How Can Geospatial "Big Data" Help Disaster Response and Track Disease Outbreaks?

Transform Innovative Geospatial Technology to Solve Real World Problems.

Dr. Ming-Hsiang Tsou

mtsou@mail.sdsu.edu, Twitter **@mingtsou**

Director of **the Center for Human Dynamics in the Mobile Age**
Professor, Department of Geography , San Diego State University
Assistant Director, National GeoTech Center

August 19, 2015

GeoTECH GEOTECHCENTER.ORG HDMA @SDSU THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

HDMA @SDSU THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — San Diego State University

What is Big Data?

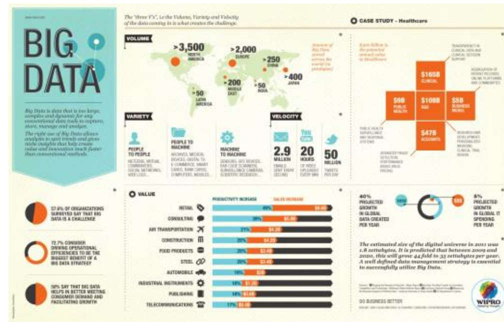


Image source: http://visual.ly/big-data (definition from IBM, and WIPRO)

Is this a good definition of "Big Data"?

HDMA @SDSU THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — San Diego State University

Big Data is Human-Centered Data

*Big Data is a large **dynamic** dataset created by or derived from **human activities, communications, movements, and behaviors**. (Tsou, 2015).*

The term, Big Data, refers to **big ideas, big impacts, and big changes** for our society in addition to a **big volume** of datasets.

Tsou, M.H. (2015, In Press), Research Challenges and Opportunities in Mapping Social Media and Big Data, *Cartography and Geographic Information Science.*

The Challenge of Big Data Analytics:

Big Data are very Messy, Noisy, and Unstructured!

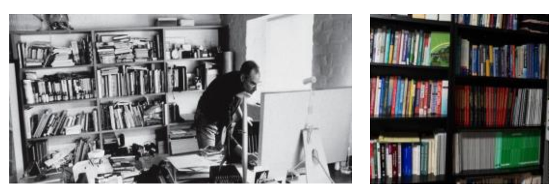


Image Source: http://www.contentverse.com/office-pains/10-messy-desks-successful-people/

Require collaboration efforts from linguistics, **geographers (GIS experts),** computer scientists, data mining experts, statisticians, physicists, modelers, and domain experts.

Human Dynamic in the Mobile Age (HDMA) San Diego State University

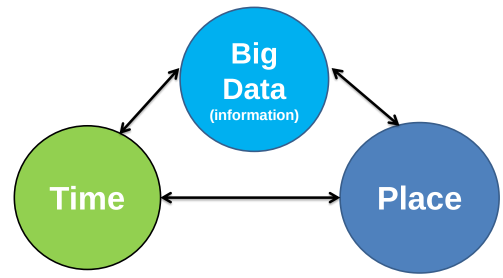
Question #1:

Are you a Geographer? (YES/NO)

Are You familiar with Geospatial Technology? (YES/NO)

Geography (place and time) is the KEY for Understanding and Integrating Big Data

(Tsou and Lietner, 2013)



Big Data Category (Tsou, 2015).

Social life data: social media services (Twitter, Flickr, Snapchat, YouTube, Foursquare, etc.), online forums, online video games, and web blogs.

Health data: electronic medical records (EMR) from hospitals and health centers, cancer registry data, disease outbreak tracking and epidemiology data.

Business and commercial data: credit card transactions, online business reviews (such as Yelp and Amazon reviews), supermarket membership records, shopping mall transaction records, credit card fraud examination data, enterprise management data, and marketing analysis data.

Transportation and human traffic data: GPS tracks (from taxi, buses, Uber, bike sharing programs, and mobile phones), traffic censor data (from subways, trolleys, buses, bike lanes, highways), and mobile phone data (from data transmission records and cellular network data).

Scientific research data include earthquakes sensors, weather sensors, satellite images, crowd sourcing data for biodiversity research, volunteered geographic information, and census data.

Geography (place and time) is the KEY for understanding Big Data!

Data Integration / Data Fusion

Explore their **spatiotemporal relationships** in both network space and geographical space.

Health or Disaster Data Layer

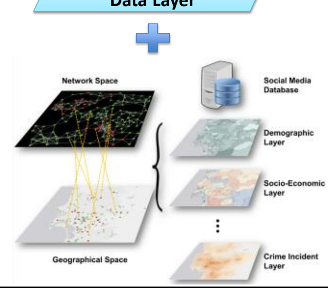


Image provided by Dr. Atshushi Nara (Associate Director of HDMA Center).

