Follow along at `bit.ly/GTCSE2020`

"Hot CSE"

Some hot topics in
Computational Science and Engineering
at Georgia Tech

Presented by Kasimir Gabert & Srinivas Eswar
August 12, 2020

Georgia Tech College of Computing
Computational Science and Engineering

Kasimir Gabert

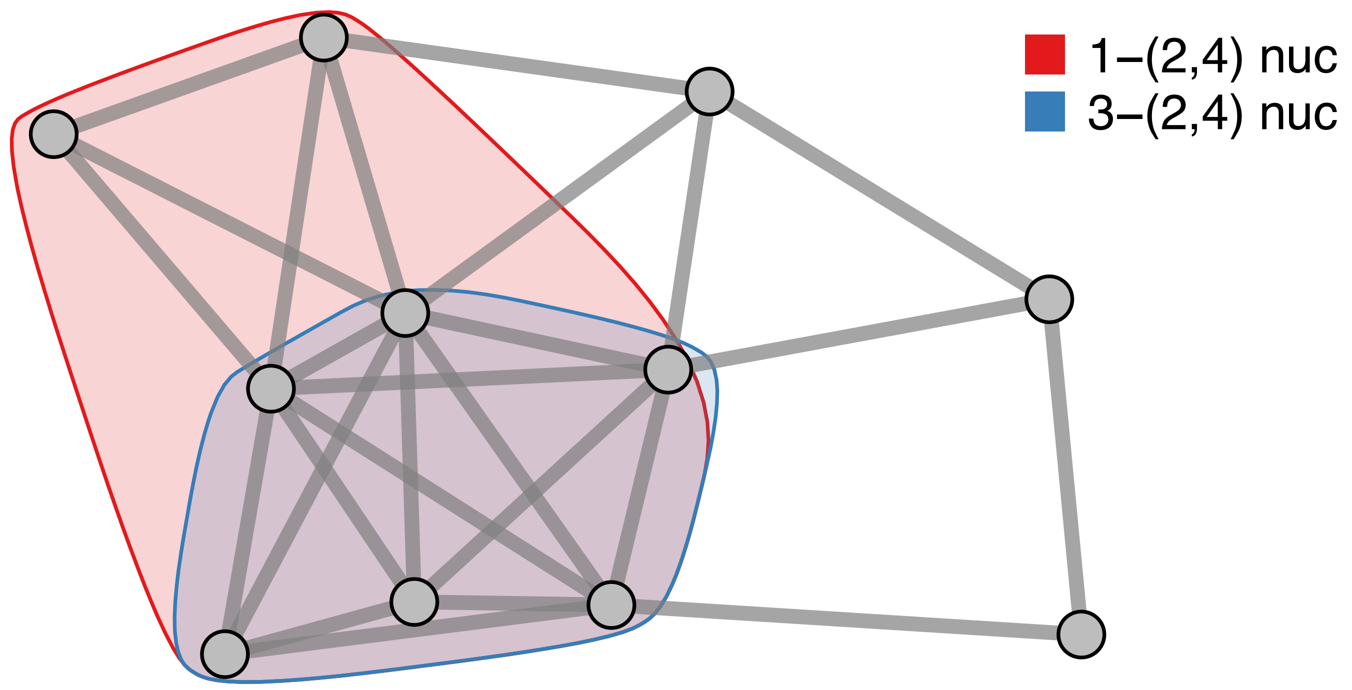
Ph.D. Student in CSE
5th year

Advised by Dr. Ümit Çatalyürek
Research areas: Dynamic Graph Algorithms / Systems
Interned at Sandia National Labs
Worked at Sandia National Labs

M.S. CS from Georgia Tech (2012)
B.S. Math and CS from New Mexico Tech (2011)

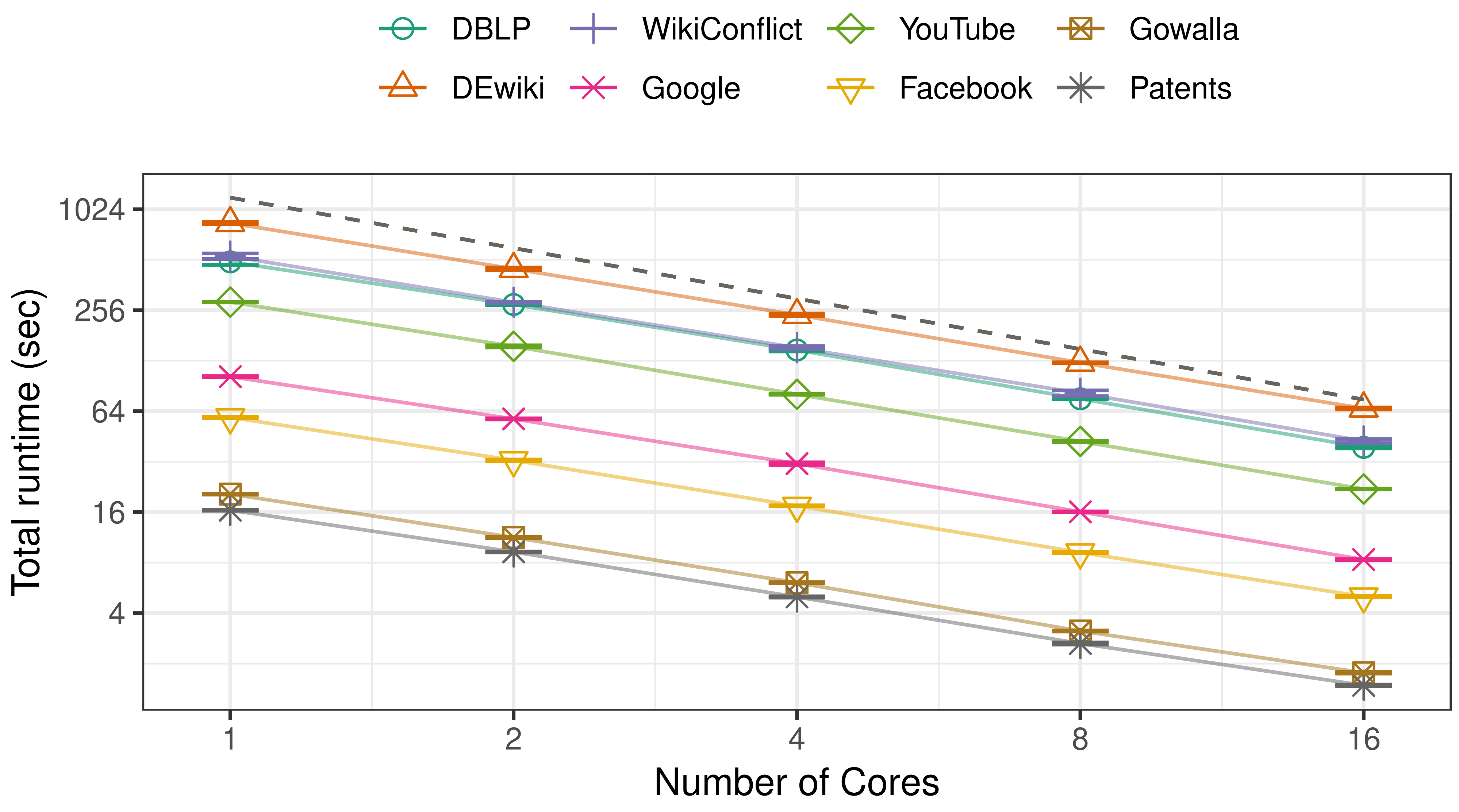
Discovering and Maintaining Dense Hierarchies

Goal: Uncover dense regions and hierarchies in graphs that **continuously change**



Approach: Reframe the more effective **nucleus** problem into a common **core** problem

Solve this quickly and in **parallel** with **local algorithms**

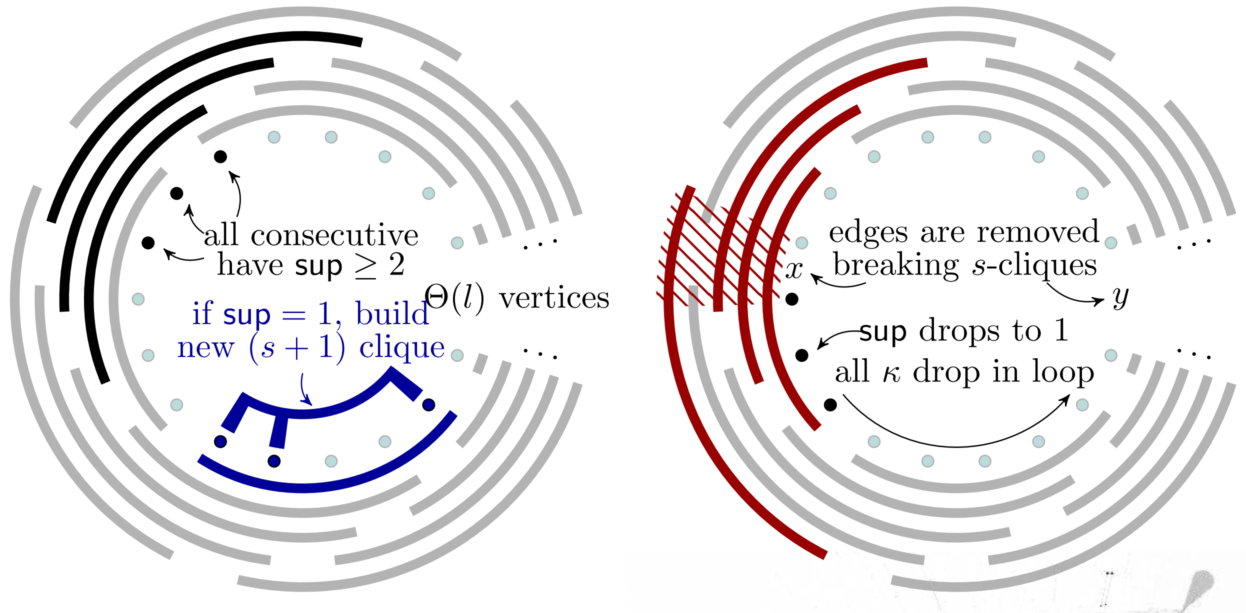


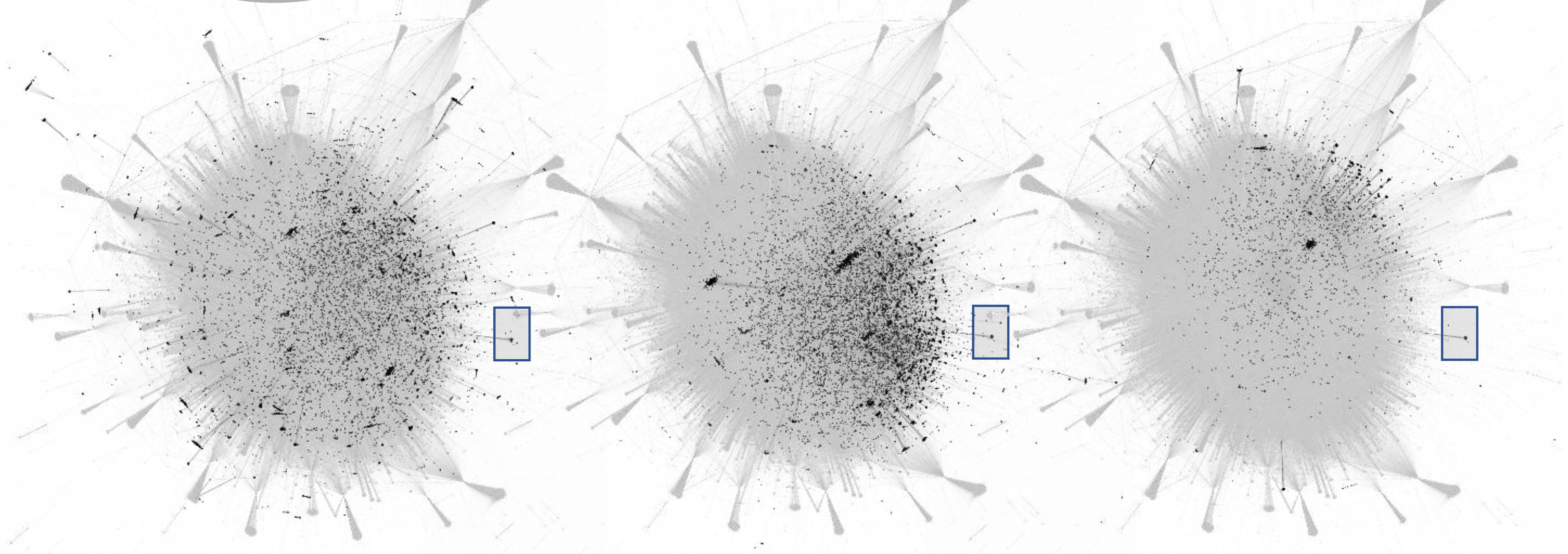
Addressing Variability in Dynamic Graphs

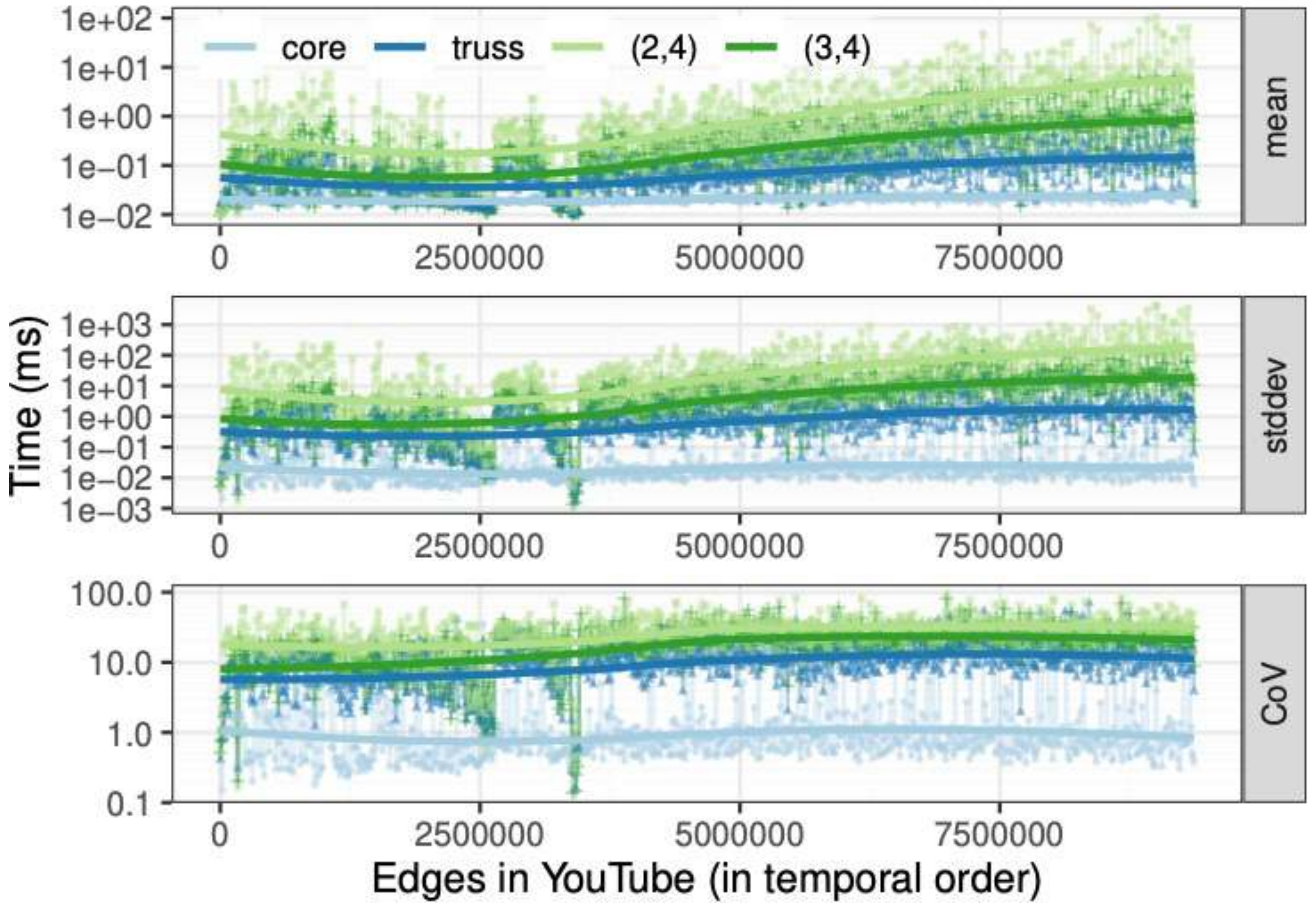
Problem: dynamic graph algorithms experience **significant variability** in latency and even output

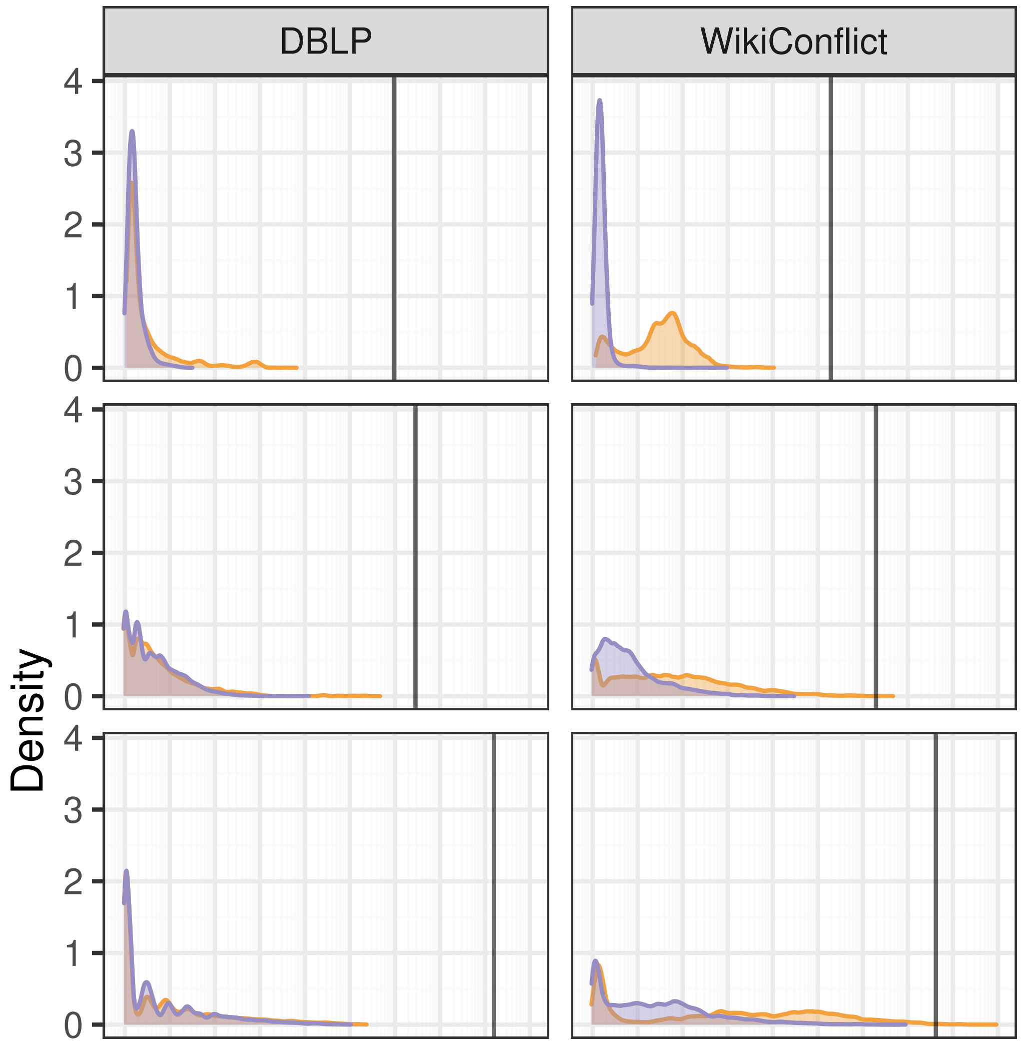
This is seen both **theoretically** and **empirically**

We need to address this with **new systems and algorithms** that address variability directly


all consecutive have sup ≥ 2
if sup $= 1$, build new $(s+1)$ clique
$\Theta(l)$ vertices
...
...
edges are removed breaking s-cliques
x
y
sup drops to 1
all κ drop in loop
...
...







