Integrating Real-time GIS and Social Media for Qualitative Transportation Data Collection

PI: Dr. Hongmian Gong
Co-PI: Dr. Carsten Kessler
Mario Giampieri
Why social media data?

• Transportation planning has traditionally focused on collecting quantitative data for modelling, focusing on how much people travel.
• Qualitative data allows us to model why people use different modes for different purposes.
• This study seeks to determine how real-time transportation information affects daily travels.
Research Participants

• Participants have installed a GPS tracking app that report their locations

• Users allow us to inspect data they publish on Twitter, and have that data analyzed using machine learning algorithms
Server infrastructure setup

- Amazon Virtual Machine Instance running Windows Server 2008
- ESRI ArcServer 10.2
- Esri Geoevent Processor for Server 10.2
- mySQL database
ESRI Geoevent Processor
Ingest Twitter data
Twitter developer API

Application Management

utrc

Details Settings Keys and Access Tokens Permissions

UTRC response app
http://www.utrcweettracker.com

Organization
Information about the organization or company associated with your application. This information is optional.

Organizer: None
Organization website: None

Application Settings
Your application’s Consumer Key and Secret are used to authenticate requests to the Twitter Platform.

Access level: Read and write (modify app permissions)
Consumer Key (API Key): 641TSeW/hyQpcG1fVnw7.WebServlets (manage keys and access tokens)
Callback URL: None
Callback URL Locked: No
Sign in with Twitter: No
App-only authentication: https://api.twitter.com/oauth/authorize
Request token URL: https://api.twitter.com/oauth/request_token
Authorize URL: https://api.twitter.com/oauth/authorize
Access token URL: https://api.twitter.com/oauth/access_token
Twitter developer API
GeoEvent processor Twitter connector
GeoEvent processor Twitter connector
GeoEvent processor
Add filter
GeoEvent processor
Add processor
Map fields to remove extraneous data
GeoEvent processor
Print tweet data to console
Print tweet data to console
GeoEvent processor
Stream Tweet data to web map
Stream Tweet data to web map
GeoEvent processor
Write Tweet data to JSON
Write Tweet data to JSON
Classify tweets

C:\dir > java –jar TweetClassifierFile.jar tweets.json tweets.classifier
Classify tweets

C:\dir > java –jar TweetClassiferFile.jar tweets.json tweets.classifier
"classifier":{"none":0.011179990893958146,"shopping":
0.3480679872338228,"workschool":0.08498662517955215,"social":
0.01098021382705481,"eat":0.01570851139120672,"workother":
0.0021255868285692445,"dropoff":
0.14357428557518584,"recreation":
0.23121059429062144,"otherfam":0.0823702210443366}
import os

out_format = '{\"\"}.format

with open('2013-10-13-09-11.json', "r") as file:
    outfile = open('cleaned.json', "w")
    for line in file:
        outfile.write(line)
    outfile.close()

clease
Next steps

• Send classified tweet information to mySQL database and associate with user

• Build public facing web map application that will allow users to modify ML classifications if necessary
Questions?

Thanks!

mariogiampieri@gmail.com