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — SAN DIEGO STATE UNIVERSITY

Research Showcase #1:

Geo-Targeted Social Media (Twitter) Analytics for Tracking Flu Outbreaks in U.S.

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — SAN DIEGO STATE UNIVERSITY

Question #2:

Do you have a Twitter Account? (YES/NO)

Do you send out tweets regularly? (more than once per week). (YES/NO).

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — SAN DIEGO STATE UNIVERSITY

Why Choose Twitter?

80% academic researchers are using Twitter APIs to get their social media data.

1. **Free and Open** Access Data from **APIs** (you can write a program in your desktop to download Twitter data (**tweet**s) automatically). But the free APIs has the 1% data limit.
2. **Large** User Base (+500 million users) and very popular in U.S., Europe, and Japan. But not in China, Taiwan, and Korea (China has a similar platform called "**Weibo**").
3. **Easy to program** in Python or PHP (Tweepy, TwitterSearch, etc.). Many available API libraries to use now.
4. **Historical data** and 100% data can be purchased from Twitter (but very expensive).
5. Rich **[Metadata]** tags in each tweet (time stamp, user, follower, platform, time zone, text, URL, Retweet, language, devices).

Other possible social media APIs: **Flickr, Instagram**, Foursquare, Yelp, YouTube. Why not **Facebook**? (Facebook Graph APIs are **VERY LIMITED and PROTECTIVE**. **No Public data feed**). You need to have "internal connections" to Facebook staff to conduct research.

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — SAN DIEGO STATE UNIVERSITY

What we can get from Twitter data? Where to find geospatial information?

CREATED_AT_LOCAL	TEXT	FROM_USER	FROM_USER_NAME	LOCATION
6/23/2015 23:55	Join us tonight for Free Rapid HIV Testing! 5:30pm to 8pm! All are welc	Rai		Concord, CA, USA
6/23/2015 23:54	June 25, 26, & 27 Free HIV Testing at select @Walgreens. Rapid HI	Fem		Los Angeles, CA
6/23/2015 22:50	It's #NationalHIVTestingWeek! Do you know where to get a free #HIV t	BTG	n	San Diego, CA
6/23/2015 22:05	RT @CDCgov: Join our Natâ€™l #HIV Testing Day Thunderclap support	biwi	ns	West Orange, TX
6/23/2015 21:59	RT @DrFriedenCDC: Knowing your #HIV status is empowering. Take the	biwi	ns	West Orange, TX

Example: Use Twitter Search API to search for keyword "**HIV test**" or "**HIV testing**"
Only **1%** - 7% of Tweets have **X, Y GEO-coordinates** (from GPS or Geo-tagged).
But **50%** - 70% Tweets have **city-level locations** provided by their **user profile**.
Time Zone (spatial meaning)

URLS	HASHTAGS	IS_RETWEET	FOLLOWERS	FRIENDS_	STATUSES	TIME_ZONE	GEO	SOURCE	LANGUAG
NULL	NULL	FALSE	239	218	1768	NULL	-122.03107, 37.977	Facebook	en
NULL	NHTD	FALSE	291	726	1267	NULL	-118.24368, 34.052	Twitter for iPad	en
http://fb.me/	NationalHIVTestir	FALSE	32	137	175	NULL	-117.15726000000	Facebook	en
http://bit.ly/1	HIV	TRUE	240	477	2141	NULL	-93.75822, 30.0821	Twitter for Android	en
http://ow.ly/	HIV, NHTD	TRUE	240	477	2141	NULL	-93.75822, 30.0821	Twitter for Android	en
NULL	NULL	FALSE	50	169	1971	NULL	-117.90728, 33.703	Twitter for iPhone	en
NULL	NHTD, NPINChat	TRUE	2229	201	120858	Pacific Time (	-122.23635, 37.485	Twitter Web Client	en
http://ow.ly/	GetTested	FALSE	13	21	31	NULL	-97.74306, 30.2671	Hootsuite	en
http://1.usa.g	ONE	FALSE	83	202	1684	Brasilia	-101.48765, 33.299	twitterfeed	en

Data Filtering, Mining, and Visualization

HDMA @SDSU THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — San Diego State University

Application Programming Interfaces (API)

Geo-Targeting Data Collection (Twitter **APIs**)

Filter → Machine Learning → Trend Analysis → Spatial Analysis

Analysis

SMART Dashboard

Visualization

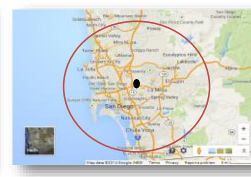
How to Collect Geo-targeted Social Media Data?

