

School of Engineering Center for Cyber–Physical Systems and the Internet of Things

# Digital Technologies for COVID-19

A Virtual Mini-Workshop Friday 3rd April, 2020

Hosted by CCI, https://cci.usc.edu/



#### Introduction

#### Session 1: Data collection, mobile apps, IoT and social media analytics

- Peter Kuhn
- Cyrus Shahabi
- Bhaskar Krishnamachari
- Emilio Ferrara
- Kristina Lerman
- Yan Liu and Karishma Sharma

#### Session 2: Modeling, prediction, knowledge extraction

- Viktor Prasanna and Ajitesh Srivastava
- Aram Galstyan and Fred Morstatter
- Pedro Szekely
- Jay Pujara
- Jonathan May
- Paul Bogdan and Xiong Ye Xiao

#### Wrap up

### Session 1: Data collection, Mobile Apps, IoT and Social Media Analytics

### Peter Kuhn

# **Data Challenges in COVID-19**

### The infection: SARS CoV2 virus The disease: COVID-19



Convergent Science Institute in Cancer

### Testing: Why?

•What does broad surveillance look like and what does privacy preservation mean.

- •Amongst those infected:
  - 50% are asymptomatic (recent data from Iceland).
    40% have mild/moderate symptoms (home care)
    7% have severe symptoms (need clinical support)
    3% have critical symptoms (ICU).
    About half of the critical patients die.

### Testing How

- Two key challenges that need data support:
   Reduce spread by identifying emergence of an outbreak (hotspots)
  - Reduce progression from mild/moderate to critical, which is currently happening at about a rate of 10% (data from China).

### **Testing What**

- Testing has two aspects:
  - Antigen testing for identifying the Viral RNA (genome)
  - Antibody testing for identifying the immune response.
  - oBoth have advantages and disadvantages.

### Cyrus Shahabi

### An end-to-end system for data acquisition, storage, access, analysis and visualization of COVID-19 spread http://imsc.usc.edu/

Cyrus Shahabi, Ph.D. Professor of Computer Science, Electrical Engineering & Spatial Sciences Helen N. and Emmett H. Jones Professor of Engineering Chair, Department of Computer Science Director, Integrated Media Systems Center (IMSC) Viterbi School of Engineering University of Southern California Los Angeles, CA 900890781 shahabi@usc.edu

### **Research Background**

- *Contact tracing* is an effective method for prevention and control of epidemic disease outbreaks COVID-19
- An obvious scalable solution is to use people's mobile phones, but concerns:
- 1. Ineffective (Infrequent or coarse location tracking):
  - Use social relationship inferred from people's historical trajectories
    - H. Pham, C. Shahabi, and Y. Liu. EBM: an entropy-based model to infer social strength from spatiotemporal data. In Proceedings of the ACM SIGMOD International Conference on Management of Data, SIGMOD 2013, New York, NY, USA, June 22-27, 2013, pages 265–276. ACM, 2013.

#### 2. Violate Privacy:

- Searchable encryption w spatial indexes (slow but accurate)
  - NSF IIS-1910950: PE4GQ Practical Encryption for Geospatial Queries on Private Data
- Geo-indistinguishability (fast but noisy)
  - R. Ahuja, G. Ghinita, and C. Shahabi. A utility-preserving and scalable technique for protecting location data with geo-indistinguishability. In Advances in Database Technology - 22nd International Conference on Extending Database Technology, EDBT 2019, Lisbon, Portugal, March 26-29, 2019, pages 217–228. OpenProceedings.org, 2019.



### **Current Projects**

- COVID-19 Mobile Data Collection
  - w/ Peter Kuhn, Bhaskar Krishnamachari and M. Naveed, developed the iOS app
- COVID-19 Data Management and Analysis
  - NSF RAPID Proposal in collaboration with Emory University and UT-Health 
     I risk analysis
- Map Visualization of COVID-19 Spread
  - W/ Gabe Kahn at Annenberg Crosstown project, we created an online map visualization of COVID-19 spread, see: <u>https://coronavirus.xtown.la/</u>
- The combination of these 3 projects will enable us to create an end-to-end system for acquisition, storage, access, analysis and visualization of COVID-19 spatiotemporal data.