Srinivas Eswar

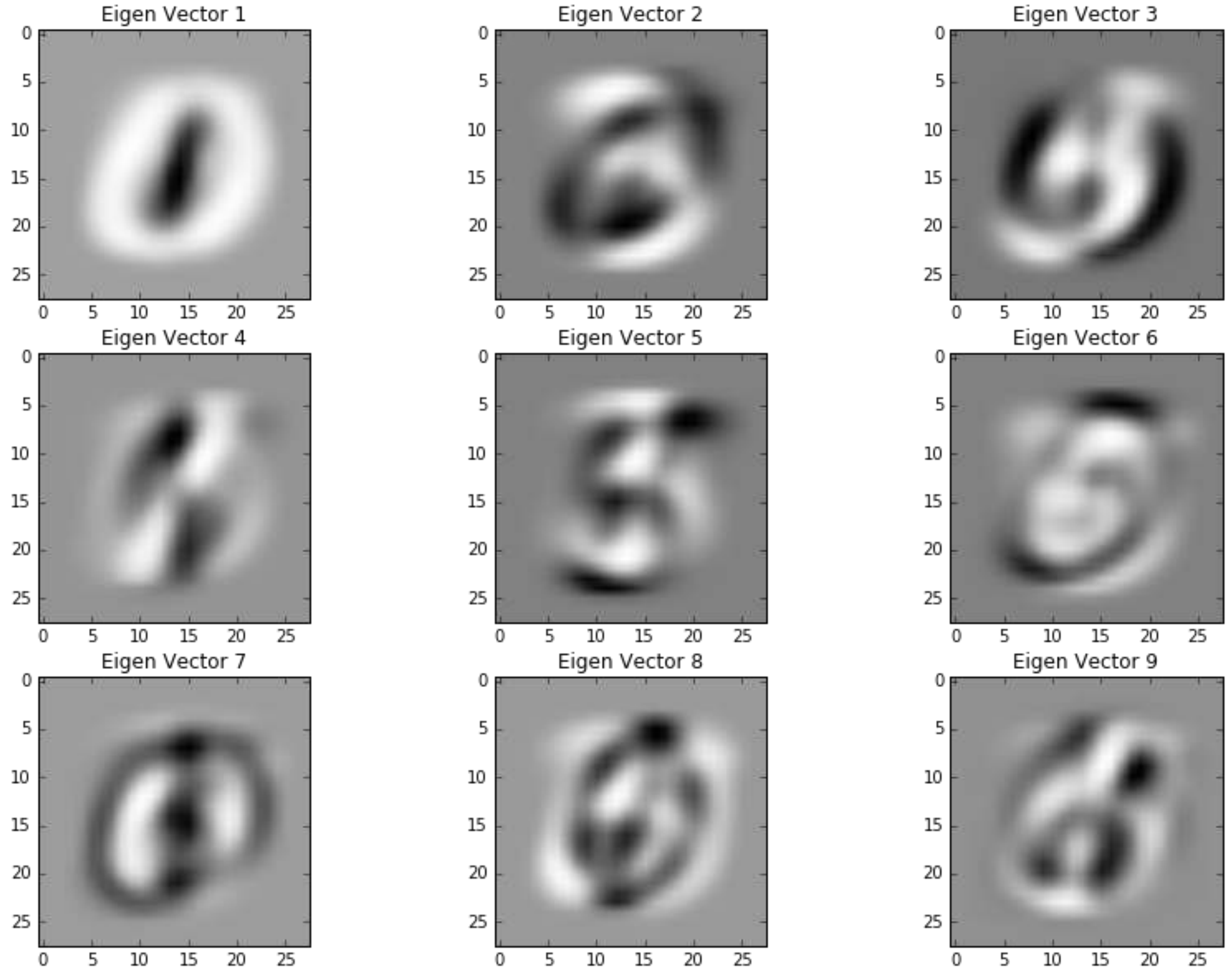
Ph.D. Student in CS
5th year

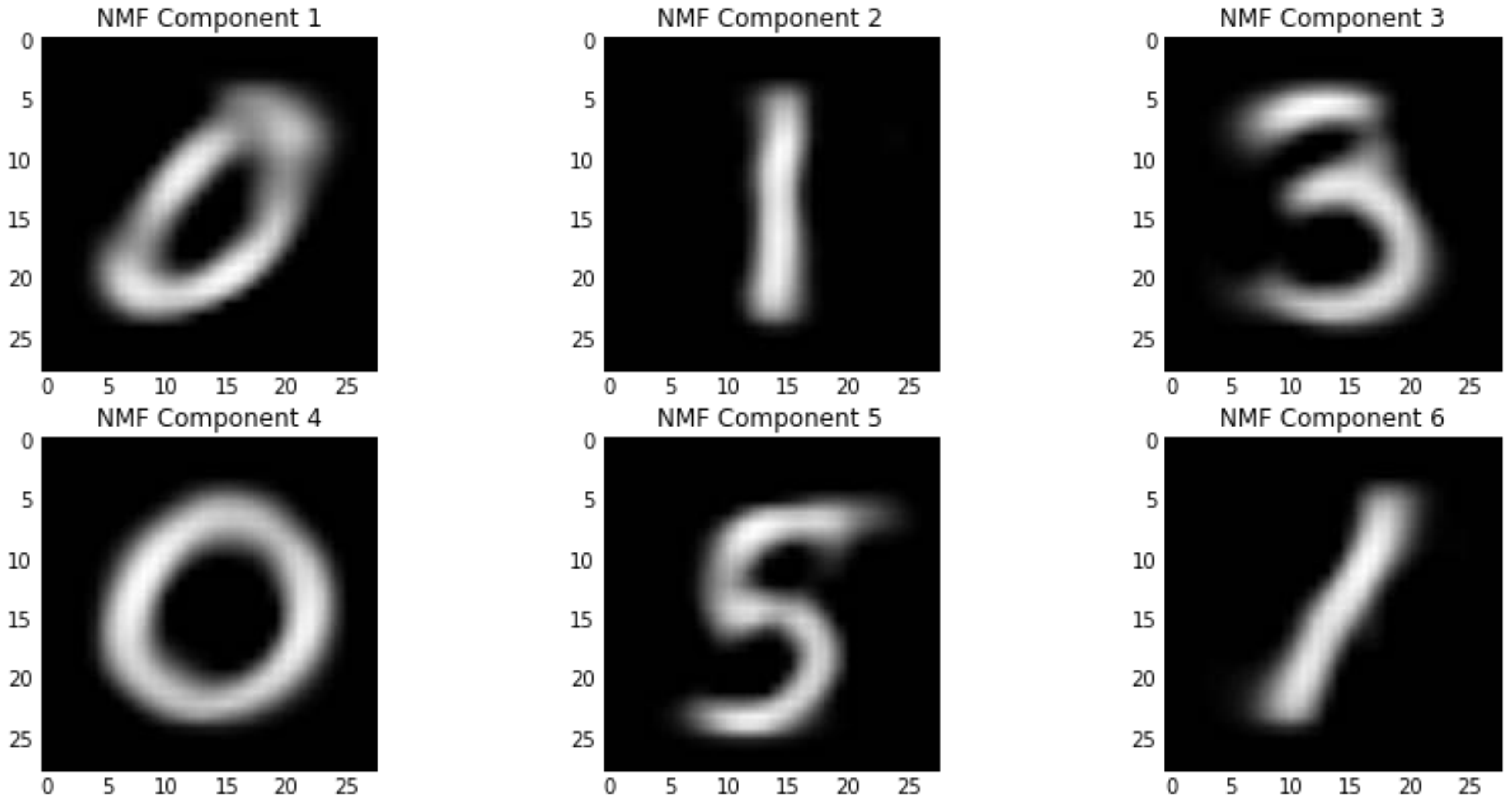
Advised by Dr. Richard Vuduc and Dr. Haesun Park
Research areas: HPC, Matrix/Tensor Factorizations
Interned at Arm and Oak Ridge National Laboratory
Worked at Citibank

M.S. CS from Georgia Tech (2016)
B.E. and M.Sc. Computer Science and Mathematics from BITS-Pilani (2012)

Nonnegative Matrix Factorization

- Decomposing signals into components
  - Alternative approach to Principal Component Analysis
  - Provides **interpretable** models with parts-based features
  - Can be adapted to various data types via regularization





Who is who, and doing **what**, in CSE @ GT?

Samples of papers from the last few years.

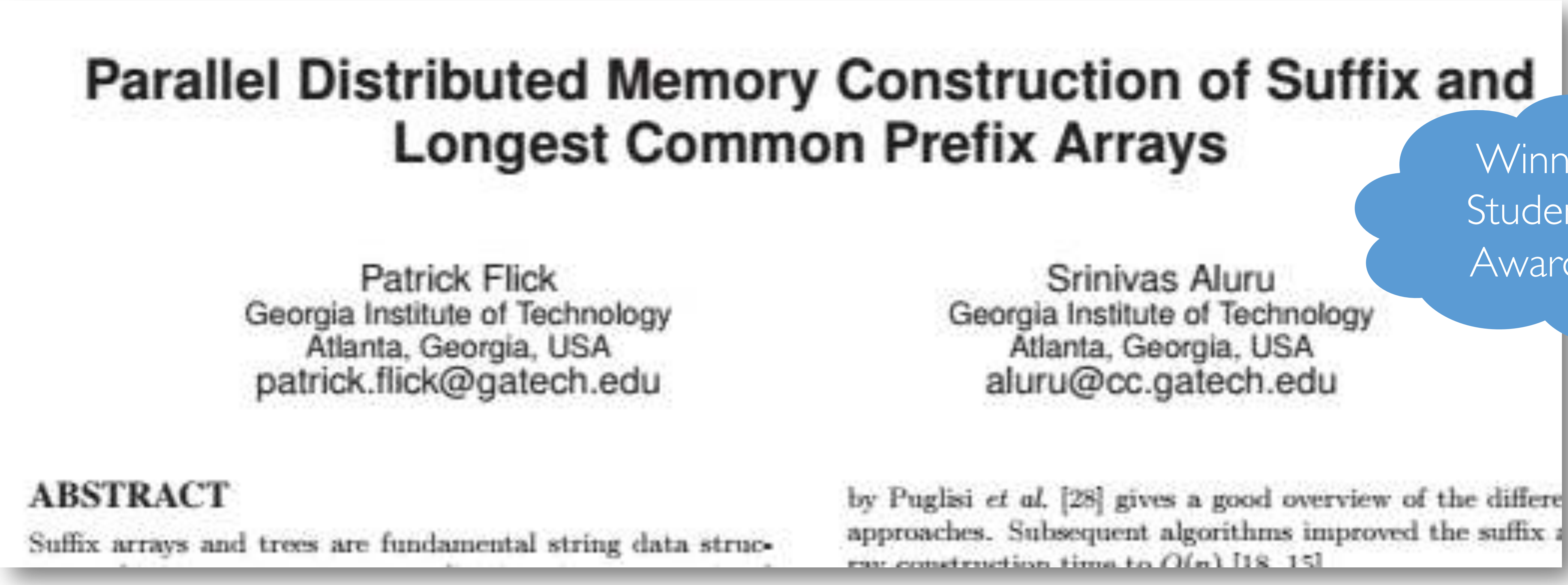
Follow along at bit.ly/GTCSE2020

Srinivas **Aluru**

Executive Director of IDEaS
AIAA, IEEE, SIAM Fellow

HPC in biology

Using parallel computing in bioinformatics and the modeling and analysis of complex biological systems.



Parallel Distributed Memory Construction of Suffix and Longest Common Prefix Arrays

Patrick Flick
Georgia Institute of Technology
Atlanta, Georgia, USA
patrick.flick@gatech.edu

Srinivas Aluru
Georgia Institute of Technology
Atlanta, Georgia, USA
aluru@cc.gatech.edu

ABSTRACT

Suffix arrays and trees are fundamental string data struc-

by Puglisi *et al.* [28] gives a good overview of the differe approaches. Subsequent algorithms improved the suffix a ray construction time to $O(n)$ [18, 15]

Winner: Best Student Paper Award, SC'15

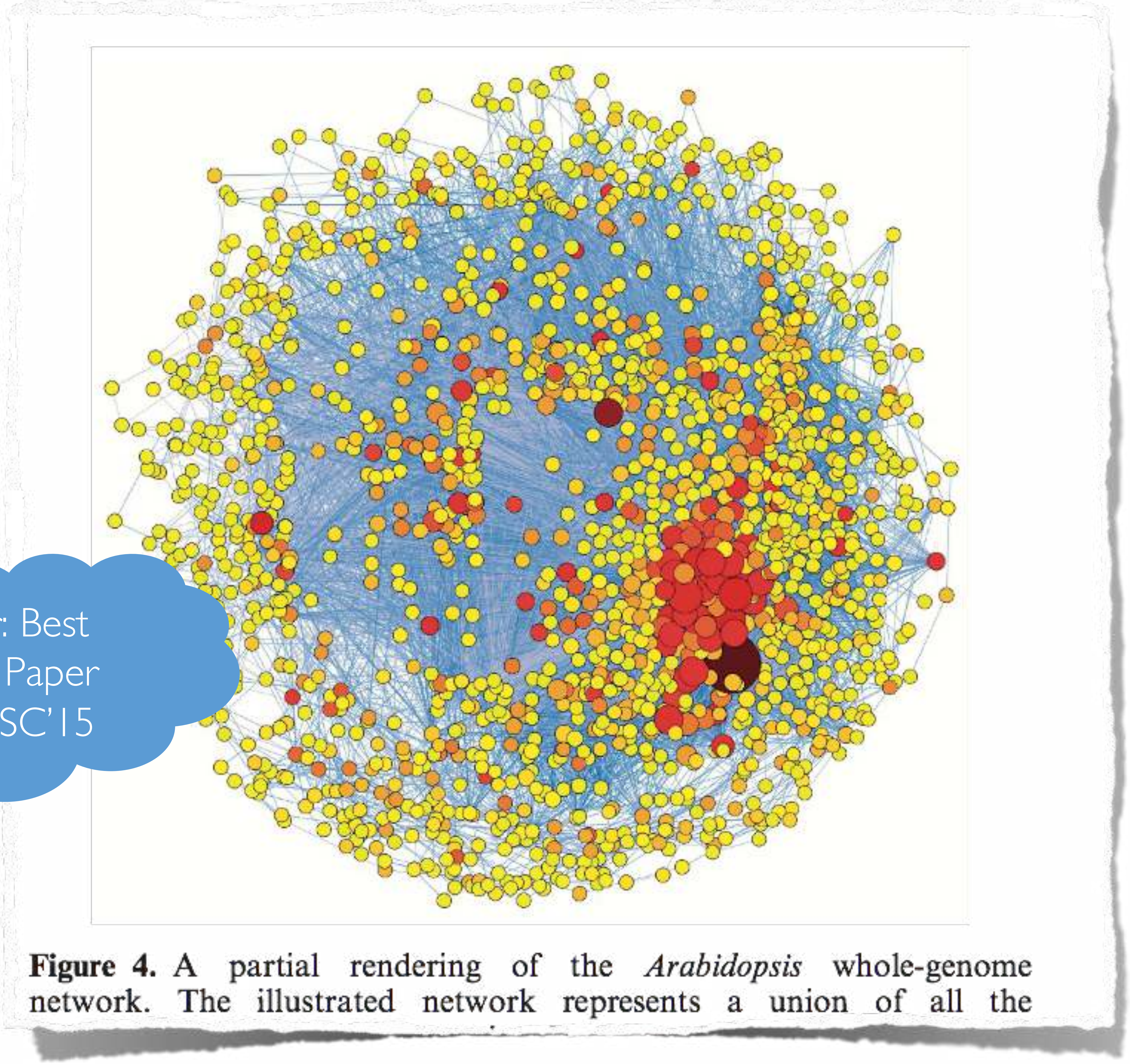


Figure 4. A partial rendering of the *Arabidopsis* whole-genome network. The illustrated network represents a union of all the

Follow along at bit.ly/GTCSE2020

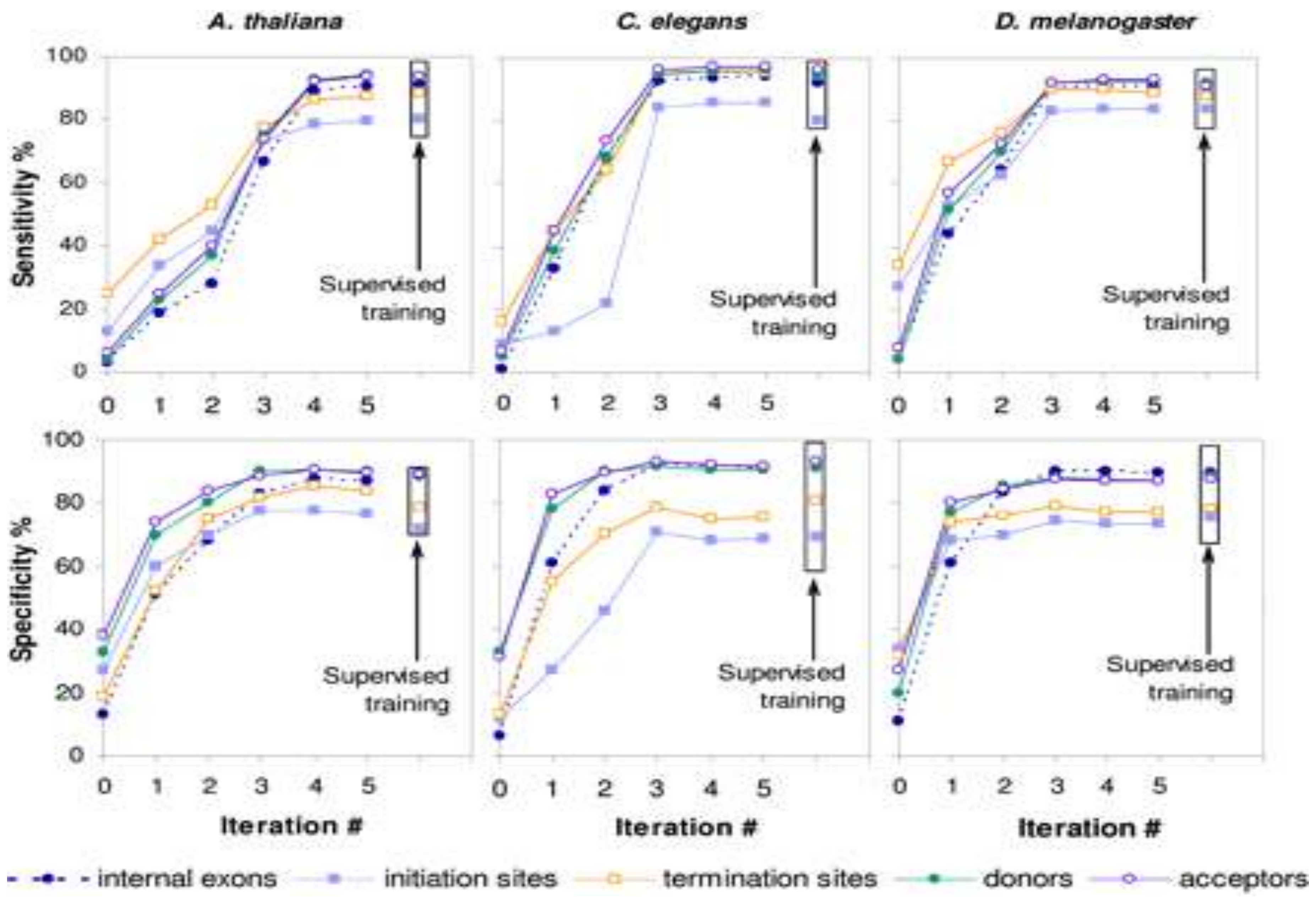
Mark Borodovsky

AIMBE Fellow, Regents’ Professor
Joint with Biomedical Engineering
Founder of the GT Bioinformatics graduate program

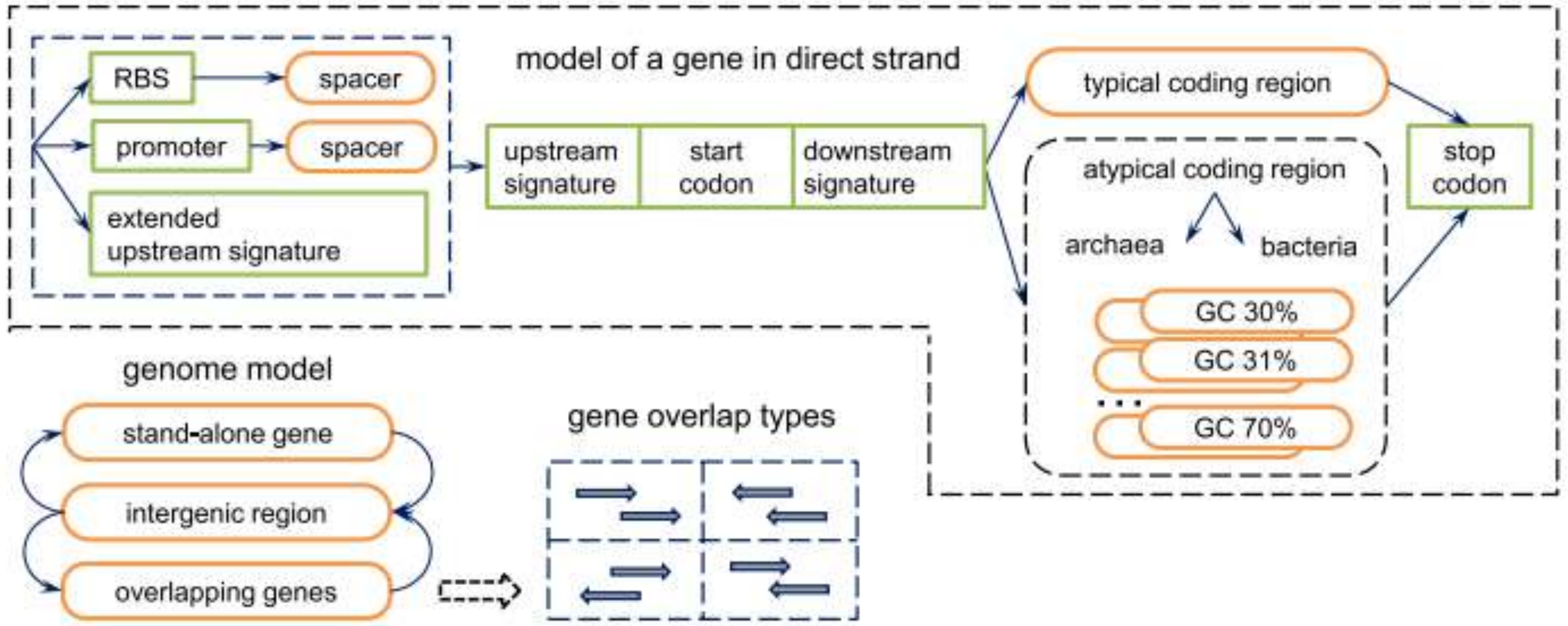
Bioinformatics algorithms

Developing algorithms for inferring genome function from its primary structure, particularly the gene finding algorithms