Source: Twitter **Search APIs** Spatial Search Function to collect public tweets (Geo-targeted) within a city.

Keywords: "flu" and "influenza".

Region: 17 or 20 miles radius from the center of 31 U.S. Cities.

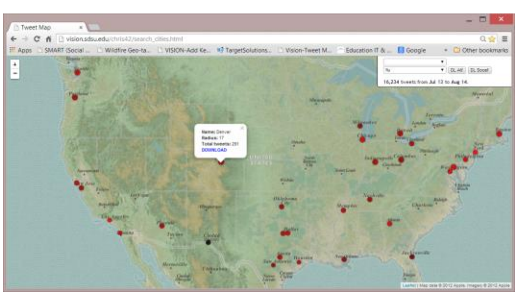
Time: **September 30, 2013** (Week 40) – **March 23, 2014** (Week 12)



Twitter Search API: based on the user profile (locations) and gazetteers (San Diego: include La Jolla, La Mesa, Chula Vista) (FREE for Search APIs)

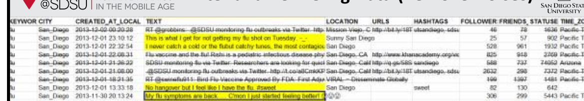
Collect Tweets from Top 31 U.S. Cities (17 miles radius)

31 different cities across the United States (chosen based on their population sizes): *Atlanta, Austin, Baltimore, Boston, Chicago, Cleveland, Columbus, Dallas, Denver, Detroit, El Paso, Fort Worth, Houston, Indianapolis, Jacksonville, Los Angeles, Memphis, Milwaukee, Nashville-Davidson, New Orleans, New York, Oklahoma City, Philadelphia, Phoenix, Portland, San Antonio, San Diego, San Francisco, San Jose, Seattle, and Washington, D.C.*



Filter and Refine Big Data (Remove Noises)

HDMA @SDSU THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — San Diego State University



STEP-1: Collect tweets that contain the word "flu" in their text from **31 U.S. Cities** with 17 miles radius.

STEP-2: Remove all retweets (RT) and tweets with a URL address (Web Pages).

STEP-3: Remove all tweets that contain the word "stomach" in the text. Since "stomach flu" is not relevant to influenza.

STEP-4: Remove users that have the character string "flu" in their **username**, as well duplicate users from within the same week.

STEP-5: **Machine Learning Procedure (scikit-learn)** for identifying "flu" relevant texts. Remove tweets that are classified as "invalid" by a support vector classifier that has been trained with a manually tagged training set

STEP-6: Calculate tweeting rates by dividing the number of filtered tweets in each city by the population of the city (within the 17 miles radius) from which they were collected.

Number of tweets
10,678
5,398
4,947
4,944
3279

Machine Learning

Total Flu tweets collected: 307,070.

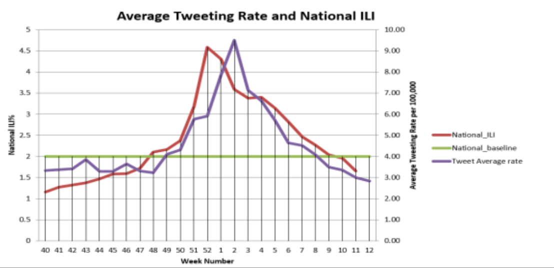
Final valid flu tweets: 88,979.

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

Real-Time Monitoring of Flu Outbreaks in U.S.
(National Scale – combined 31 Cities), 2013 – 2014 flu season

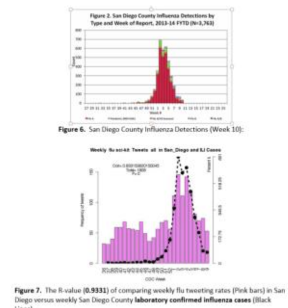
RED Line: National ILI
Purple Line: Weekly Tweeting Rate

(R) value = **0.8494**



HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

Trend Analysis at the Municipal Scale (San Diego)
with the Lab-tested confirmed flu cases



San Diego: Lab confirmed Flu Cases vs Tweeting Rate:
(R) value = **0.9331**

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

Next Step:

Monitor Flu Outbreaks in Real-Time?