#### App Demo

Carrier 🗢 5:07 PM 💻			
To keep our community safe, take our symptom survey.			
0 active cases USC		53 active cases Los Angeles	
329 active cases California		2730 active cases United States	
Worldwide	)		
76151	(48%)	Active	
74461	(47%)	Recovered	
5821	(3%)	Deaths	
156433		Total	
Summary	Sur	Vey Contact	

### Bhaskar Krishnamachari

### Privacy-Sensitive Contact Tracing via Mobile App





For more details see <u>https://link.medium.com/oNay6EVI84</u> Paper in submission with Gowri S. Ramachandran, Arvin Hekmati

#### 2 Protocols:

- send encrypted or random anonymous beacons continually via Bluetooth
  - infected patients release logged data on a server after certification
- Other can check privately if they have been near an infected patient

### Multi-Community Data-driven Epidemic Prediction



- Using data released by LA County Public Health department to
- Estimating parameters for multi-community SEIR-type models that can be used to predict progress of COVID-19 (and therefore, hospital bed requirements).
- Exploring how to enhance these models with IoT, Call Data, Mobility and Traffic Data

Work in progress with Mehrdad Kiamari and Gowri S. Ramachandran Data and scripts online at <u>https://anrg.usc.edu/www/covid19.html</u>

### **Emilio Ferrara**

# COVID-19 & Social Media

Ongoing research @ Emilio Ferrara's Lab & many associates: Lerman @ISI, Cowan @Annenberg, De la Haye @Keck

### How to Enable Human-centric research?

#### COVID-19 Social Media Watch:

Multilingual, real-time data collection, shared with research community **Dataset**:

- Online manipulation of COVID-19 chatter Tracking forms of manipulation (bots, trolls, fake news, etc.) & influence
- **COVID-19 emotion dynamics: fear, anger, hate** How do fear and emotional swings contribute to the spread of panic? Can we classify the kinds of public responses via social media cues? Do emotional trends (hate, fear) change with outbreak major events?
- **2020 US Presidential Elections & COVID-19** How has COVID-19 changed sentiment in US political discourse? Can political engagement be predicted by online emotional contagion? Has online fear spread to election-related events (turnout)?



# Predicting events in the offline world during outbreaks

- **Predicting insurgency during epidemic outbreaks** Food/cash shortages, service disruption, outages: All can trigger panic, unrest, and insurgency. We are developing methods to predict violence from *social sensors*, *wearables*, *spatio-temporal* and *network* data.
- **Predicting food insecurity and food shortages** Using mobility data, purchasing behavior data, supply chain data, we can determine changing food needs/shortages for outbreaks.
- **Predictability limits of real-time outbreak surveillance** What is the intrinsic limit to predicting epidemic spread under limited testing, infection reporting, and inaccurate/sampled data?



### Kristina Lerman

# Monitoring Impacts of COVID-19

Research labs of Kristina Lerman and Emilio Ferrara

### Forecasting impacts of Covid19



- Current **hotspots** for Covid19 in LA
- Merge with census data to forecast where the next hotspots will be
- Also exploring the economic impacts:
  - How busy are businesses now?
  - Which businesses benefit, or suffer the most from COVID-19?
  - How many businesses close for good?
  - Inequalities of impact
- Other work
  - Elections and the pandemic
  - Pandemic misinformation

### Impact of Social Distancing Interventions

- Health professionals argue for *social distancing* such as avoiding large and small crowds
- How well can social distancing isolate vulnerable populations?
- Create a contact network from mobility @ data
- We find that contact network remains connected until almost all interaction are removed
- We are exploring more efficient methods that may be easier to enforce



### Yan Liu and Karishma Sharma

### COVID-19 - Social media and misinformation analysis

• Misinformation on social media can have dire consequences on public health

CALM DOWN! BREAKING EXCLUSIVE: The Coronavirus Fatality Rate Reported by Media Is COMPLETELY INACCURATE. Actual Rate Is LESS THAN THE FLU – MEDIA LYING AGAIN! Are we surprised? @RudyGiuliani has a #coronavirus podcast: https://t.co/nWV2zaaXm6 https://t.co/0rXKmMuOIY [2020-03-16T14:33:14]