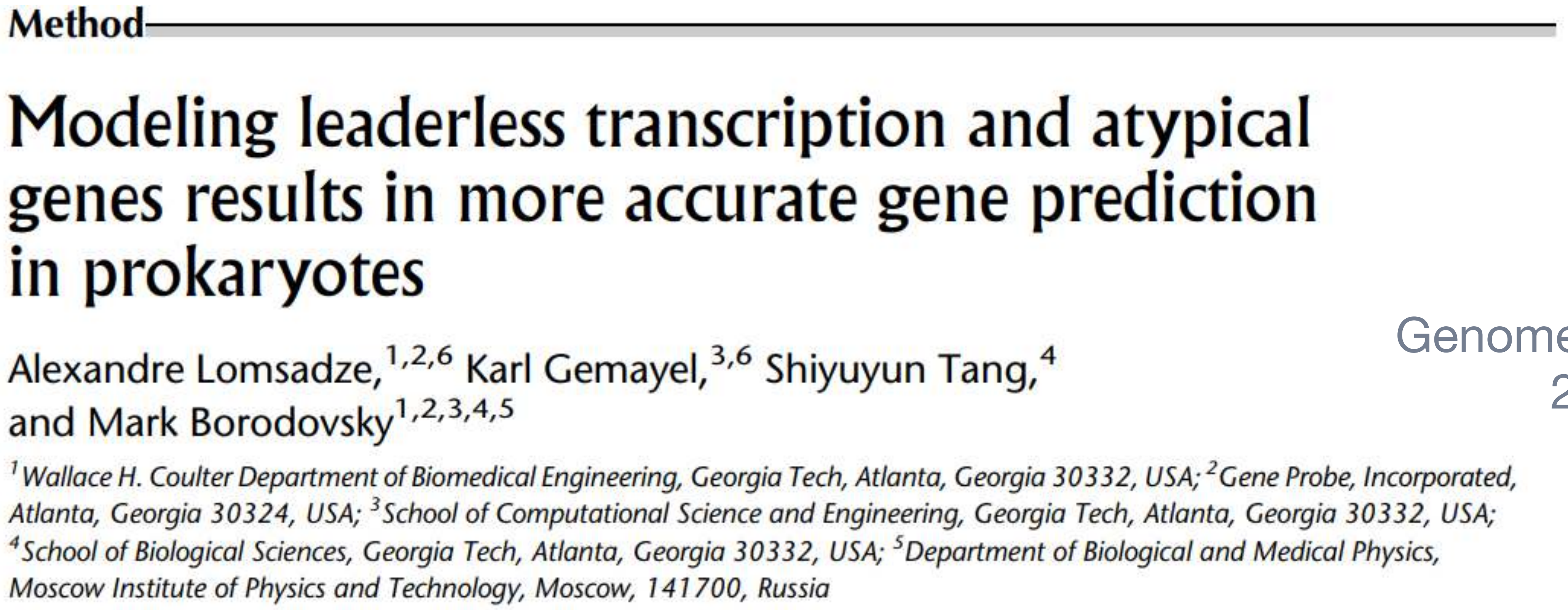
GeneMark-ES gene finding algorithm with unsupervised training for eukaryotic genomes



GeneMark-ES has no analogs among bioinformatics tools



Principal state diagram of the generalized hidden Markov model (GHHM) of prokaryotic genomic sequence. GeneMarkS-2 – self-training gene finder is a core of NCBI genome annotation pipline (805 citations since 2016)



Method

Modeling leaderless transcription and atypical genes results in more accurate gene prediction in prokaryotes

Alexandre Lomsadze,[1,2,6] Karl Gemayel,[3,6] Shiyuyun Tang,[4] and Mark Borodovsky[1,2,3,4,5]

[1]Wallace H. Coulter Department of Biomedical Engineering, Georgia Tech, Atlanta, Georgia 30332, USA; [2]Gene Probe, Incorporated, Atlanta, Georgia 30324, USA; [3]School of Computational Science and Engineering, Georgia Tech, Atlanta, Georgia 30332, USA; [4]School of Biological Sciences, Georgia Tech, Atlanta, Georgia 30332, USA; [5]Department of Biological and Medical Physics, Moscow Institute of Physics and Technology, Moscow, 141700, Russia

Genome Research, 2018

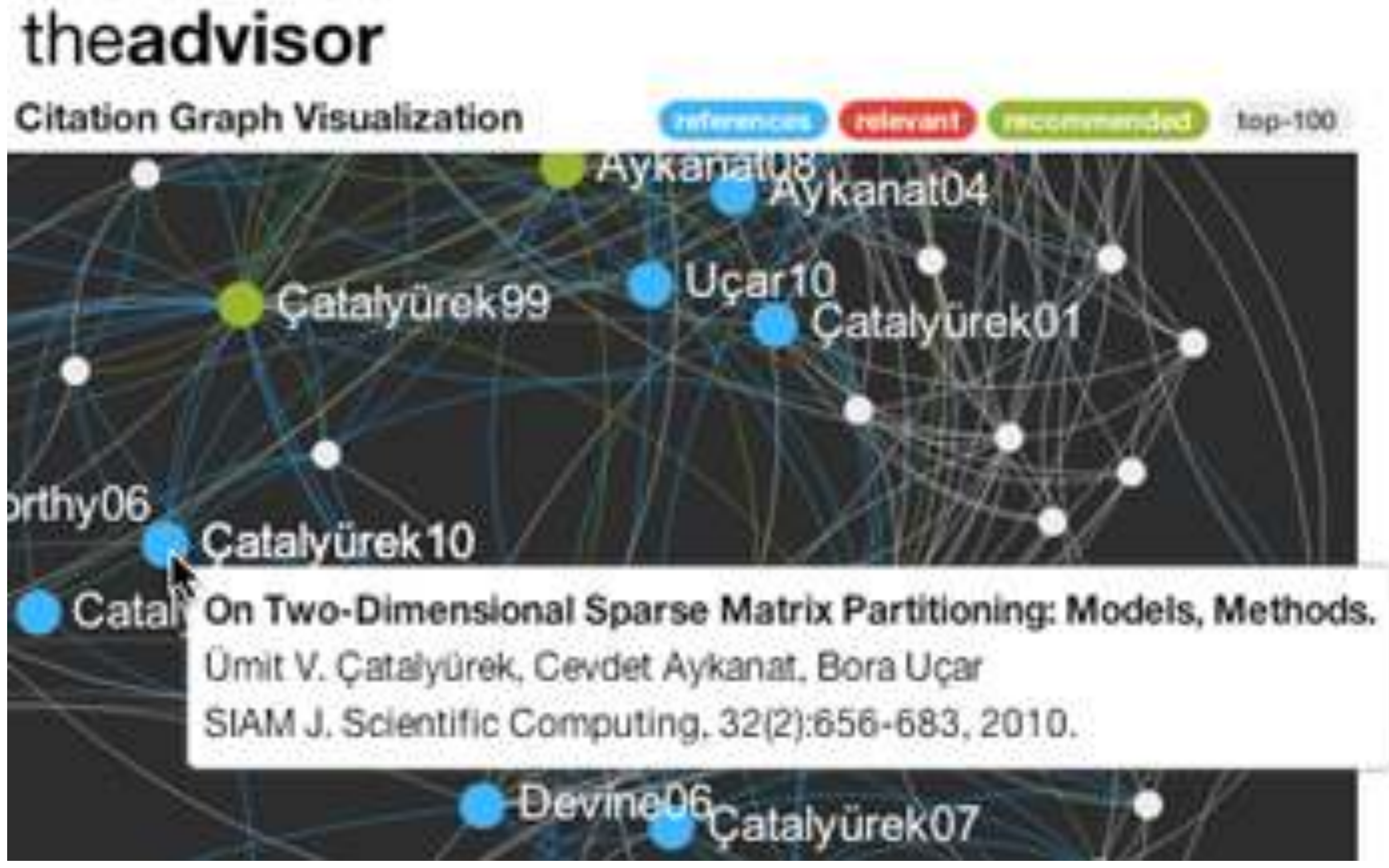
Follow along at bit.ly/GTCSE2020

Ümit Çatalyürek

IEEE & SIAM Fellow, CSE Assoc. Chair
board game geek

HPC & Translational data analytics

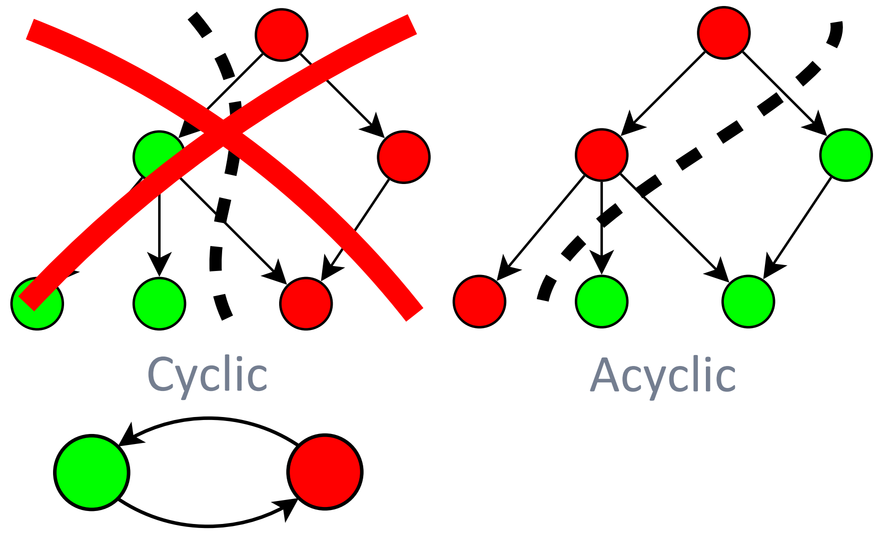
Developing scalable algorithms and systems to solve large-scale scientific computing, genomic and biomedical problems.

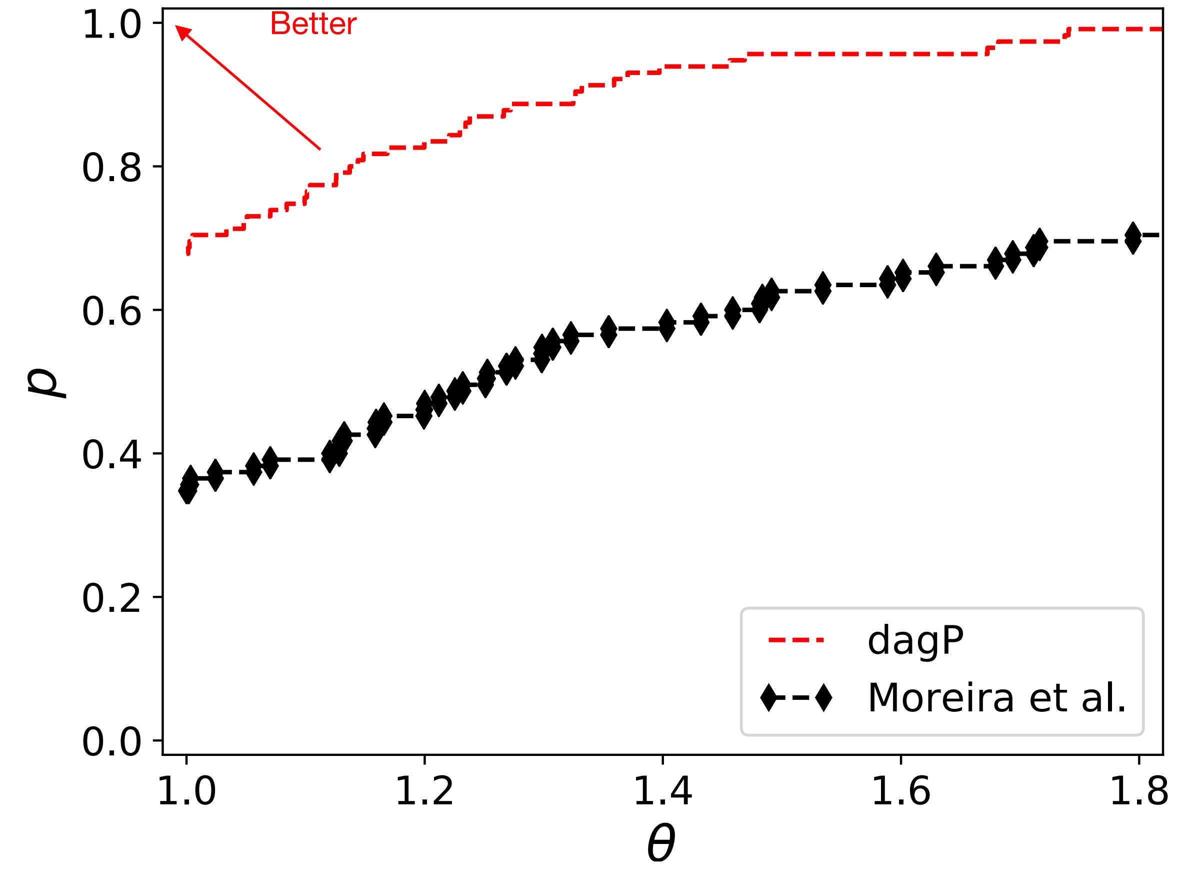


On Two-Dimensional Sparse Matrix Partitioning: Models, Methods.
Ümit V. Çatalyürek, Cevdet Aykanat, Bora Uçar
SIAM J. Scientific Computing, 32(2):656-683, 2010.

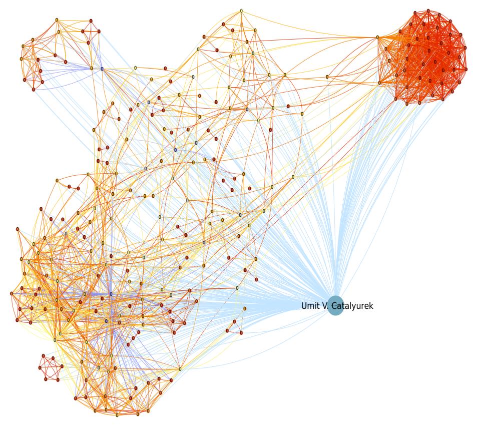
DAG Partitioning

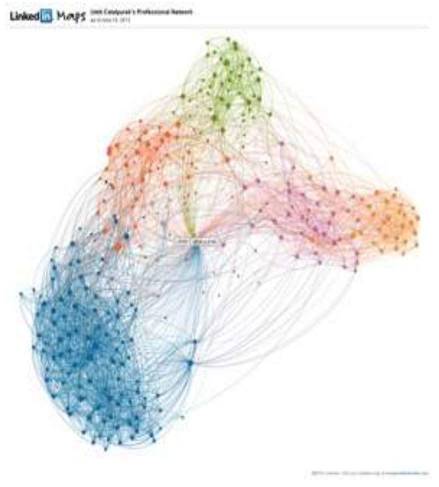
Partitioning a DAG in a way that the resulting partitions also form a DAG.





Research Track Paper


DBLP


Linkedin

KDD 2018, August 19–23, 2018, London, United Kingdom

An Iterative Global Structure-Assisted Labeled Network Aligner

Abdurrahman Yaşar
School of Computational Science and Engineering
Georgia Institute of Technology
Atlanta, Georgia
ayasar@gatech.edu

Ümit V. Çatalyürek
School of Computational Science and Engineering
Georgia Institute of Technology
Atlanta, Georgia
umit@gatech.edu

ABSTRACT

Integrating data from heterogeneous sources is often modeled as merging graphs. Given two or more "compatible" but not isomorphic

Merging two graphs involves identifying each vertex in a graph with a corresponding vertex (i.e., representing the same entity) in the other graph, whenever such corresponding vertices exist.

Follow along at bit.ly/GTCSE2020

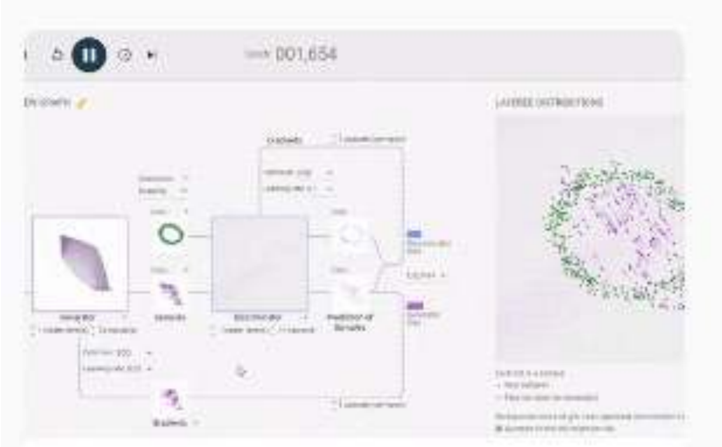
Polo **Chau**

"Data sensemaker," fights bad guys | Co-director, MS Analytics | ML Area Leader

Machine Learning + Visualization

Scalable, **interactive** & **interpretable** tools for understanding billion-scale data & ML models.