The HDMA Center developed the SMART Dashboard:

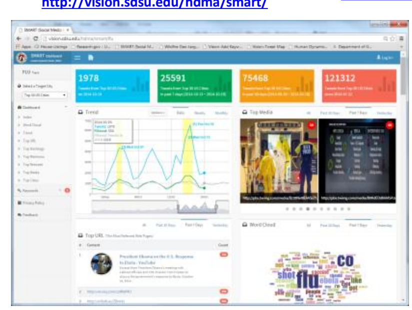
SMART Dashboard: Social Media Analytic and Research Testbed
http://vision.sdsu.edu/hdma/smart/ (Beta)

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

SMART Dashboard

Social Media Analytic and Research Testbed
http://vision.sdsu.edu/hdma/smart/

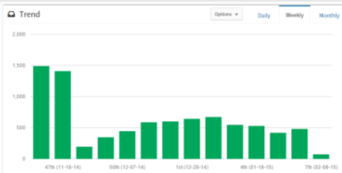
YouTube Video for 3 Mins



Real-time social media analytics (Trend Analysis, Word Clouds, Top URL, web pages, Top Hashtags/Mentions/Stories).

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

Tracking Flu Outbreaks in 2014/2015 Flu Season



# of Filtered ILI Tweets, Top 30 US Cities, as of February 9, 2015 (from SMART dashboard)

Only 1% -4% tweets has Geo-tagged coordinates.

Problems!!! Twitter broke its Search APIs on 11/20/2014 and only returned Geo-tagged tweets only. (Reduce 90% -95% of tweets collected)

CDC Influenza Positive Tests, *National Data Summary*, through Weeks 40-3, 2014-2015 Season

2014-2015 Comparison between ILI and Geotagged-only Tweets among 30 U.S. Cities

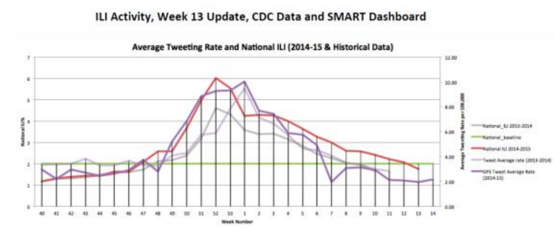


Figure 1. The comparison between National ILI Rate and the 31 Cities Tweeting Rate, with prediction up to Week 14. Red: National ILI, Purple: *GPS Only Tweets* Tweeting Rate, multiplied by 10 for 2014-2015.

**NOTE: Week 7 GPS Tweet Average (2014-15) is missing tweets from 2/15-2/19 due to internal server error, the week 7 Tweeting rage is significantly lower and incomplete. Week 8 is back to the normal collection process.*

Human Dynamic in the Mobile Age (HDMA) San Diego State University

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

Question #3:

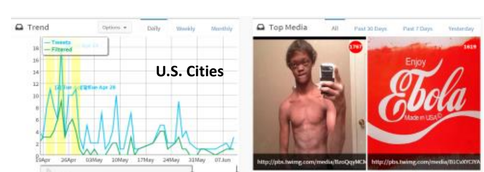
Does any of your friends get flu last year? (YES/NO).

Does any of your friends get Ebola last year? (YES/NO).

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

Ebola http://vision.sdsu.edu/hdma/smart/ebola

Different regions will have different "Voices" and "Opinions"

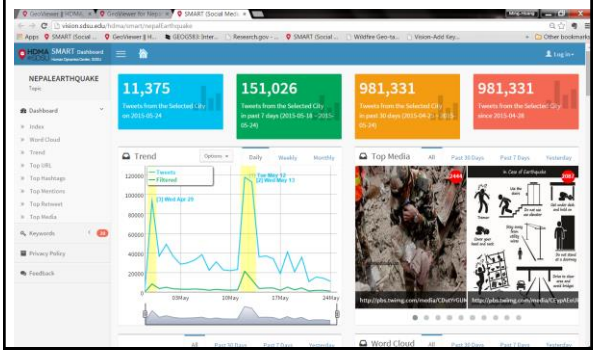


HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — SAN DIEGO STATE UNIVERSITY

Nepal Earthquake

Monitoring social media voices and discussion about this disaster.