RT @realTuckFrumper: Trump Is Withholding Coronavirus Testing Kits From Governors https://t.co/aszIIk1dOq [2020-03-08T23:28:43]





https://usc-melady.github.io/COVID-19-Tweet-Analysis/

https://arxiv.org/abs/2003.12309

### COVID-19 - Social media and misinformation analysis

<sup>D</sup> Misinformation on social media can have dire consequences on public health

CALM DOWN! BREAKING EXCLUSIVE: The Coronavirus Fatality Rate Reported by Media Is COMPLETELY INACCURATE. Actual Rate Is LESS THAN THE FLU – MEDIA LYING AGAIN! Are we surprised? @RudyGiuliani has a #coronavirus podcast: https://t.co/nWV2zaaXm6 https://t.co/0rXKmMuOIY [2020-03-16T14:33:14]

RT @realTuckFrumper: Trump Is Withholding Coronavirus Testing Kits From Governors https://t.co/aszIIk1dOq [2020-03-08T23:28:43]



### credible Source

Credible/Not

- Unreliable, Fake
- Clickbait
- Political, Biased



- Tweet features
- Propagation features

https://usc-melady.github.io/COVID-19-Tweet-Analysis/

https://arxiv.org/abs/2003.12309

### COVID-19 - Social media and misinformation analysis

#### Social media is pivotal for communication during global crisis



Top Countries (most positive on Twitter) 🥥

Saudi Arabia, Bangladesh, Uganda, Maldives, Tanzania, Ireland, Ghana, Nigeria, Pakistan, India

#### Top Countries (most negative on Twitter) 🔴

Hong Kong, Venezuela, Norway, United States of America, Sweden, Mexico, Netherlands, Greece, Spain, Portugal

+ Positive

Sentiment

- Negative

Sentiment

#### https://usc-melady.github.io/COVID-19-Tweet-Analysis/

#### https://arxiv.org/abs/2003.12309

# Session 2: Modeling, Prediction and Knowledge Graphs

### Ajitesh Srivastava and Viktor Prasanna

# **ReCOVER:** Accurate Fine-grained Predictions and Resource Allocation for COVID-19 Epidemics Response

Ajitesh Srivastava, Viktor K Prasanna



#### Motivation

Preparedness and resource management needs state/county/city level predictions:

- How many masks, testing kits, beds are needed tomorrow/next week at a given hospital
- How to distribute state/country resources across all the hospitals in a state/country
- **Solution:** Greedy marginal value maximization is optimal under certain assumptions on the resource values. Needs accurate predictions

CHIKV epidemic: Country-level predictions. Weekly over 8 months, 55 countries One of 10 winners of DARPA Grand Challenge 2015 for predicting CHIKV epidemic

#### **Research Plan**

- 1. Challenge: Noisy and not enough data
- Model unreported cases
- Use infection rates from "similar" regions
- Mobility plays a major role
- Identify and evaluate accuracy vs specificity trade-off
- 2. Forgetting factor to incorporate **changing trends** and policies (social distancing)
- 3. Prior flu seasonal patterns and COVID19+weather data for **long-term** predictions
- 4. Various **training and constraints strategies** for model parameters to compare and create ensembles.
- 5. Interactive tool generating predictions, analyzing scenarios and performing resource allocation



#### Progress

Already obtained accurate country-level and US state-level short-term predictions



### Aram Galstyan and Fred Morstatter

# Forecasting the Impact of COVID-19 with SAGE

### **Aram Galstyan and Fred Morstatter**

MACHINE INTELLIGENCE AND DATA SCIENCE



#### April 3, 2020

### **Artificial Intelligence Components**

• Provide background info with **qualitative** data.

Search articles
Angela Merkel will not seek re-election as German Chancellor
German Chancellor Angela Merkel announced on Monday that she would not seek re-election when her term expires in 2021. CNN
The Communards: Berlin's lack of housing prompts radical expropriation initiative The capital's lefty government is mulling ever more extreme measures to close the housing gap, even as burdensome regulations slow new building

www.handelsblatt.com

#### Angela Merkel REJECTS AfD claims she ENDANGERED German citizens following mass riots

ANGELA Merkel has responded to allegations she endangered the lives of her citizens with her open border immigration policy as the Alternative for Germany party (ATD) attack the German Chancellor.