Human-centered AI



GAN Lab

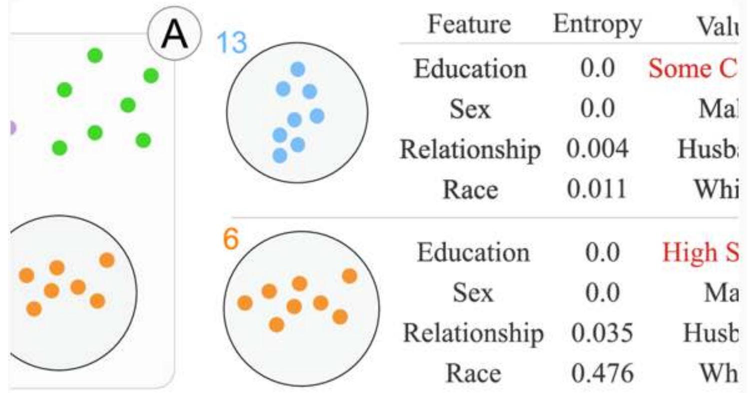
Playing with Generative Adversarial Networks in Browser

Google

ActiVis

Visual Exploration of Facebook Deep Neural Network Models

Deployed Facebook



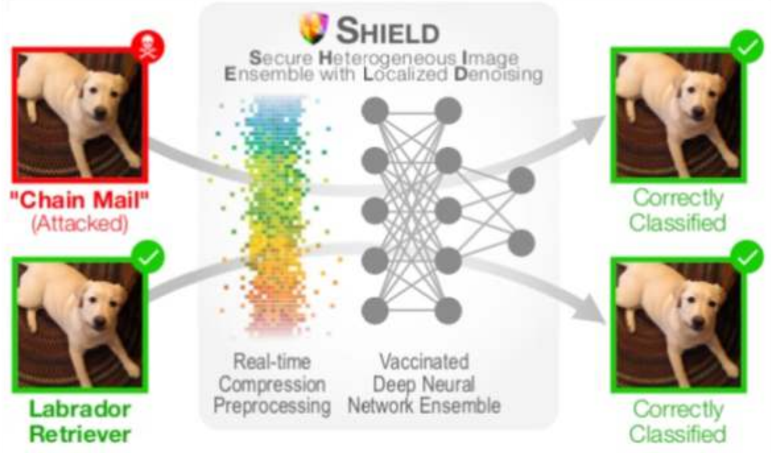
Feature	Entropy	Valı
Education	0.0	Some C
Sex	0.0	Mal
Relationship	0.004	Husb
Race	0.011	Whi

Feature	Entropy	Valı
Education	0.0	High S
Sex	0.0	Ma
Relationship	0.035	Husb
Race	0.476	Wh

Discovering Intersectional Bias

Discovery of Intersectional Bias in Machine Learning Using Automatic Subgroup Generation

ML security & Fraud


SHIELD
Secure Heterogeneous Image Ensemble with Localized Denoising
"Chain Mail" (Attacked)
Labrador Retriever
Real-time Compression Preprocessing
Vaccinated Deep Neural Network Ensemble
Correctly Classified
Correctly Classified

SHIELD

Fast, practical defense for deep learning

Audience Appreciation Award, Runner-up

ShapeShifter

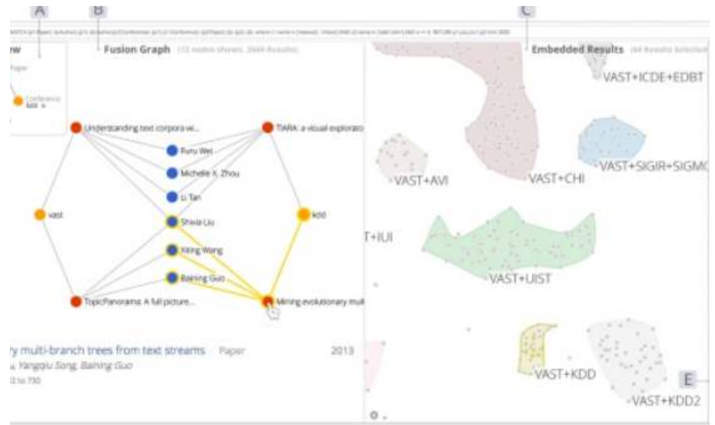
1st Targeted Physical Attack on Faster R-CNN Object Detector

MARCO

Fake Review Detection

SDM'14 Best Student Paper

Large Graph Mining & Visualization



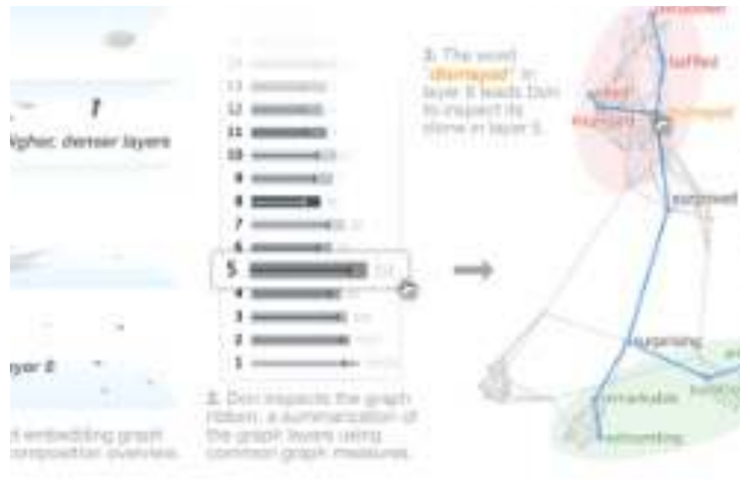
VIGOR

Interactive Visual Exploration of Graph Query Results

Symantec

M-Flash

Billion-Scale Graph Computation by Bimodal Block Processing



Atlas

Local Graph Exploration in a Global Context

Social Good & Health

DeepPop

Deep Learning on Satellite Imagery for Population Estimation

Microsoft AI for Earth

Firebird

Predicting Fire Risk in Atlanta

KDD'16 Best Student Paper, runner-up

Deployed Atlanta Fire Rescue Department

mHealth Visual Discovery Dashboard

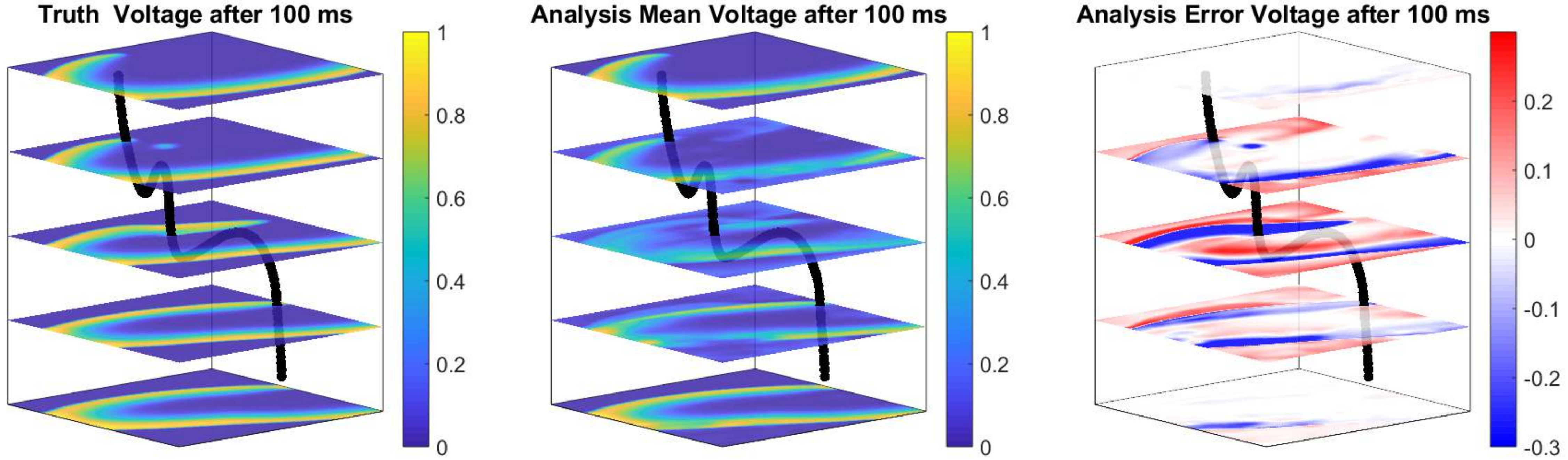
Making Sense of Mobile Health Data

Elizabeth **Cherry**

Computational modeling of cardiac arrhythmias

Computational modeling of heart

Improving the understanding of cardiac electrical dynamics in normal & diseased states; Designing advanced strategies for prevention & treatment of arrhythmias using mathematical modeling and sim.



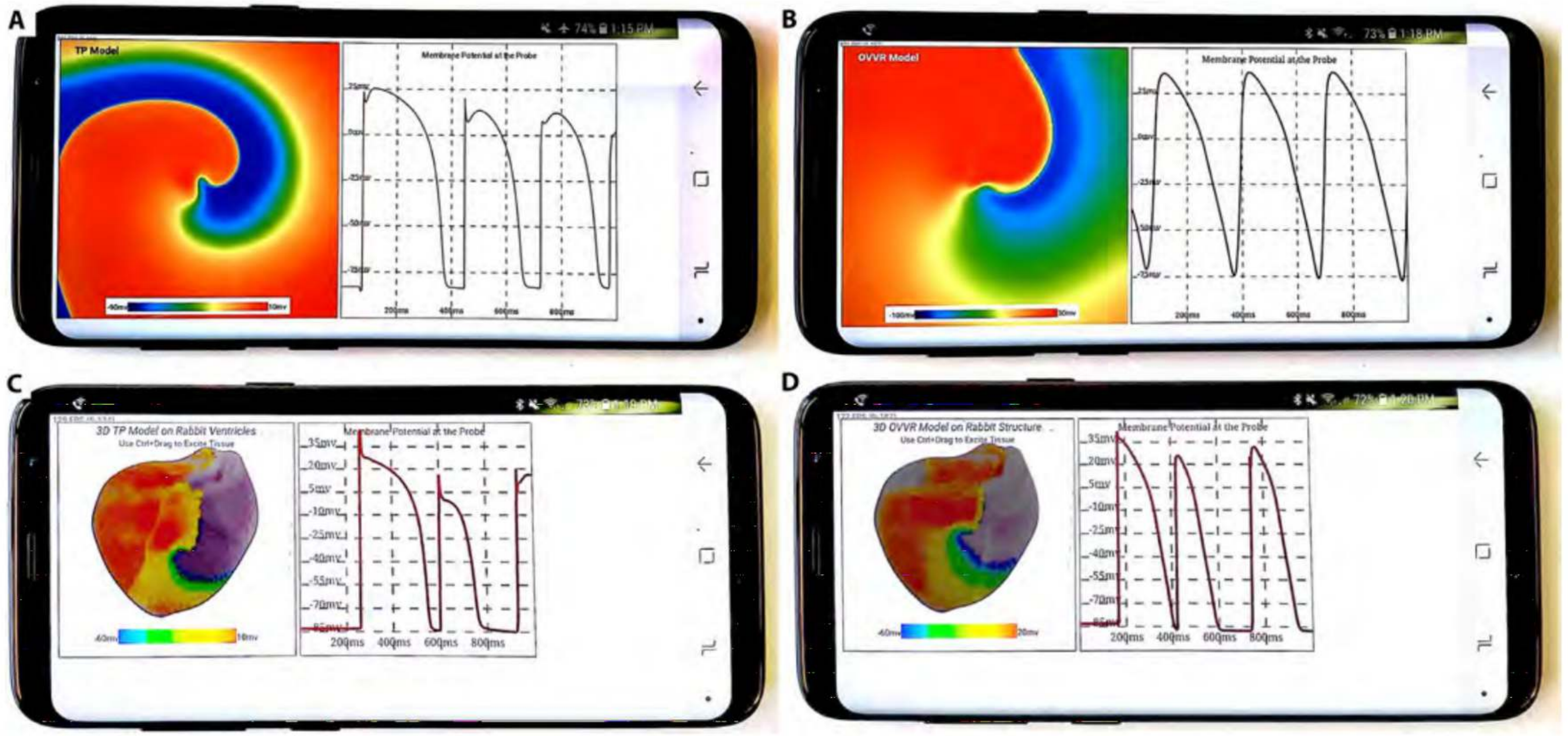
Using data assimilation to improve model predictions. Here, synthetic noisy observations of a reentrant electrical scroll wave (ventricular fibrillation) in a 3D tissue slab at the top and bottom surfaces are combined with model forecast to better match the known truth state.

SCIENCE ADVANCES | RESEARCH ARTICLE | 2019

COMPUTER SCIENCE

Real-time interactive simulations of large-scale systems on personal computers and cell phones: Toward patient-specific heart modeling and other applications

Abouzar Kaboudian[1], Elizabeth M. Cherry[2], Flavio H. Fenton[1]*


A
TP Model
Membrane Potential at the Probe
B
OVVR Model
Membrane Potential at the Probe
C
3D TP Model on Rabbit Ventricles
Membrane Potential at the Probe
200ms
400ms
600ms
800ms
D
3D OVVR Model on Rabbit Structure
Membrane Potential at the Probe
200ms
400ms
600ms
800ms

GPU simulations of reentrant electrical waves in 2D and 3D hearts can run interactively in near real time even on cell phones. (solving **1.7 billion** differential equations per second)

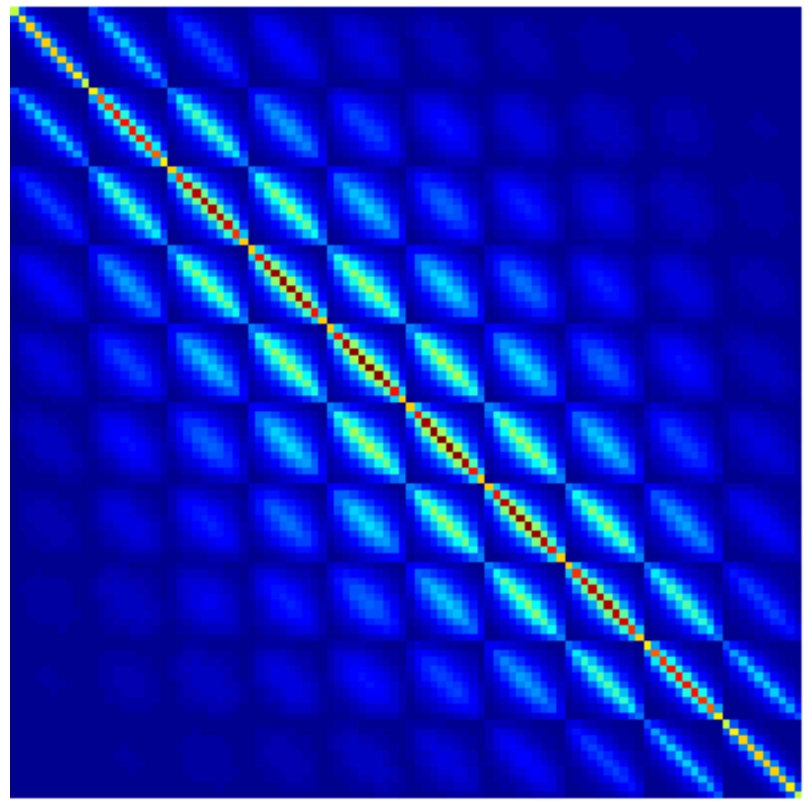
Follow along at bit.ly/GTCSE2020

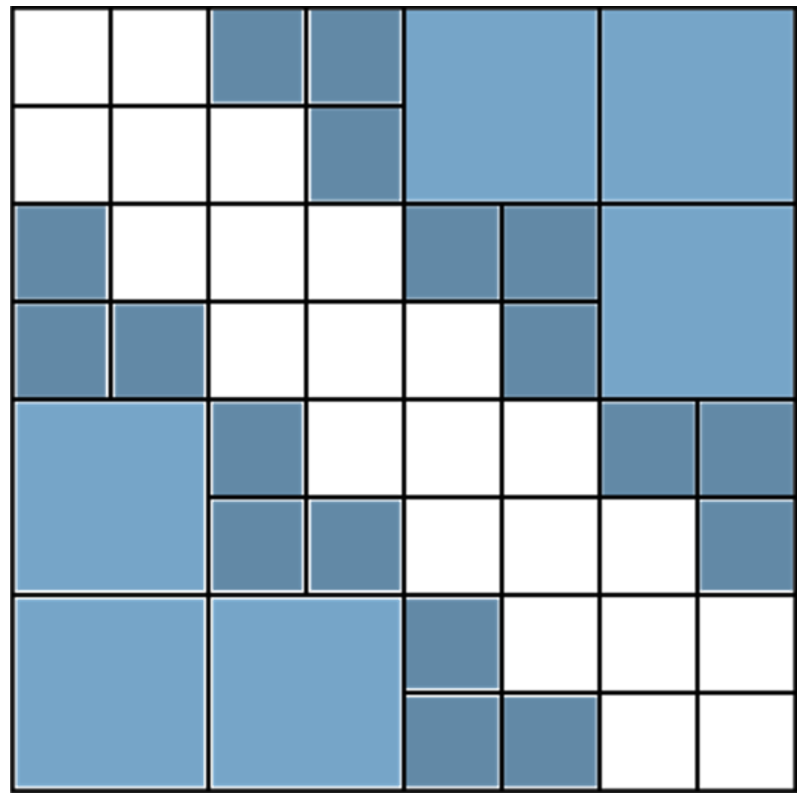
Edmond **Chow**

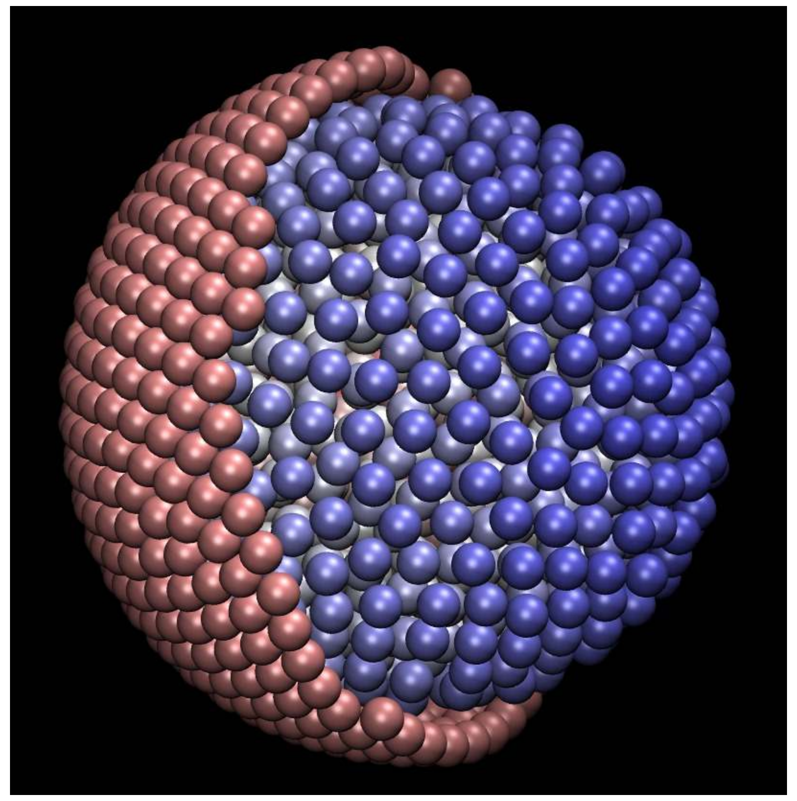
PECASE for research + best CSE cocktails

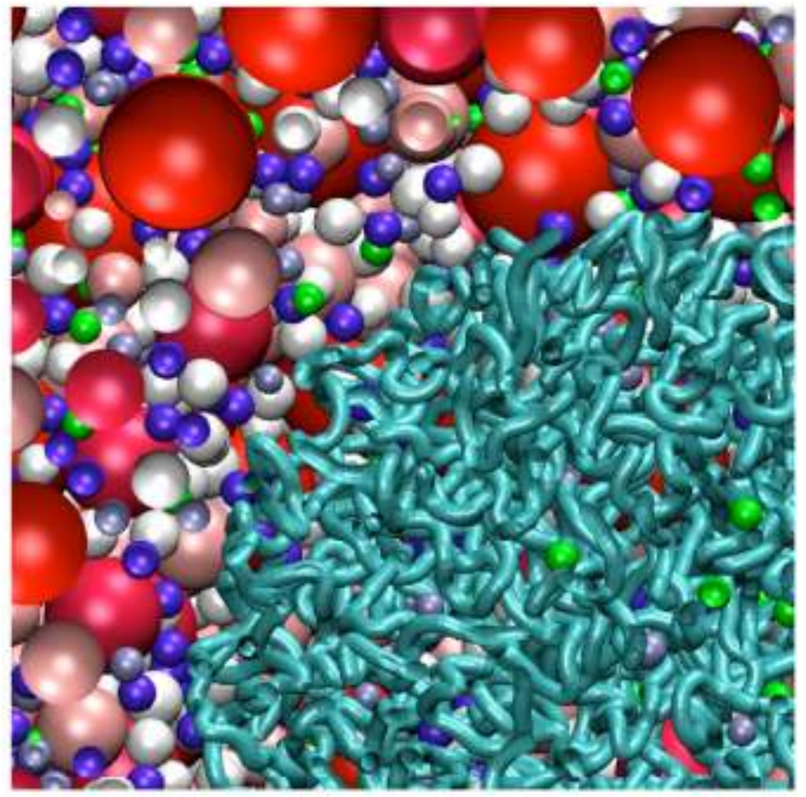
Scientific Computing at Extreme Scales

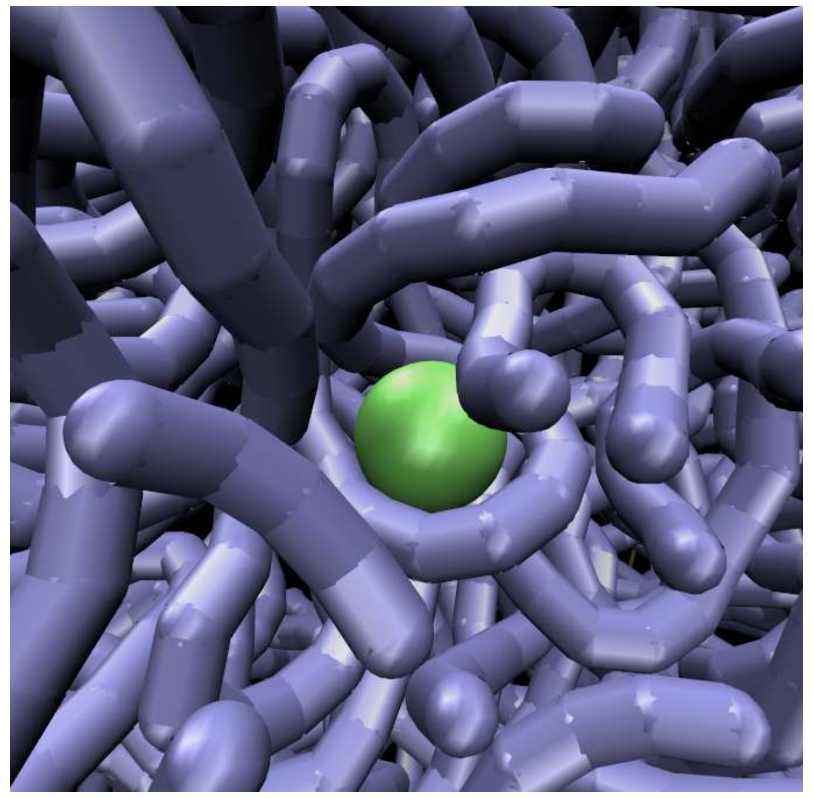
- Numerical linear algebra – design of highly parallel methods
- Scalable algorithms for computational physics – FMM and hierarchical matrices
- Quantum chemistry and materials science on GPU clusters
- Scientific machine learning – ML for scientific computing and numerical methods for ML