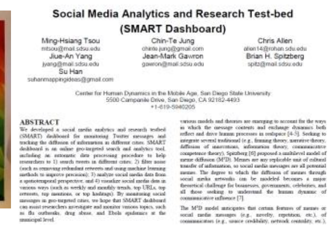
http://humandynamics.sdsu.edu/NepalEarthquake.html



HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — SAN DIEGO STATE UNIVERSITY

SMART Dashboard Won the BEST METHOD PAPER Award

in the 2015 International Conference of Social Media and Society, Toronto, Canada. http://dl.acm.org/citation.cfm?doid=2789187.2789196



Social Media Analytics and Research Test-bed (SMART Dashboard)

Ming-Hsiang Tsou, Chin-Te Jung, Chris Allen, Jiue-An Yang, Jean-Mark Gawron, Brian H. Spitzberg, Su Han

ABSTRACT

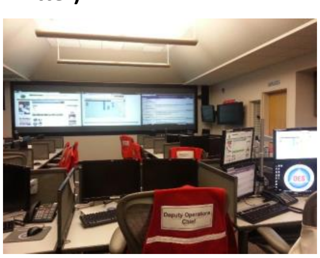
SMART Dashboard Examples and Links:

http://humandynamics.sdsu.edu/SMART.html

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — SAN DIEGO STATE UNIVERSITY

Research Showcase #2:

Real-time Situation Awareness Viewer for Monitoring Disaster Impacts Using Location-Based Social Media Messages (Twitter).



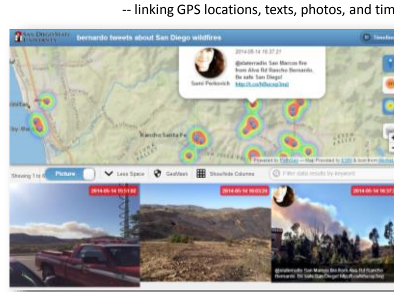
San Diego County: Office of Emergency Services (OES)

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — SAN DIEGO STATE UNIVERSITY

Geo-targeted Event Observation (GEO) Viewer 1.0

http://vision.sdsu.edu/hdma/wildfire/

Provide dynamic map display about event **ground truth observation**
-- linking GPS locations, texts, photos, and time.



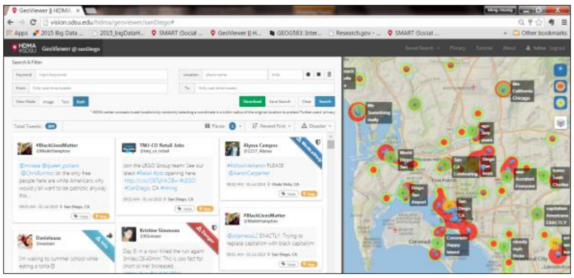
(GPS Tweets)

Monitor Disaster impacts, Recovery Activities and Victims' Needs

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

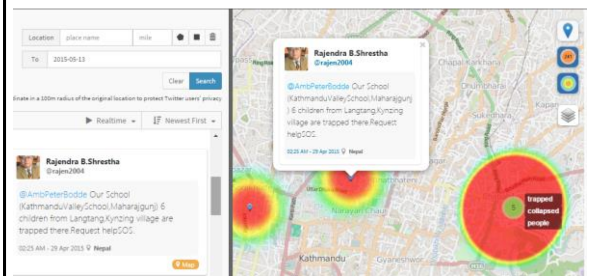
GeoViewer Tool v.2.2

(Video demo) EC2: http://vision.sdsu.edu/ec2/geoviewer/sanDiego (Live)



How to find out critical information from thousands of GPS-tagged tweets or hundreds of thousands of Non-GPS-tagged tweets?

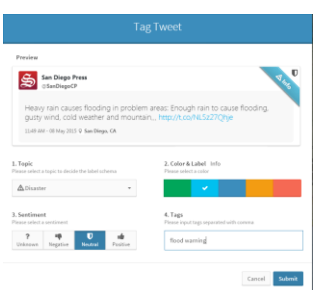
Nepal Earthquake Example: (keyword search: "trap")



One Possible Solution: Manual labeling (first 1000 tweets by volunteers) + Machine Learning Classification (built-in).

Human Dynamic in the Mobile Age (HDMA)

Digital Volunteers may help us identify and select important Tweets (for machine learning) during and after the disaster events.



Need Some programming and design help from OES, RedCross, and 211:

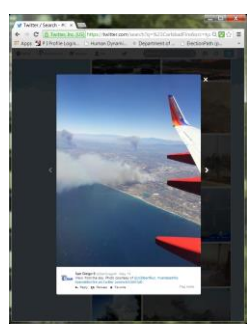
1. How to combine multiple volunteers' Inputs and Integration Systems (ranking system).
2. Which category and color schemes/labels should we use for each types of disasters (flooding, wildfires, earthquake, hurricanes).
3. Which tags might be useful?
4. Who are the target users? What kinds of "Output" system should we create? (for OES staff? For RedCross staff?)
5. Other suggestions?

