chore information within my household
- Collaboration on a work, school, or other community group project
- Documenting a life event, such as with photos or videos
- Connecting, staying in touch with, or having a conversation with friends or family
- Advertising, Marketing, Coupons or other Commercial Email