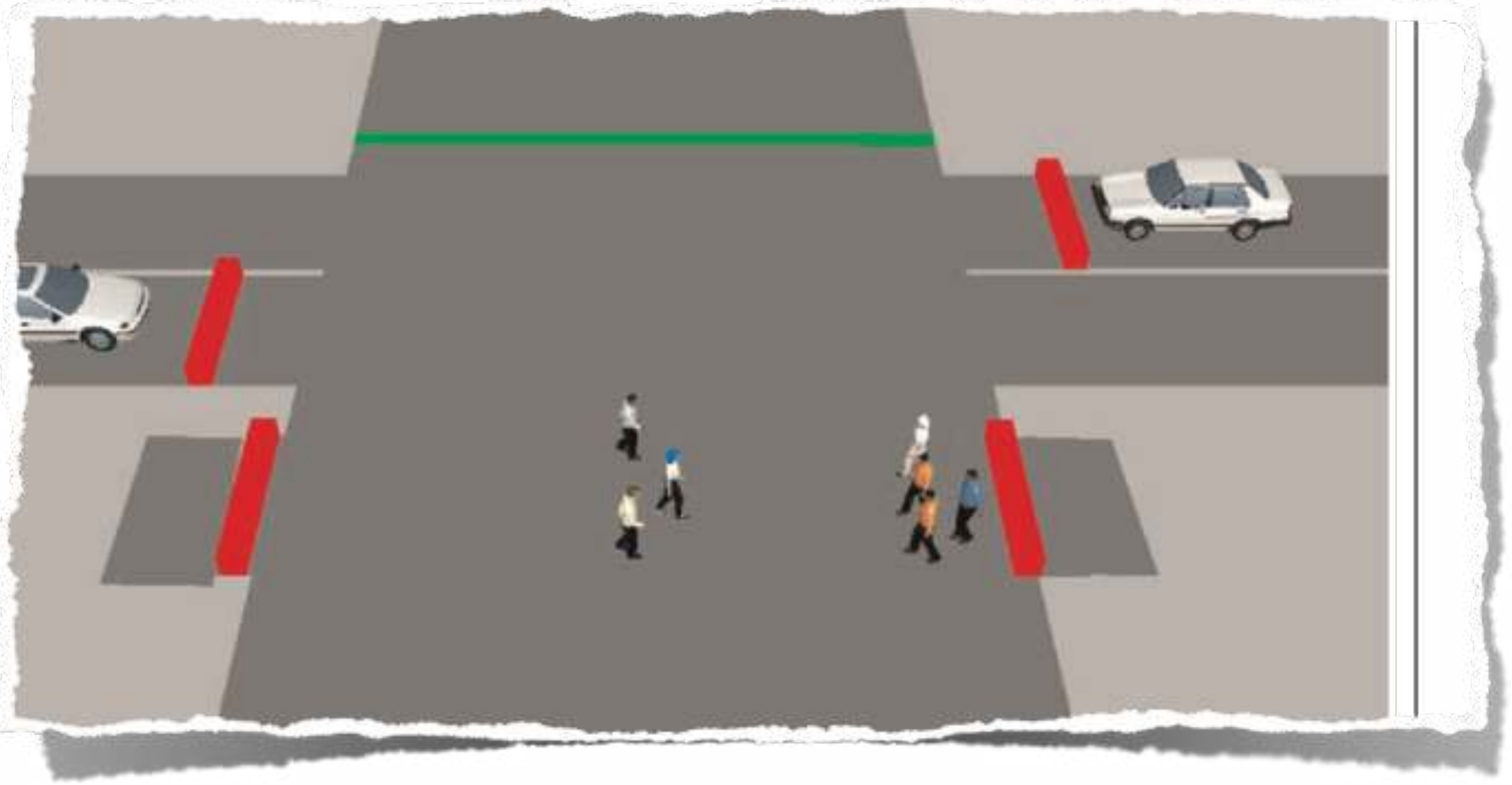




Follow along at bit.ly/GTCSE2020

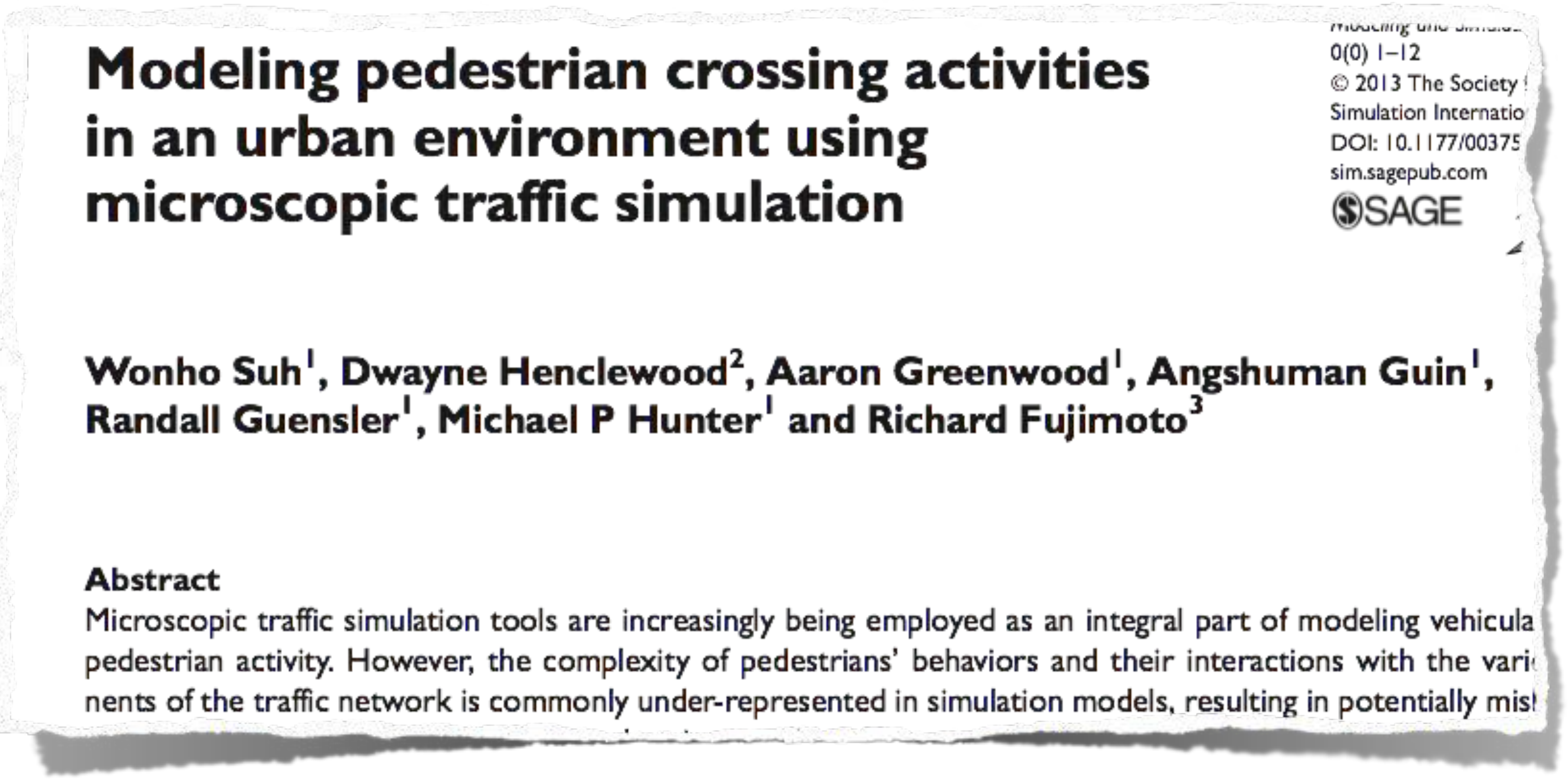
Richard **Fujimoto**

Regents' Professor, and all-around ninja



Parallel discrete-event simulation

Analyzing transportation systems, communication networks, and defense systems, on machines from mobile real-time platforms to supercomputers.



Modeling and Simulati
0(0) 1–12
© 2013 The Society f
Simulation Internatio
DOI: 10.1177/00375
sim.sagepub.com
SAGE

Modeling pedestrian crossing activities in an urban environment using microscopic traffic simulation

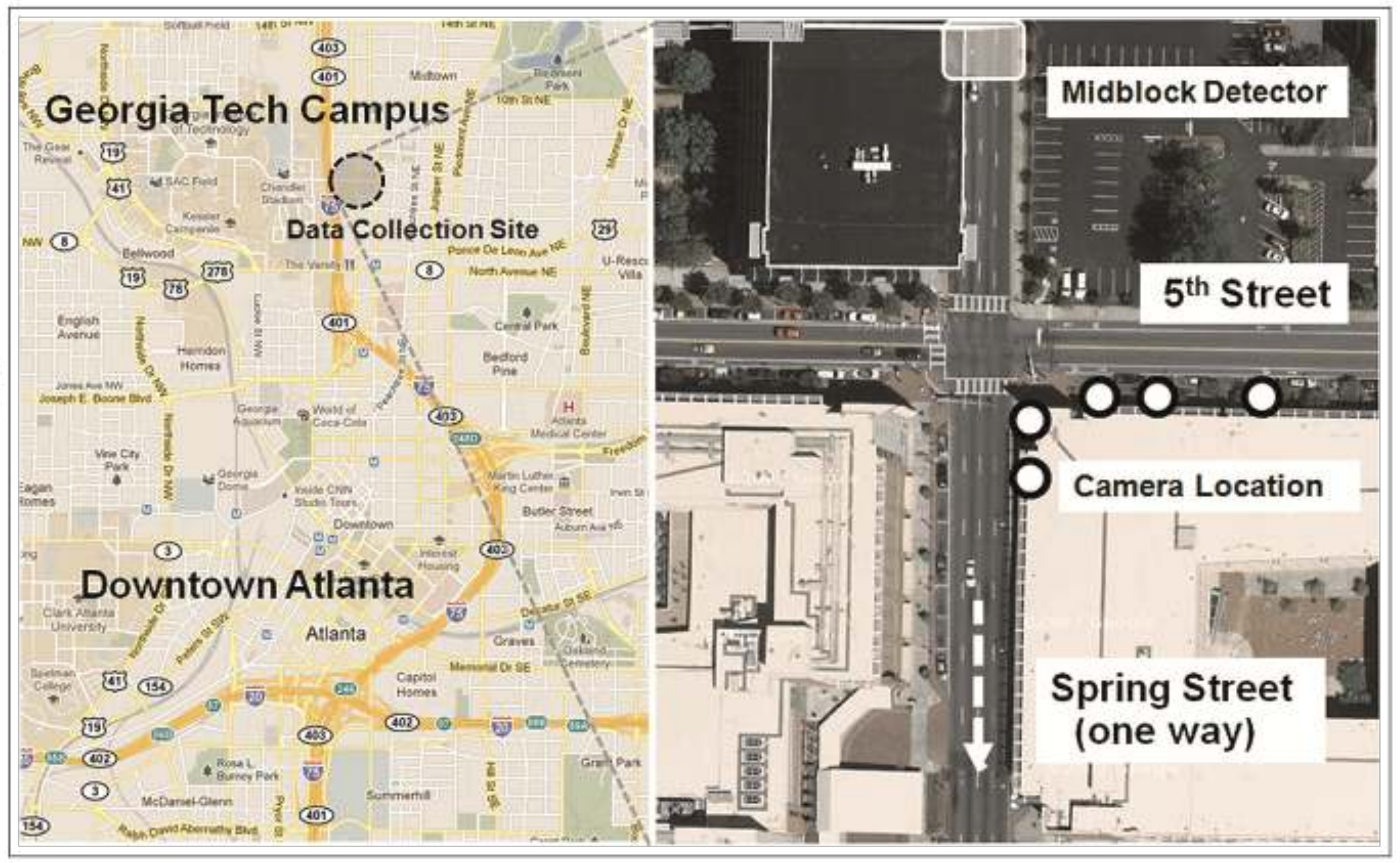
Wonho Suh[1], Dwayne Henclewood[2], Aaron Greenwood[1], Angshuman Guin[1], Randall Guensler[1], Michael P Hunter[1] and Richard Fujimoto[3]

Abstract

Microscopic traffic simulation tools are increasingly being employed as an integral part of modeling vehicula pedestrian activity. However, the complexity of pedestrians' behaviors and their interactions with the vari nents of the traffic network is commonly under-represented in simulation models, resulting in potentially mis

nd traffic data as a part of a National Science Foundation

pedestrian stepped from this zone into the crosswalk (see

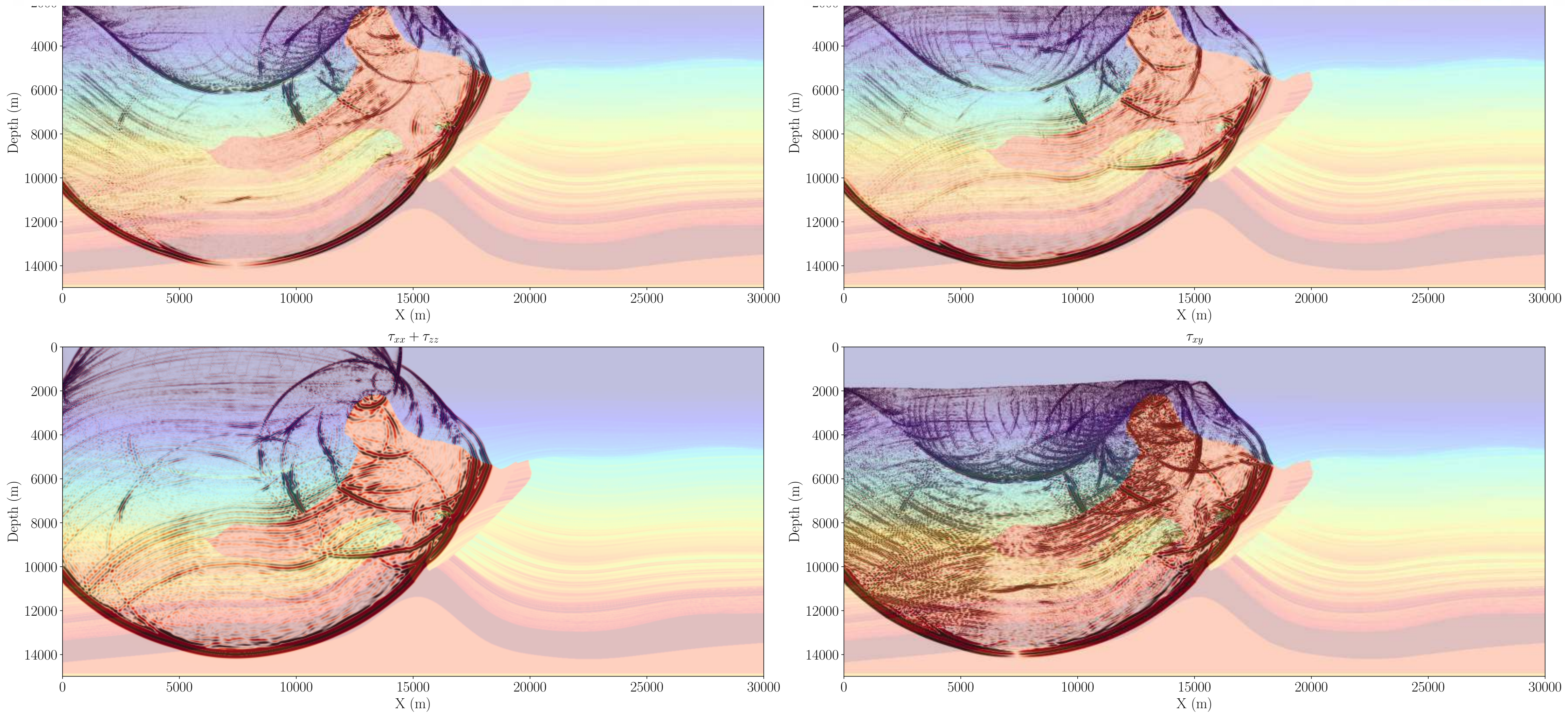


Follow along at bit.ly/GTCSE2020

Felix J. **Herrmann**

GRA Eminent Scholar, Chair in Energy
2019 Distinguished Lecturer SEG
https://slim.gatech.edu/

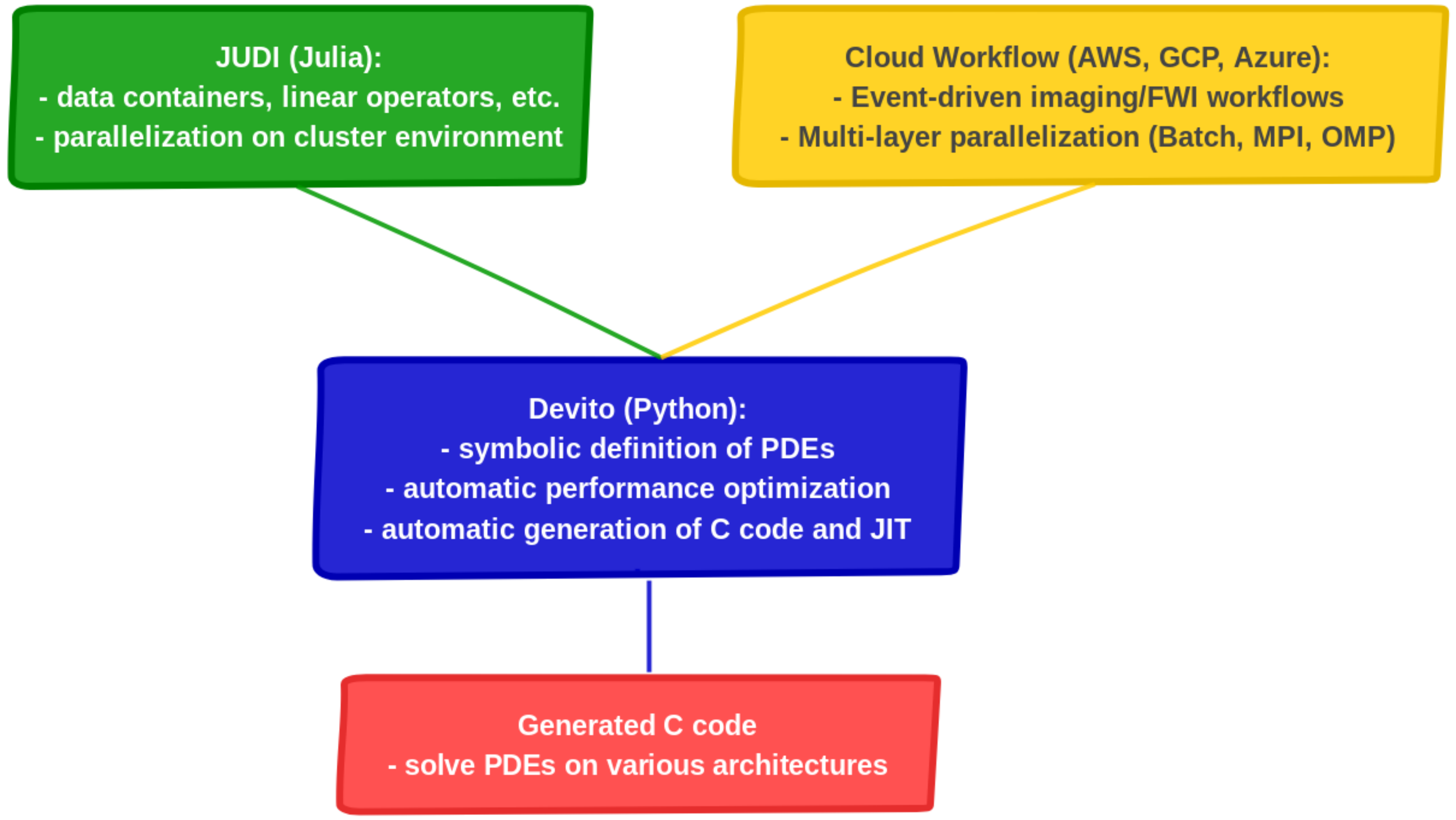
Seismic Laboratory for Imaging and Modeling



ML & HPC in seismic imaging

Using DSLs, serverless Cloud Computing, Machine Learning, and Compressive Sensing to tackle large-scale wave-based inversion problems.