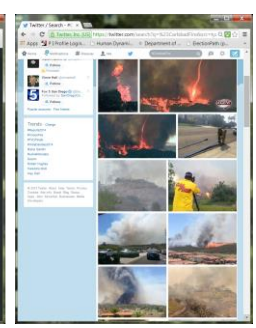
Human Dynamic in the Mobile Age (HDMA)

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

San Diego Wildfire 2014, May 13, 14, Case Study

(Real time + People's Need + Public Opinion + Communications)





HDMA @SDSU THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

San Diego State University

Social Media (Twitter) Data Collection from May 10 to May 19, 2014

- Total tweets collected

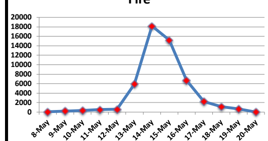
Keyword / Topic	Total Tweets	Tweets with GPS
Fire *(many unrelated tweets)*	**51,495** (not a good keyword)	**4 % with GPS coordinates**
wildfire	**2,885**	**3 % with GPS**
evacuation	**6,477**	**3 % with GPS**
san marcos	**11,781**	**1 % with GPS**
bernardo	**2,871**	**4 % with GPS**
carlsbad	**10,513**	**2 % with GPS**

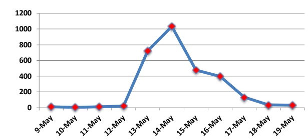
HDMA @SDSU THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

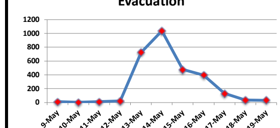
Temporal Analysis (Wildfire Tweets)

San Diego State University

Tweet Intensity (tweets per day) by Topics







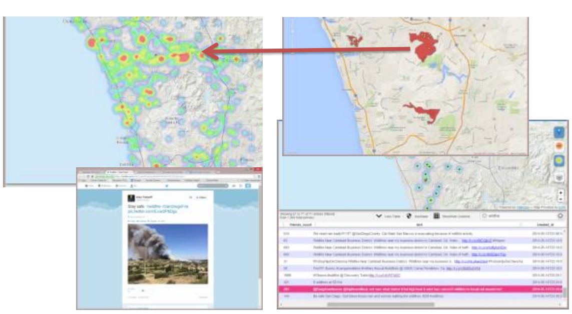
Most relevant tweets started on May 13, 2014. The peak date is May 14, 2014. Then the number of tweets decreased after May 15, 2014.

HDMA @SDSU THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

Spatial Clustering (Wildfire Tweets)

San Diego State University

Spatial clusters (hotspots) of tweets are nearby the actual locations of wildfires (events)



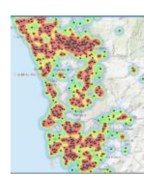
HDMA @SDSU THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

Dynamic Spatiotemporal Patterns

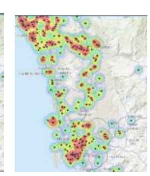
San Diego State University

Temporal Change of Geo-tagged Tweets Related to “fire”, “wildfire”, and “evacuation”.

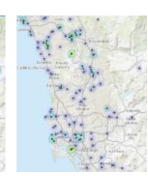
5/14/2014



5/15/2014



5/16/2014



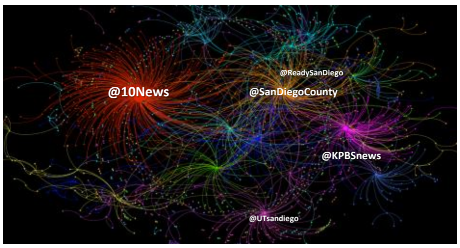
5/17/2014

HDMA @SDSU THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

SAN DIEGO STATE UNIVERSITY

2014 San Diego Wildfire Tweets

Social Network Analysis (SNA)



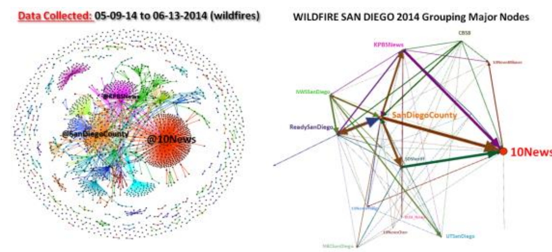
- Identify the **network influence** for each individuals (who are the opinion leaders?)
- Predicting the **Spreads** (Speed, Scale, and Range) of Social Media Messages in Different Social Networks. (**following, retweets, and mentions** relationships)

HDMA @SDSU THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

SAN DIEGO STATE UNIVERSITY

Social Network Analysis in "ReTweet" Messages

San Diego 2014 Wildfire SNA
(in-degrees networks -- who is **retweeting** who?)