Express.co.uk 🖄 12 September 2018

• Generate stats and forecasts with **quantitative** data.



 Provide a base rate questions with **no data**.

SAGE





### **COVID-19 Deaths in California**







#### Become a forecaster: <u>https://sage.isi.edu/</u>

### Pedro Szekely

## **COVID 19 Knowledge Graphs**

Center On Knowledge Graphs, Pedro Szekely (szekely@usc.edu)



- Sources of data
- scientific
- publications
- databases
- CSV and Excel
- Web pages
- PDF tables

- data cleaning and
- improvement
- integration of multiple sources
- entity resolution
- graph embedding
- search and query

- Kaggle covid
- corpus
- extraction of
- entities
  - & relations
- Gene, Disease, ... DBs



### Jay Pujara

### MacroScore: Project Overview by Jay Pujara

Part of the DARPA SCORE program

<u>jpujara@usc.edu</u>

**Goal**: Assess the quality and importance of scientific articles using a combination of information extraction and network analysis



Key ideas:

- Extract "micro" features of studies extracted from scientific articles
- Identify "macro" features from social network analysis and bibliometrics
- Combine these features (and curate claims) in a knowledge graph
- Discover profiles of research patterns that lead to high-quality research
- Use explainable probabilistic models to predict research quality

Information Sciences Institute

### MacroScore Knowledge Graph

As shown in Figure 1,  $CD_{25}$  was significantly related to SSS but not to objective social status (bivariate relationships between study variables are presented in Table 2). Hierarchical regression models were separately conducted for each measure of SSS (entered on step 3) to adjust for sociodemographic covariates (entered on step 1) and objective socioeconomic status (HTFI was entered on step 2). Results for step 1 indicated that this model significantly explains variance in  $CD_{25}$  (p = .019,  $R^2 = 0.173$ ) with gender ( $\beta = -0.341$ , p = .002) being a significant predictor of  $CD_{25}$  but not





citation ego-network of a reproducible paper

### Jonathan May

### Rapid Social Media-Driven Emergency Response Tools

Jonathan May, Emilio Ferrara, Mozhdeh Gheini, Thamme Gowda, Ulf Hermjakob, Shri Narayanan, Emily Shen, Karan Singla

- Problem: How to quickly provide emergency response when reporting is in a language we don't have capabilities for?
- Solution: speech recognition, translation, name finding, and topic detection for a new data se tin 24 hours that can leverage the limited data available
- Outcome: early aid decisions in the hands of providers; prevent mass casualties



### **Extension to COVID-19 Scenario**

- What actionable situations can we detect from the millions of coronavirus-related tweets collected by Emily Chen and Emilio Ferrara?
  - Hoarding
  - PPE needs
  - Depression
  - Employment/food/shelter
  - Civil disobedience (gathering, ignoring quarantine, strikes)
- Can we expand from the manual list of relevant keywords to better support discovery?
  - Manual list: "coronavirus", "covid-19", "SocialDistancing"
  - Discovered: "drtedros", "epicenter", "paho"

### Paul Bogdan and Xiong Ye Xiao

### Multifractal Analysis of COVID-19 RNA nets

#### Generate network from COVID-19 RNA sequence

- Divide the original RNA sequence (containing 29903 bases) into 996 equal-length basic sequence (each containing 30 bases)
- Consider each basic sequence as a node
- □ Use 1,2,3,4 to denote the base A,C,G,T and calculate the Spearman correlation coefficient (SCC) between the basic sequences
  - □ If abs(SCC)>0.5, we consider two nodes are connected and use SCC as the weight of the edge
  - Analyze these weighted graphs

#### **Compare the spectrum of the original RNA net and the mutated RNA net**

We find four RNA sequence to generate networks and use MFA to compare their fractal structures

### Multifractal Analysis of COVID-19 RNA nets



### Wrap-Up