www.devitoproject.org   github.com/slimgroup/JUDI.jl

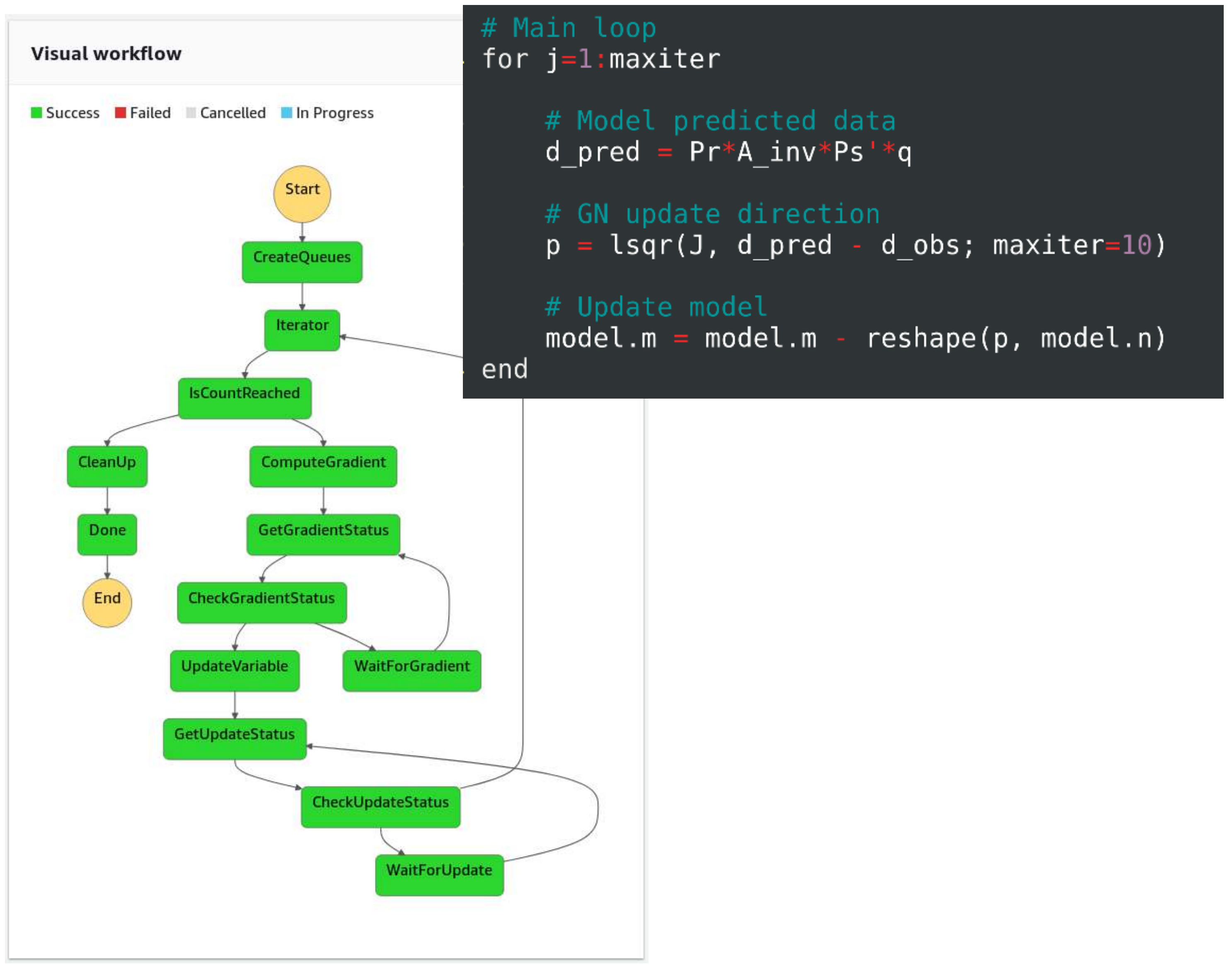


JUDI (Julia):
- data containers, linear operators, etc.
- parallelization on cluster environment

Cloud Workflow (AWS, GCP, Azure):
- Event-driven imaging/FWI workflows
- Multi-layer parallelization (Batch, MPI, OMP)

Devito (Python):
- symbolic definition of PDEs
- automatic performance optimization
- automatic generation of C code and JIT

Generated C code
- solve PDEs on various architectures



```
# Main loop
for j=1:maxiter

    # Model predicted data
    d_pred = Pr*A_inv*Ps'*q

    # GN update direction
    p = lsqr(J, d_pred - d_obs; maxiter=10)

    # Update model
    model.m = model.m - reshape(p, model.n)
end
```

arxiv.org/pdf/1807.03032.pdf

Architecture and performance of Devito, a system for automated stencil computation

FABIO LUPORINI, Imperial College London
MATHIAS LOUBOUTIN, Georgia Institute of Technology
MICHAEL LANGE, European Centre for Medium-Range Weather Forecasts
NAVJOT KUKREJA, Imperial College London
PHILIPP WITTE, Georgia Institute of Technology
JAN HÜCKELHEIM, Imperial College London
CHARLES YOUNT, Intel Corporation
PAUL H. J. KELLY, Imperial College London
FELIX J. HERRMANN, Georgia Institute of Technology
GERARD J. GORMAN, Imperial College London

Follow along at bit.ly/GTCSE2020

Tobin **Isaac**

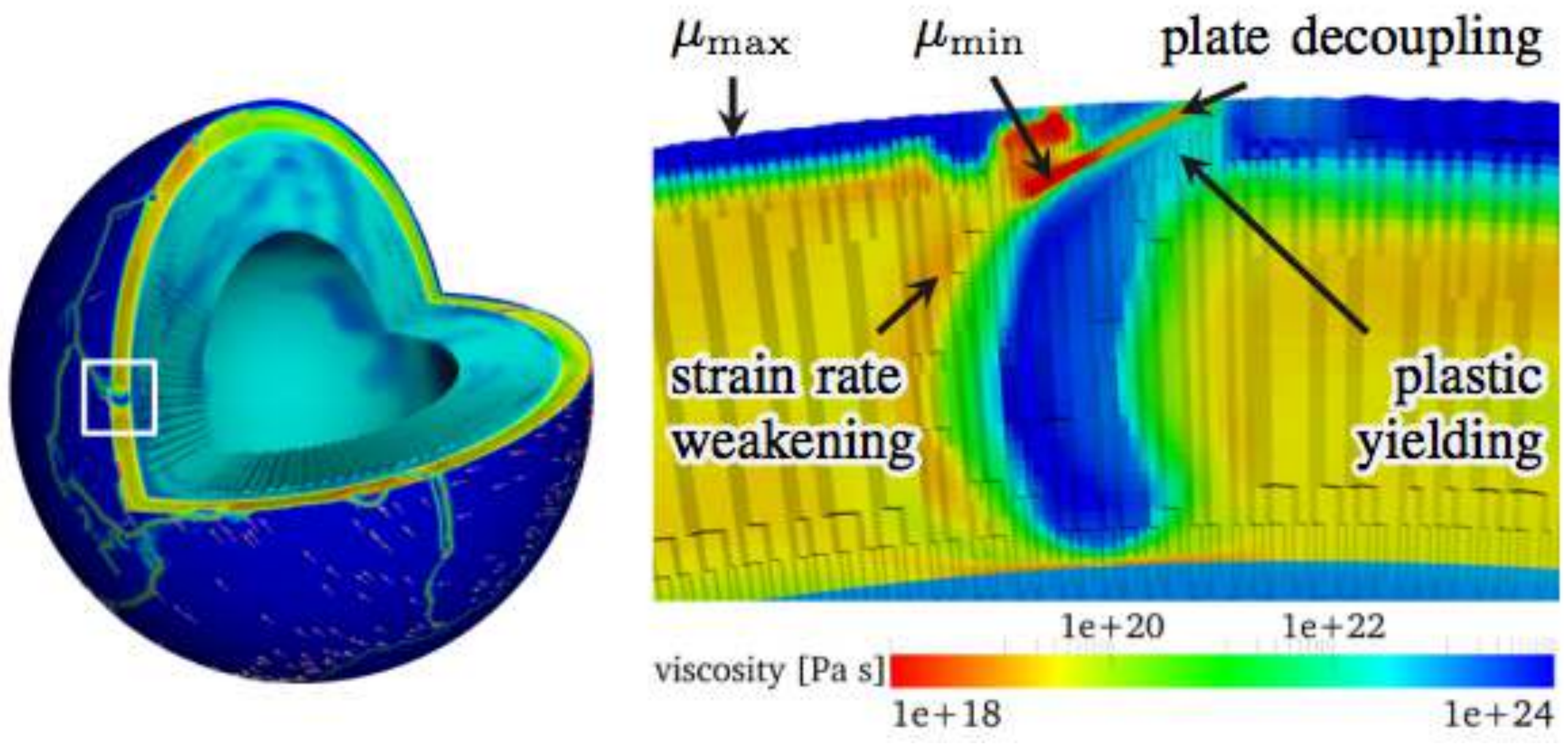
PhD (2015), UT Austin, Awarded 2016 SIAM/Supercomputing Early Career Prize

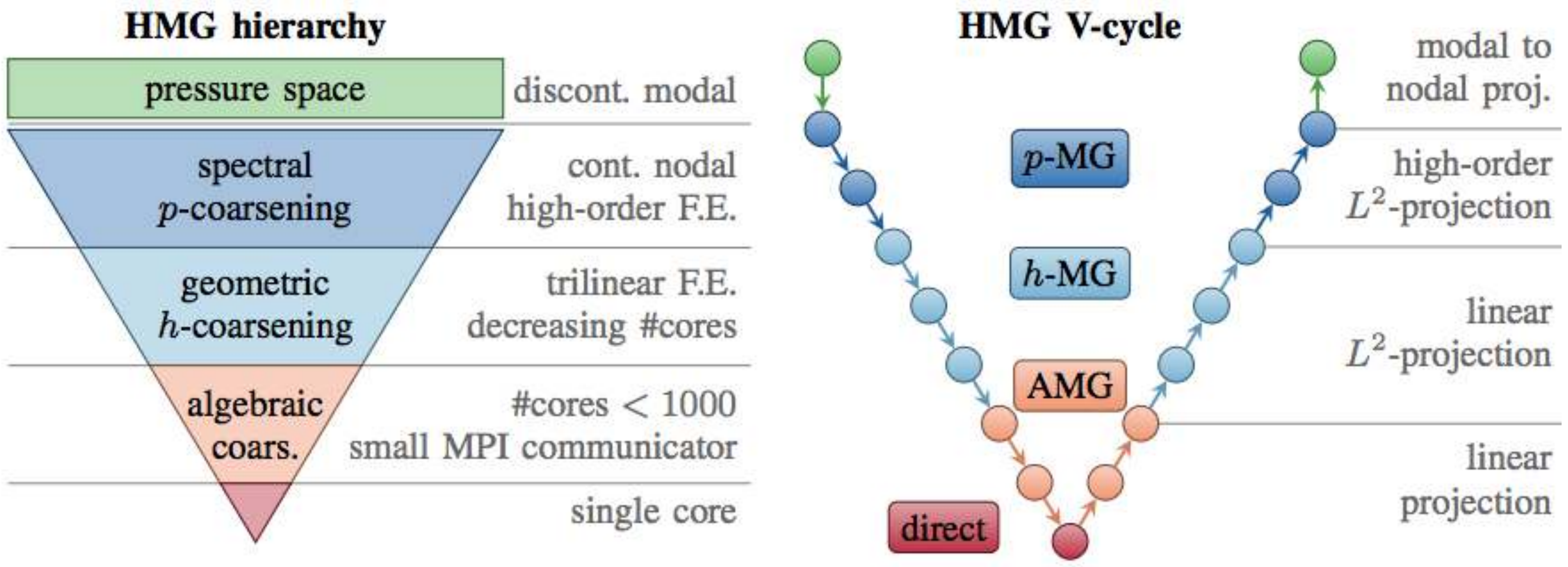
Applied math & numerical analysis

Developing Bayesian inversion techniques for models governed by large, complex systems; atmospheric modeling and weather prediction

ACM Gordon Bell Prize @SC 2015

An Extreme-Scale Implicit Solver for Complex PDEs: Highly Heterogeneous Flow in Earth's Mantle





Follow along at bit.ly/GTCSE2020

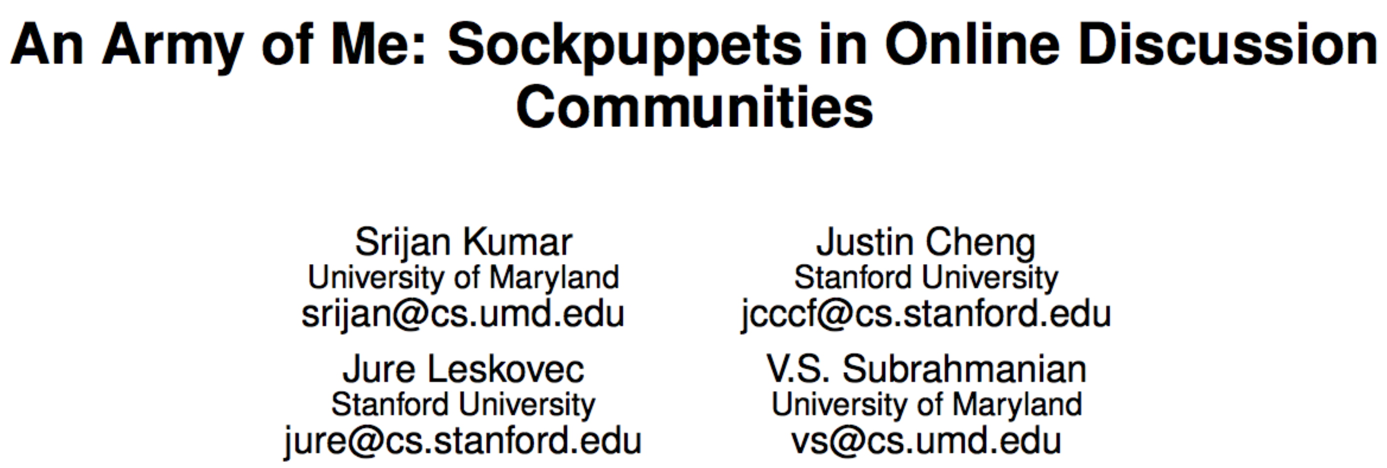
Srijan Kumar

Assistant Professor, PhD (2017), Stanford postdoc
Creating a safer web for everyone

Data Science to Improve Web Safety, Integrity, and Well-being

Inventing network science, user modeling, and machine learning methods to model human behavior and improve web and social media. We develop actionable insights for enable efficient decision making. Our models are used at Wikipedia, Facebook, Twitter, and Flipkart.

Web Integrity



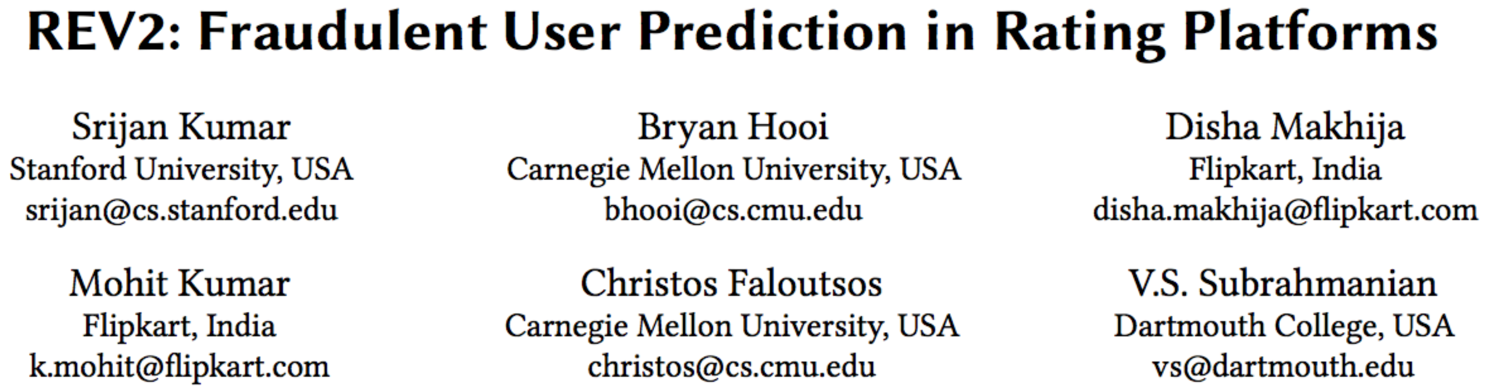
An Army of Me: Sockpuppets in Online Discussion Communities

Srijan Kumar
University of Maryland
srijan@cs.umd.edu

Justin Cheng
Stanford University
jcccf@cs.stanford.edu

Jure Leskovec
Stanford University
jure@cs.stanford.edu

V.S. Subrahmanian
University of Maryland
vs@cs.umd.edu



REV2: Fraudulent User Prediction in Rating Platforms

Srijan Kumar
Stanford University, USA
srijan@cs.stanford.edu

Bryan Hooi
Carnegie Mellon University, USA
bhooi@cs.cmu.edu

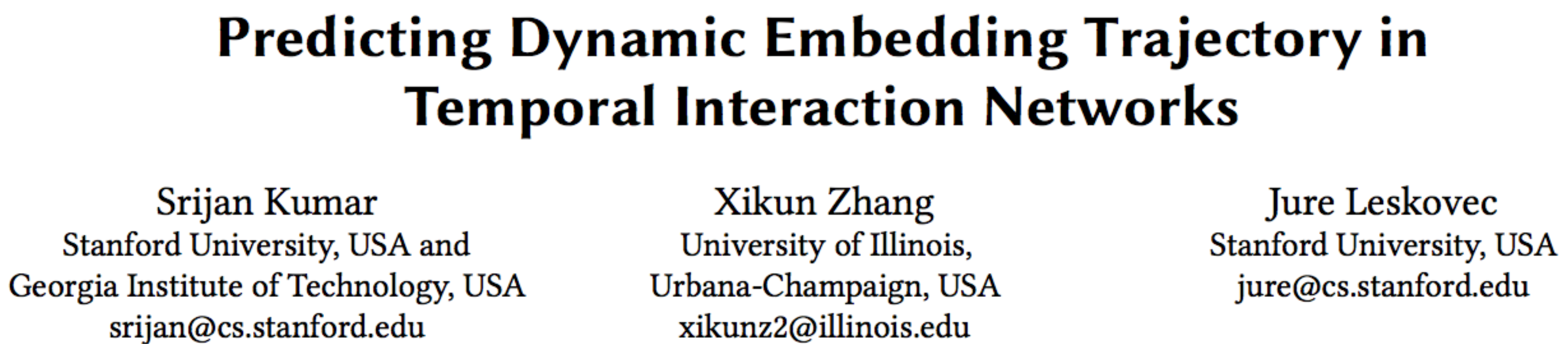
Disha Makhija
Flipkart, India
disha.makhija@flipkart.com

Mohit Kumar
Flipkart, India
k.mohit@flipkart.com

Christos Faloutsos
Carnegie Mellon University, USA
christos@cs.cmu.edu

V.S. Subrahmanian
Dartmouth College, USA
vs@dartmouth.edu

Network Modeling

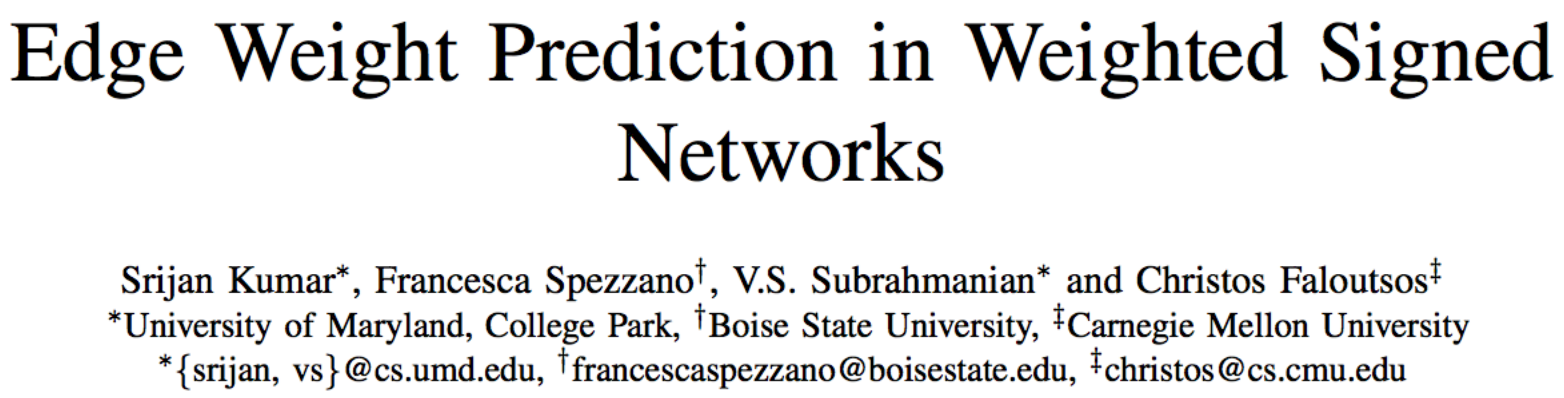


Predicting Dynamic Embedding Trajectory in Temporal Interaction Networks

Srijan Kumar
Stanford University, USA and
Georgia Institute of Technology, USA
srijan@cs.stanford.edu