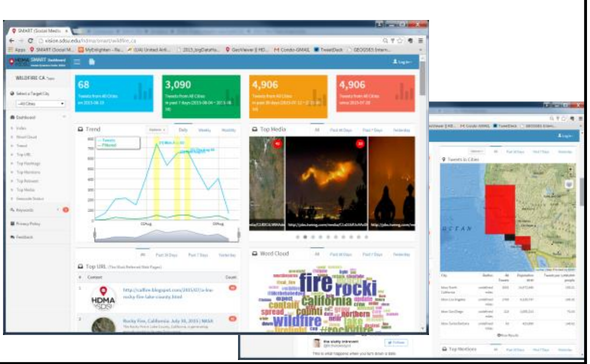


HDMA @SDSU THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

SAN DIEGO STATE UNIVERSITY

SMART Dashboard for Tracking CA Wildfires:

http://vision.sdsu.edu/hdma/smart/wildfire_ca



HDMA @SDSU THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

SAN DIEGO STATE UNIVERSITY

The Limitations and Challenges of Social Media For Public Health Research

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — San Diego State University

Question #4:

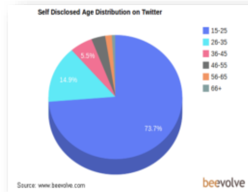
Have you ever used Facebook/Foursquare/Instagram/Twitter to "Check-in" places and restaurants? (YES/NO)

Do your friends "tag" you in their photos or messages? (YES/NO).

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — San Diego State University

Social Media User Profiles

Social Media messages can NOT represent all population, but it can provide **warning signals** and **real-time updates**.



2012 Survey

Twitter Users are

- **Young** (75% are between 15 – 25 years old).
- More **Urban** residents than rural
- Higher adoption% in African Americans
- Many Journalists and **Mass Media** staff.
- 20% are not real "human beings" (**robots**): many advertisement and marketing activities.

Using Different Keywords can get different demographic groups:

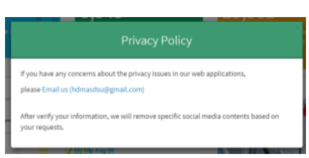
- **#Healthcare: include more senior people (Very few teenagers will tweet about "healthcare"). (We need more background study).**
- **"Keywords" as a sampling tool for social media users.**

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — San Diego State University

User Privacy Issue

* HDMA center conceals tweet locations by **randomly selecting a coordinate in a 100m radius of the original location** to protect Twitter users' privacy.

- Concerns about "**Big Brother**".
- Although all the tweets collected from APIs are "**public tweets**" (everyone can search them and retrieve them).
- Some content of tweets may contain **personal private information** (real names, locations of homes, offices, private conversations, medical situations, etc.)



HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — San Diego State University

Additional Information and Learning Resources

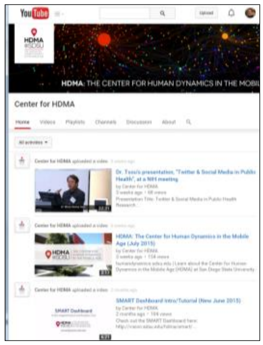
The Center for Human Dynamics in the Mobile Age
http://humandynamics.sdsu.edu/educationSource.html

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

Github and YouTube Channels

San Diego State University

https://github.com/HDMA-SDSU/HDMA-SocialMediaAPI
https://www.youtube.com/channel/UCKiEJPhr3vRgNgFi5aD49VA



HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

HDMA Center Activities (Fall 2015)

San Diego State University

Annual Big Data Science Symposium (Oct 2): project strong academic influences at national and international levels.

Annual Big Data Hackathon for San Diego (Oct 3, 4, and 10): examine important civic issues with actionable results in San Diego: **Water Drought/Conservation, Disaster Response/Assistance, and Crime Monitor/Prevention**.

Build a collaborative community for the future development of the **Open Government** and **Open Data Initiative** in San Diego (City of San Diego, County of San Diego, SANDAG, SanGIS).