Xikun Zhang
University of Illinois,
Urbana-Champaign, USA
xikunz2@illinois.edu

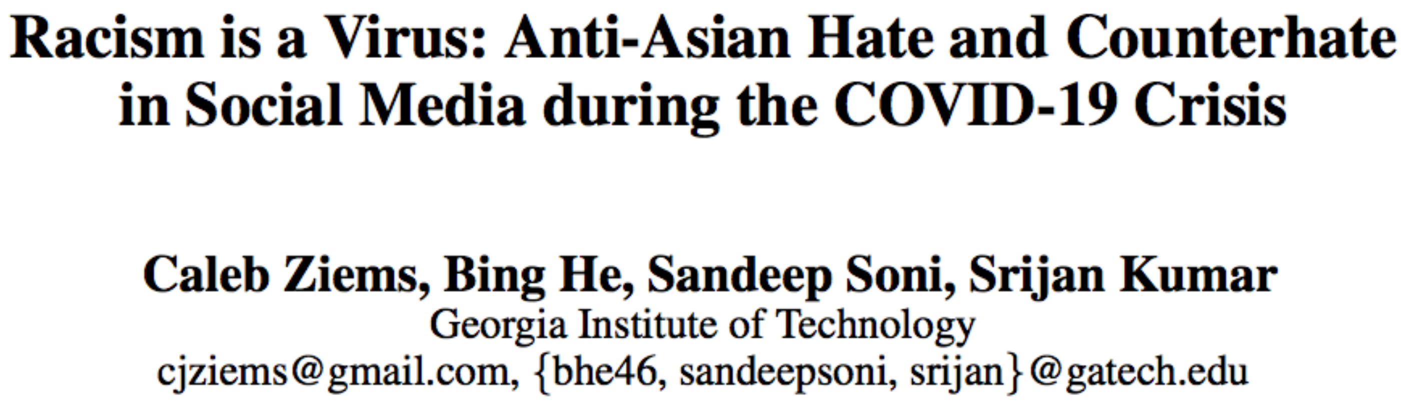
Jure Leskovec
Stanford University, USA
jure@cs.stanford.edu



Edge Weight Prediction in Weighted Signed Networks

Srijan Kumar*, Francesca Spezzano†, V.S. Subrahmanian* and Christos Faloutsos‡
*University of Maryland, College Park, †Boise State University, ‡Carnegie Mellon University
*{srijan, vs}@cs.umd.edu, †francescaspezzano@boisestate.edu, ‡christos@cs.cmu.edu

Health & Safety



Racism is a Virus: Anti-Asian Hate and Counterhate in Social Media during the COVID-19 Crisis

Caleb Ziems, Bing He, Sandeep Soni, Srijan Kumar
Georgia Institute of Technology
cjziems@gmail.com, {bhe46, sandeepsoni, srijan}@gatech.edu

National Science Foundation
WHERE DISCOVERIES BEGIN

Award Abstract #2027689

RAPID: Tackling the Psychological Impact of the COVID-19 Crisis

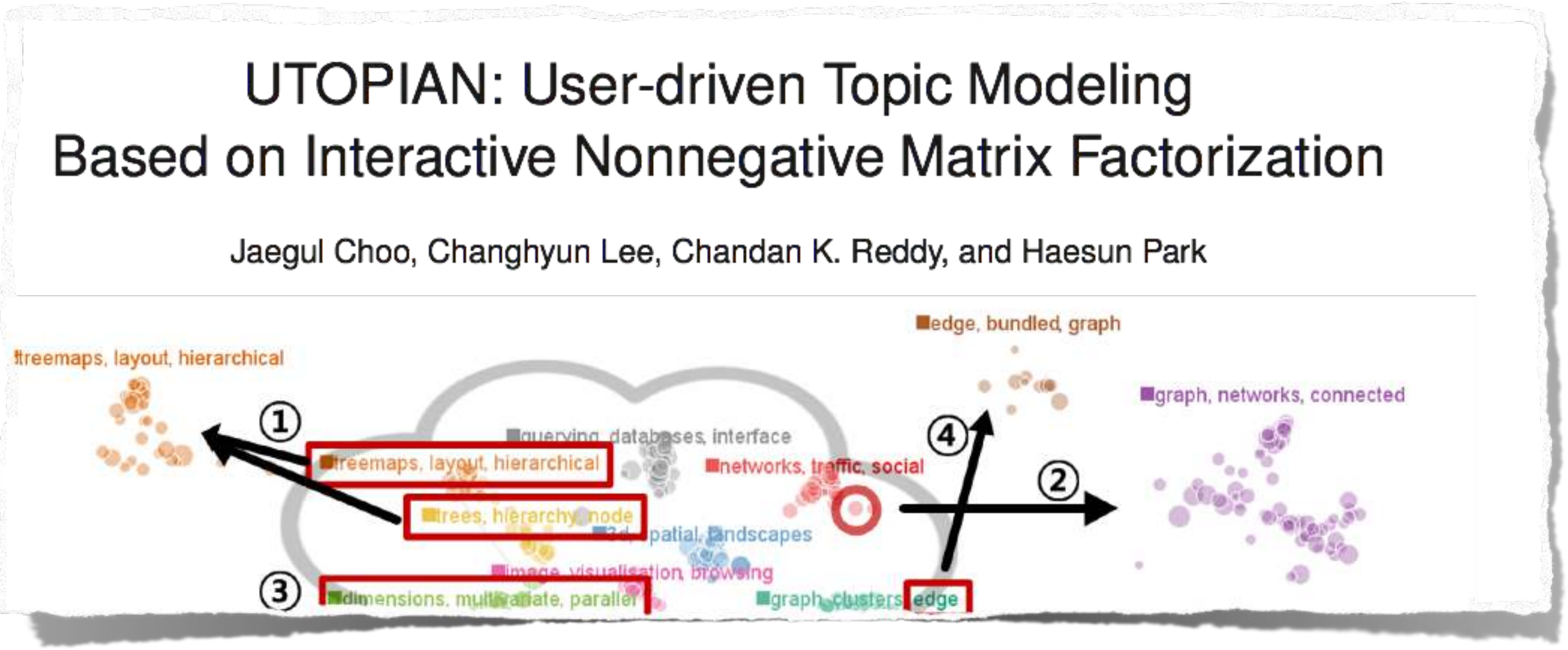
Follow along at bit.ly/GTCSE2020

Haesun **Park**

CSE Chair & Regents' Professor
SIAM & IEEE Fellow, & pianist!

Data and visual analytics

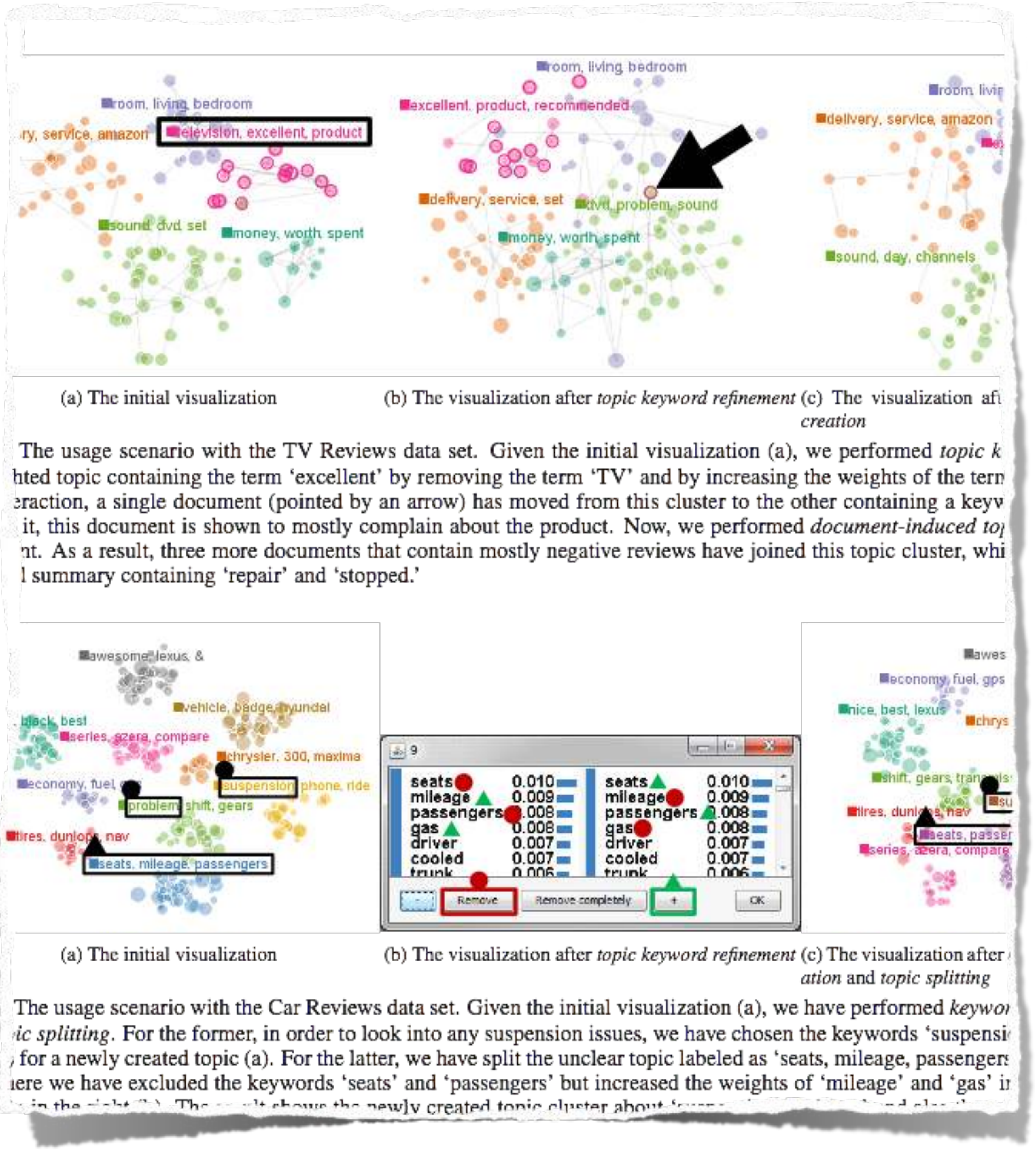
Bringing numerical linear algebra and optimization to bear new data analysis and mining, for missing value estimation, nonnegative matrix factorization, tensor computations, …



UTOPIAN: User-driven Topic Modeling Based on Interactive Nonnegative Matrix Factorization

Jaegul Choo, Changhyun Lee, Chandan K. Reddy, and Haesun Park

treemaps, layout, hierarchical
edge, bundled, graph
graph, networks, connected
①
querying, databases, interface
④
treemaps, layout, hierarchical
networks, traffic, social
②
trees, hierarchy, node
spatial, landscapes
image, visualization, browsing
③
dimensions, multivariate, parallel
graph, clusters, edge



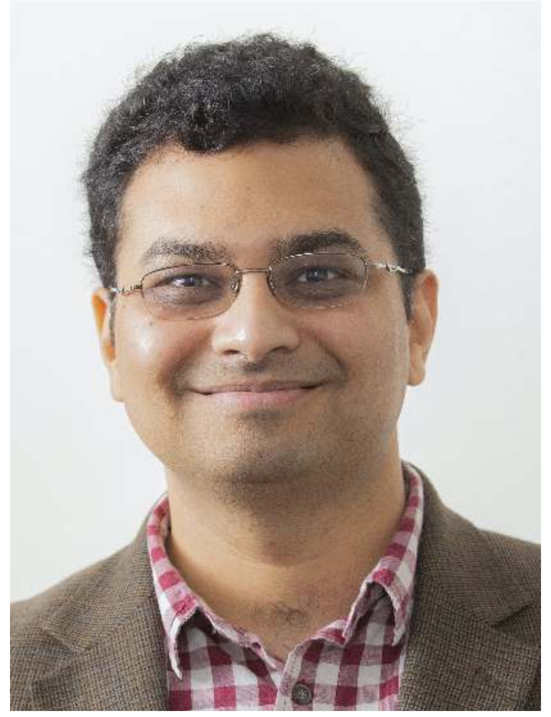
(a) The initial visualization (b) The visualization after *topic keyword refinement* (c) The visualization aft *creation*

The usage scenario with the TV Reviews data set. Given the initial visualization (a), we performed *topic k* hted topic containing the term 'excellent' by removing the term 'TV' and by increasing the weights of the tern eraction, a single document (pointed by an arrow) has moved from this cluster to the other containing a keyv it, this document is shown to mostly complain about the product. Now, we performed *document-induced to* nt. As a result, three more documents that contain mostly negative reviews have joined this topic cluster, whi l summary containing 'repair' and 'stopped.'

(a) The initial visualization (b) The visualization after *topic keyword refinement* (c) The visualization after ation and *topic splitting*

The usage scenario with the Car Reviews data set. Given the initial visualization (a), we have performed *keywor ic splitting*. For the former, in order to look into any suspension issues, we have chosen the keywords 'suspensi , for a newly created topic (a). For the latter, we have split the unclear topic labeled as 'seats, mileage, passengers iere we have excluded the keywords 'seats' and 'passengers' but increased the weights of 'mileage' and 'gas' ir in the right (b). The result shows the newly created topic cluster about

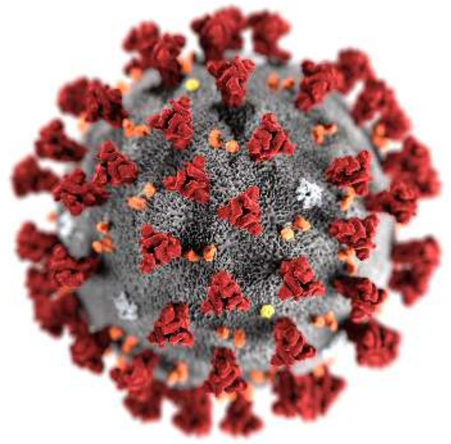
Follow along at bit.ly/GTCSE2020



B. Aditya **Prakash**

Associate Professor, PhD (2012) CMU,
loves 'data' and plays the Tabla!

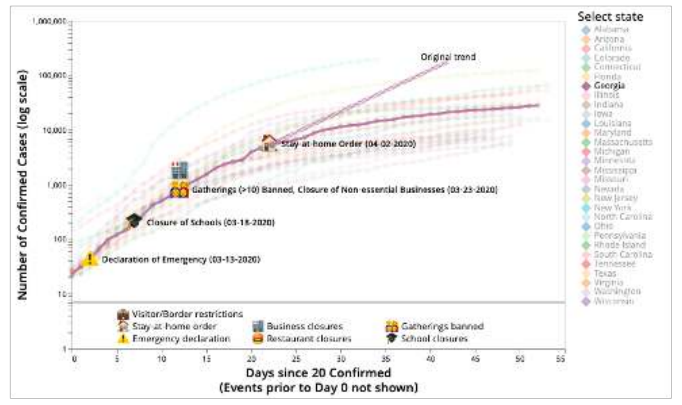
Aditya Lab



Other COVID-19 Response Activities

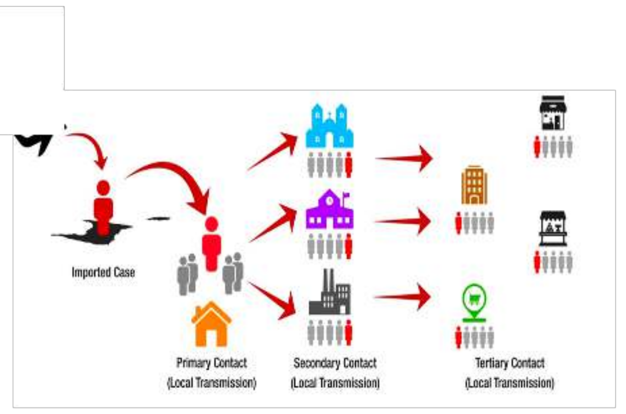
Data science and machine learning

Emphasis on solving big-data problems in networks and sequences, motivated from high-impact applications such as epidemiology, public health, urban computing, the web and security.



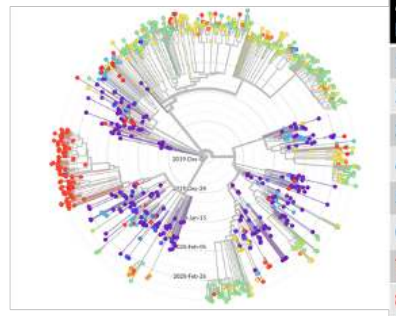
Visualizing impact of nonpharmaceutical interventions

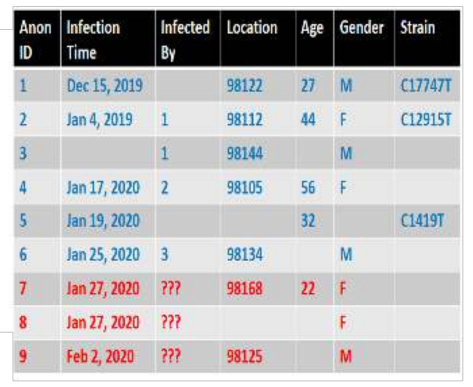
Berkeley UNIVERSITY OF CALIFORNIA



Imported Case
Primary Contact (Local Transmission)
Secondary Contact (Local Transmission)
Tertiary Contact (Local Transmission)

Mobility Analysis





Anon ID	Infection Time	Infected By	Location	Age	Gender	Strain
1	Dec 15, 2019		98122	27	M	C17747T
2	Jan 4, 2019	1	98112	44	F	C12915T
3		1	98144		M	
4	Jan 17, 2020	2	98105	56	F	
5	Jan 19, 2020			32		C1419T
6	Jan 25, 2020	3	98134		M	
7	Jan 27, 2020	???	98168	22	F	
8	Jan 27, 2020	???			F	
9	Feb 2, 2020	???	98125		M	

Adaptive surveillance

UNIVERSITY of VIRGINIA

epiDAMIK @KDD2020

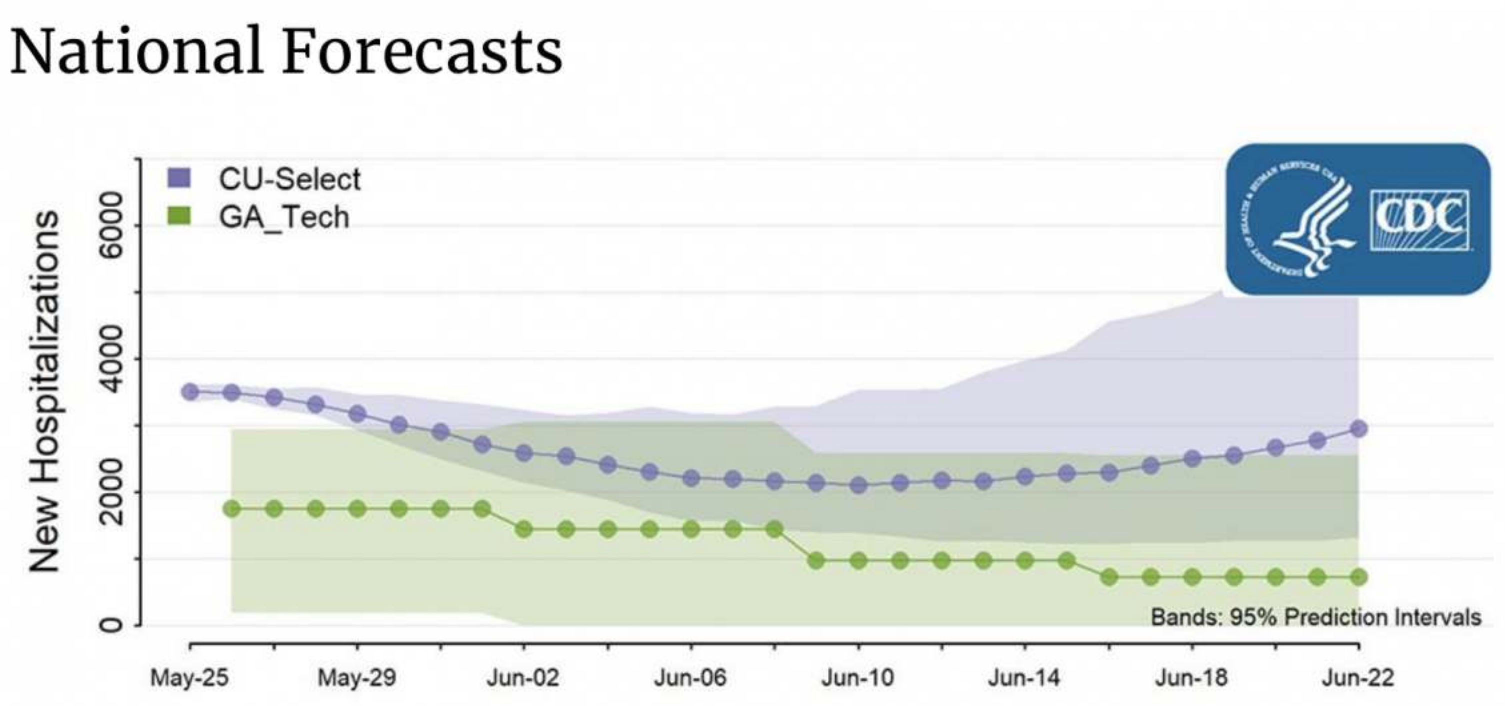
HARVARD UNIVERSITY Google Georgetown University UNIVERSITY of VIRGINIA THE UNIVERSITY OF IOWA

Data Science+Epi workshop @SIGKDD2020

Team Using Deep Learning to Forecast Pandemic in the U.S.