The City of San Diego

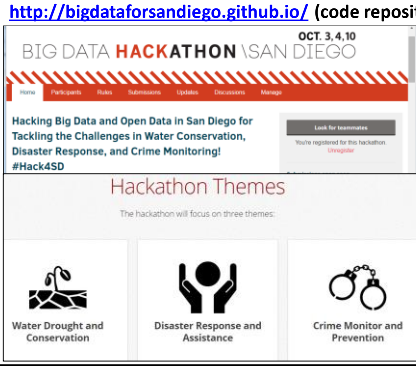
HDMA @SDSU

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

SDSU Big Data Hackathon Event (Oct 3, 4, 10)

San Diego State University

http://bigdataforsandiego.challengepost.com/ (Registration)
http://bigdataforsandiego.github.io/ (code repository)



HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

http://humandynamics.sdsu.edu/

Thank You Q & A

Director: Dr. Ming-Hsiang Tsou
mtsou@mail.sdsu.edu

Funded by
- NSF Cyber-Enabled Discovery and Innovation (CDI) program. Award # 1028177. (2010-2015) http://mappingideas.sdsu.edu/
- NSF Interdisciplinary Behavioral and Social Science (IBSS) Program, Award #1416509 (2014-2018): "Spatiotemporal Modeling of Human Dynamics Across Social Media and Social Networks". http://socialmedia.sdsu.edu/

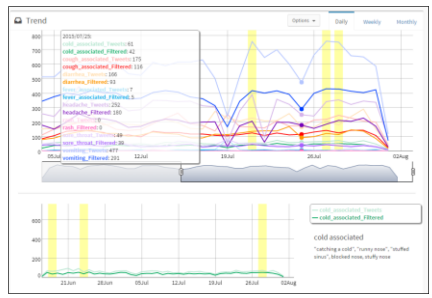
Human Dynamic in the Mobile Age (HDMA) San Diego State University

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — San Diego State University

Backup Slides

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — San Diego State University

Next Step: Syndromic Surveillance (Underdevelopment)
(tracking multiple Symptoms: fever, cold, cough, vomiting, etc.)
http://vision.sdsu.edu/hdma/smart/syndromic



Designed for **Early Detection** of "unknown" disease outbreaks, such as **Swine Flu** and **SARS**

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — San Diego State University

However, Twitter Search API is not working since Nov, 2014.

The HDMA Center has built our own Internal GeoCoder Engine for User Location Profile: using **GeoNames.org** gazetteers (Creative Commons Data).+ User defined rules.

Enable Flexible or Self-defined Geo-Target Boundaries (California, Santa Barbara, Los Angeles, San Diego – bounding boxes, or State boundaries)

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE — San Diego State University

Develop manual coding methods and automatic machine learning process for both SMART dashboard and GeoViewer. (Underdevelopment).

From the Wildfires CA collection (the Whole World):
Select 500 Original Tweets (Remove RT and URL).

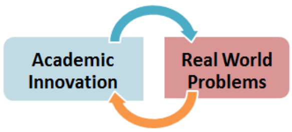
Conduct a manual Coding for (Based on Red Cross Suggestions)
DS – seeking information
DS – Need for help/service (Urgent)
DS – Need for help/service (Non-Urgent)
DS – Offer of help/service
DS – Situation reports
DS – In kind donations
DS – Emotional support
Not-DS relevant

Then Run Machine Learning to test the auto-classification results (For Wildfire Disaster).

HDMA @SDSU | THE CENTER FOR HUMAN DYNAMICS IN THE MOBILE AGE

HDMA Center

SAN DIEGO STATE UNIVERSITY



Advancing Interdisciplinary Research on BIG DATA, Human Dynamics, and the Social Web

http://humandynamics.sdsu.edu/

Five Core Faculty

Dr. Ming-Hsiang Tsou (Director)

(Director of HDMA Center, Professor of Geography) 2012 SDSU President's Leadership Fund (PLF) Awards for Excellence, expert in Web GIS, Mobile GIS, Cartography, Social Media, and Big Data

Dr. Sheldon Zhang (Board member)

(Past Chair of Sociology, Professor), expert in Human Trafficking, Transnational Organized Crime, Juvenile Delinquency, Probation and Parole, etc.

Dr. John Elder (Board member)

(Distinguished Professor of Public Health; former director and Senior Core Investigator, Institute for Behavioral and Community Health [IBACH]) 2012 SDSU President's Leadership Fund (PLF) Awards for Excellence, expert in randomized trials for chronic disease prevention in Latino communities.

Dr. Piotr Jankowski (Board member)

(Chair of Geography), expert in Spatial Decision Support Systems, Public Participation GIS, Visual Analytics, Geocomputation Methods, Volunteered Geographic Information, and Water Resource Modeling and Management.

Dr. Brian Spitzberg (Board member)

(Senate Distinguished Professor, School of Communication), expert in communication theories, communicative (in)competence, obsessive relational intrusion.