Monday, June 1, 2020

FiveThirtyEight



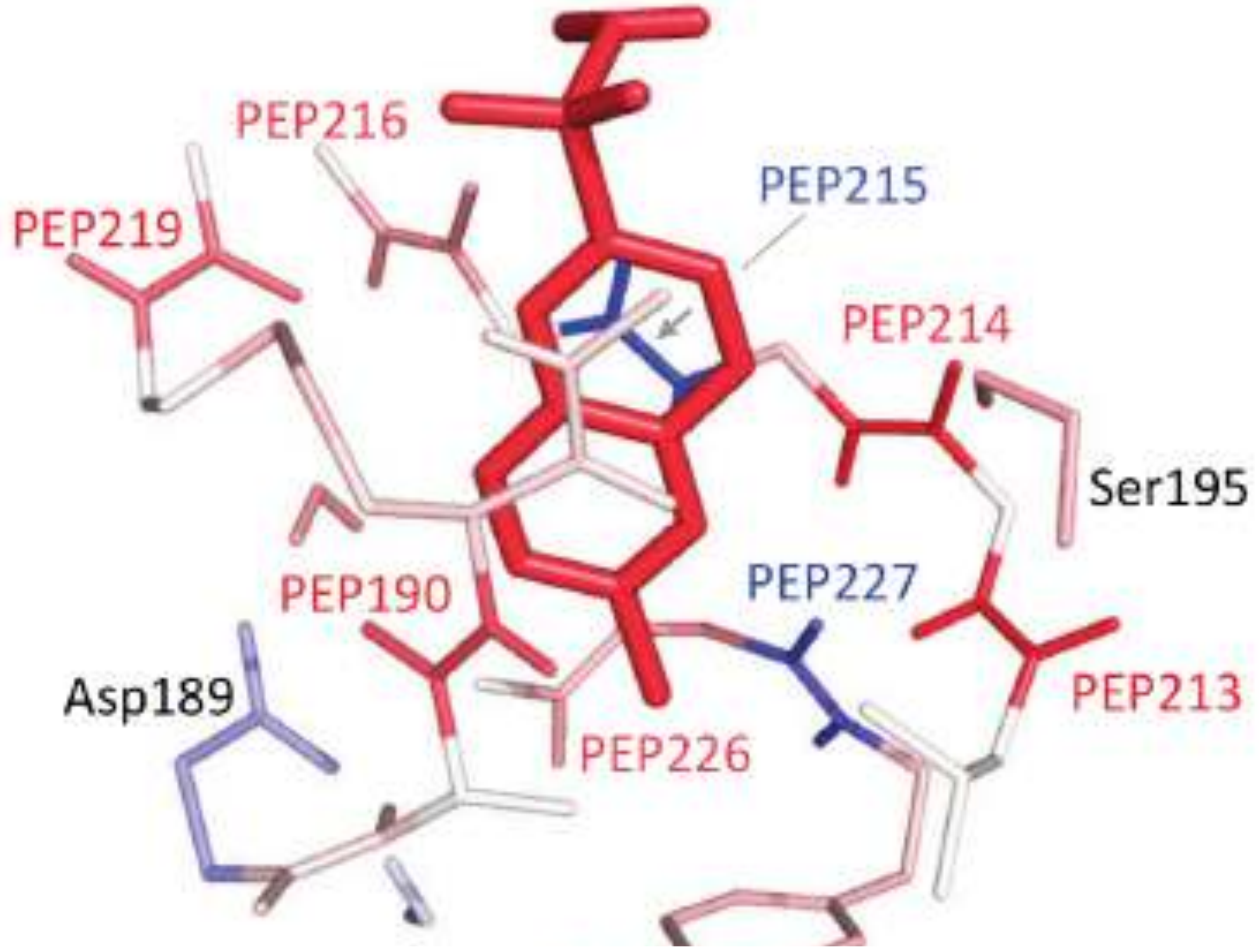
Teaching **new class** in Fall 2020! CSE 8803 EPI: Data Science for Epidemiology

David Sherrill

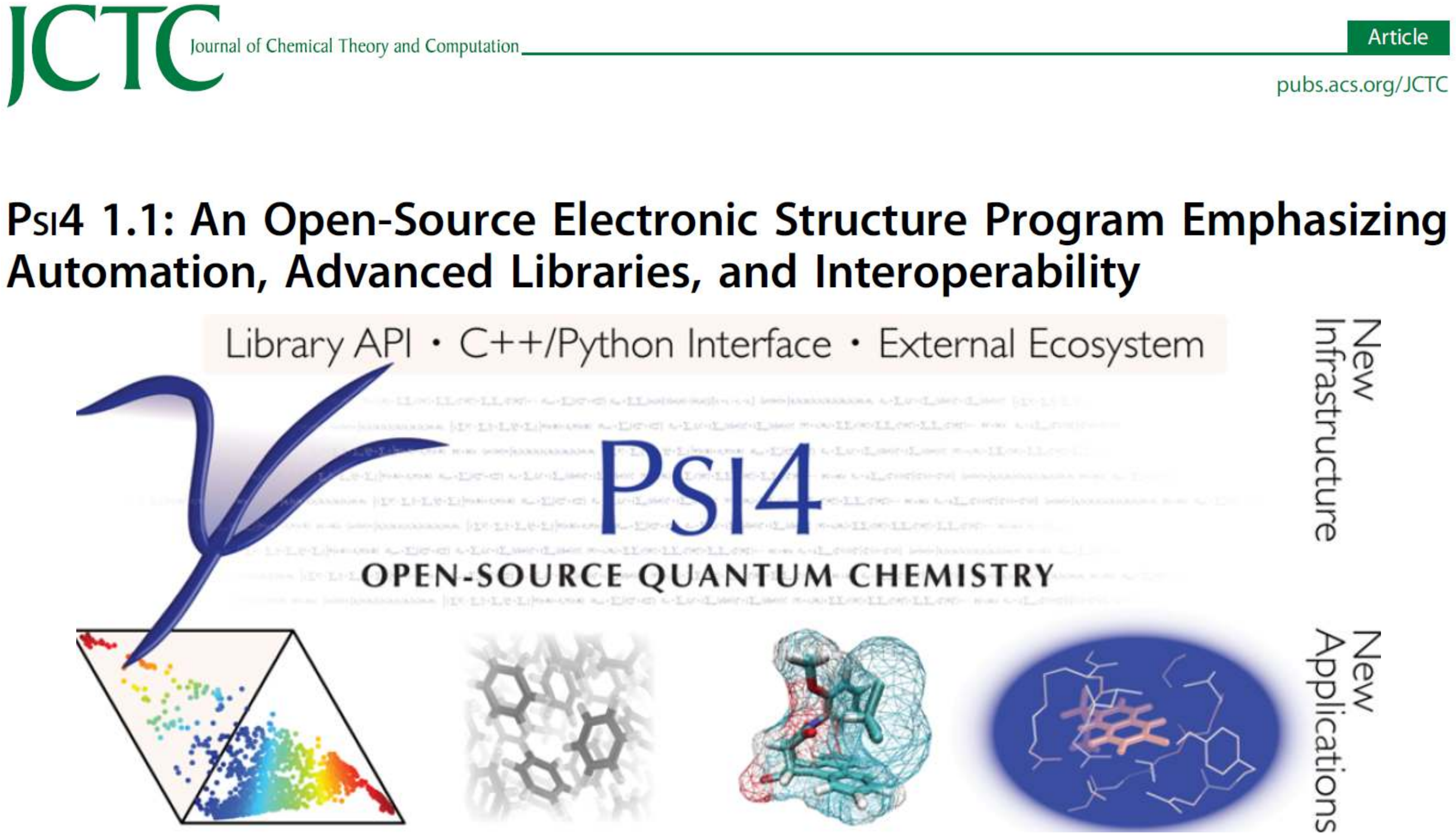
ACS, APS, AAAS Fellow
Theoretical and Computational Chemist

Models and algorithms for quantum chemistry

New approximations to the Schrödinger Equation
New algorithms for efficient implementation
The most popular open-source quantum chemistry code, Psi4



Methods for computer-aided drug design



JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability

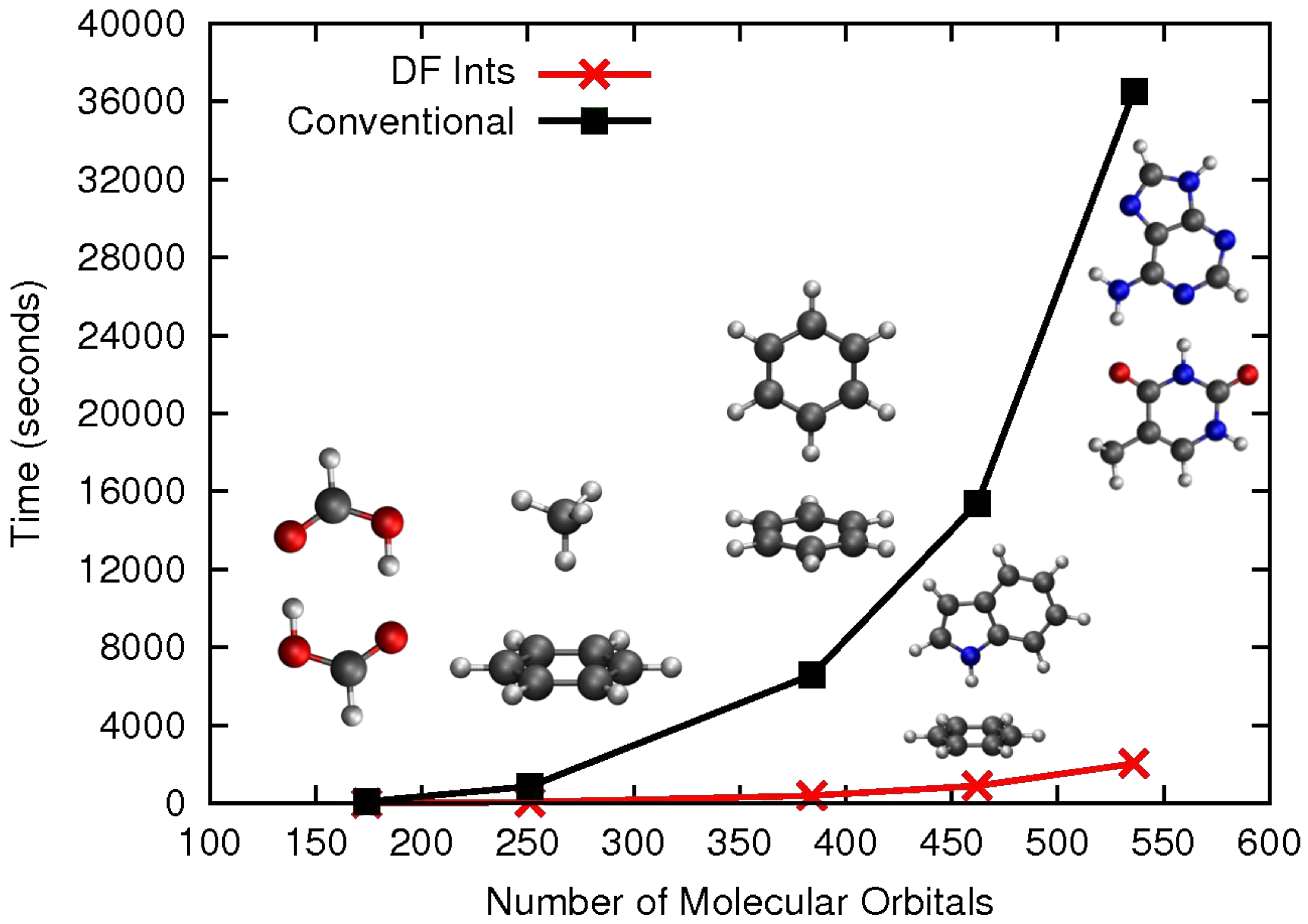
Library API • C++/Python Interface • External Ecosystem

Psi4

OPEN-SOURCE QUANTUM CHEMISTRY

New Infrastructure

New Applications



Follow along at bit.ly/GTCSE2020

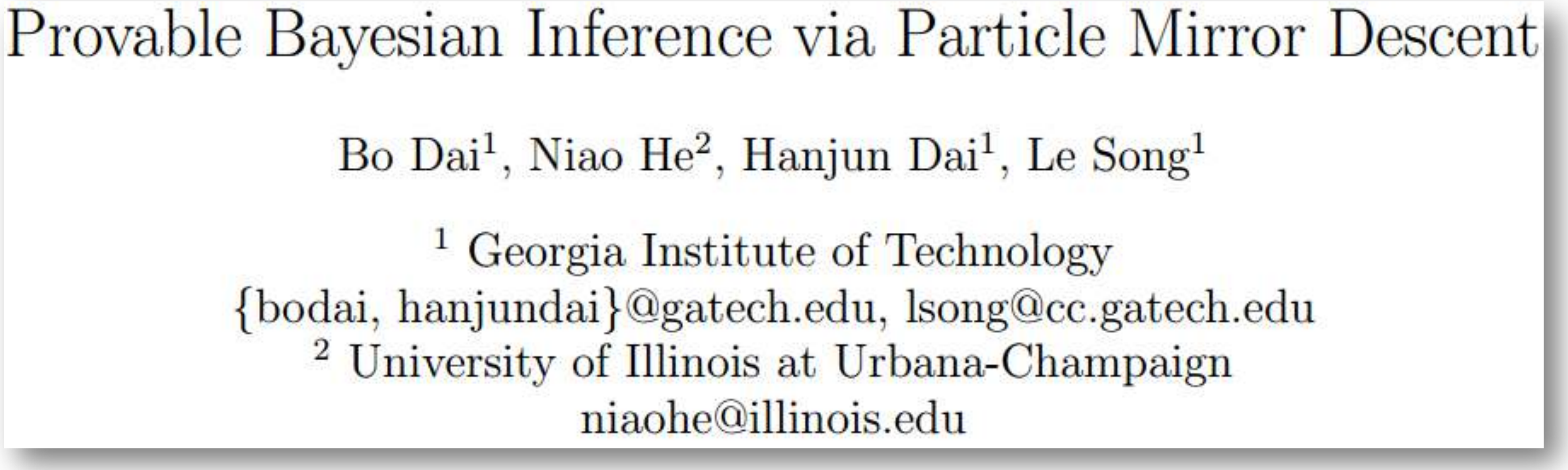
Le **Song**

Razor sharp and having unbounded energy

Statistical machine learning

Nonparametric kernel methods, graphical models, time series, distributed learning, with applications in the analysis of text, images, networks, biological systems, & social media.

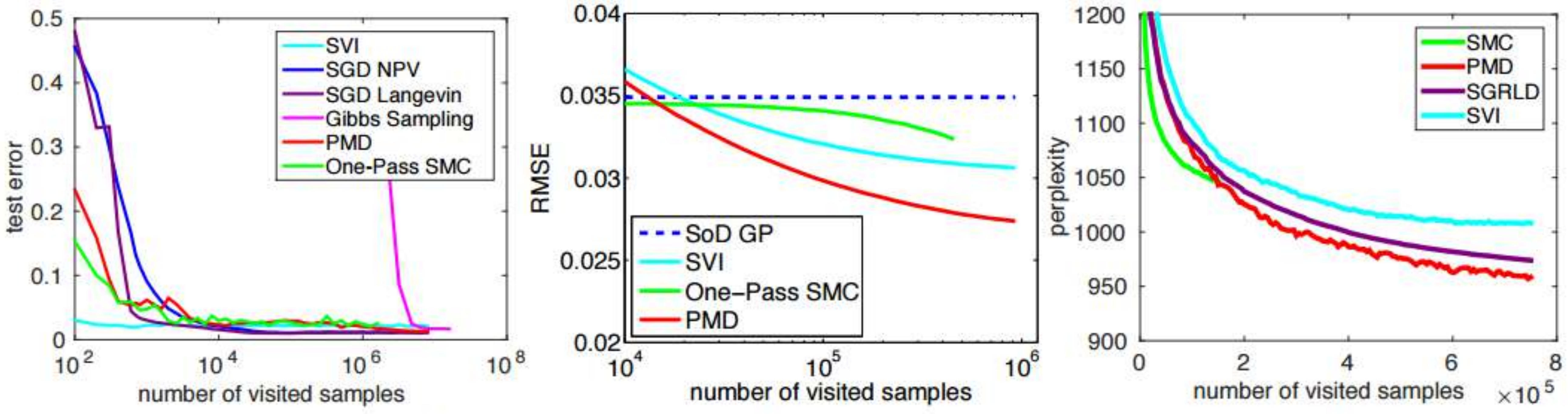
Best Student Paper @AISTATS 2016



Provable Bayesian Inference via Particle Mirror Descent

Bo Dai[1], Niao He[2], Hanjun Dai[1], Le Song[1]

[1] Georgia Institute of Technology
{bodai, hanjundai}@gatech.edu, lsong@cc.gatech.edu
[2] University of Illinois at Urbana-Champaign
niaohe@illinois.edu



(1) Logistic regression on MNIST (2) Sparse GP on music data (3) LDA on wikipedia data

Figure 2: Experimental results on several different models for real-world datasets.

Follow along at bit.ly/GTCSE2020

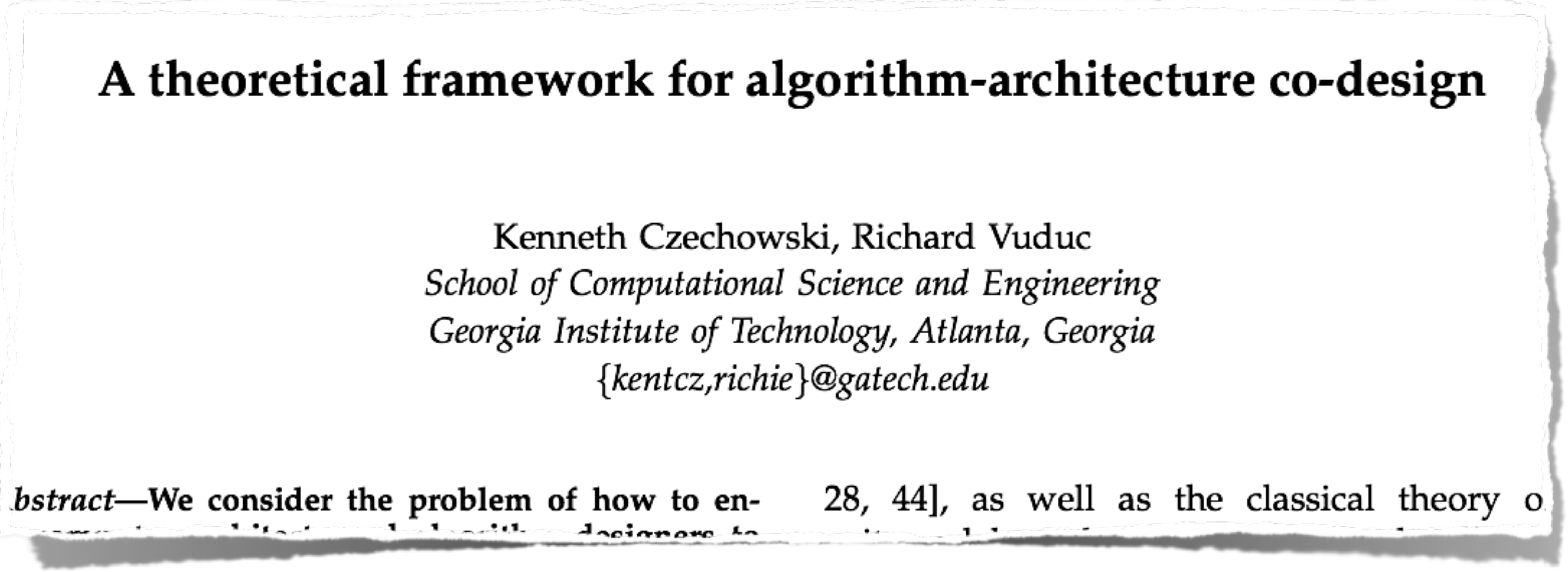
Rich **Vuduc**

Professor, HPC Guru

hpcgarage

Time, energy, power, & reliability in HPC

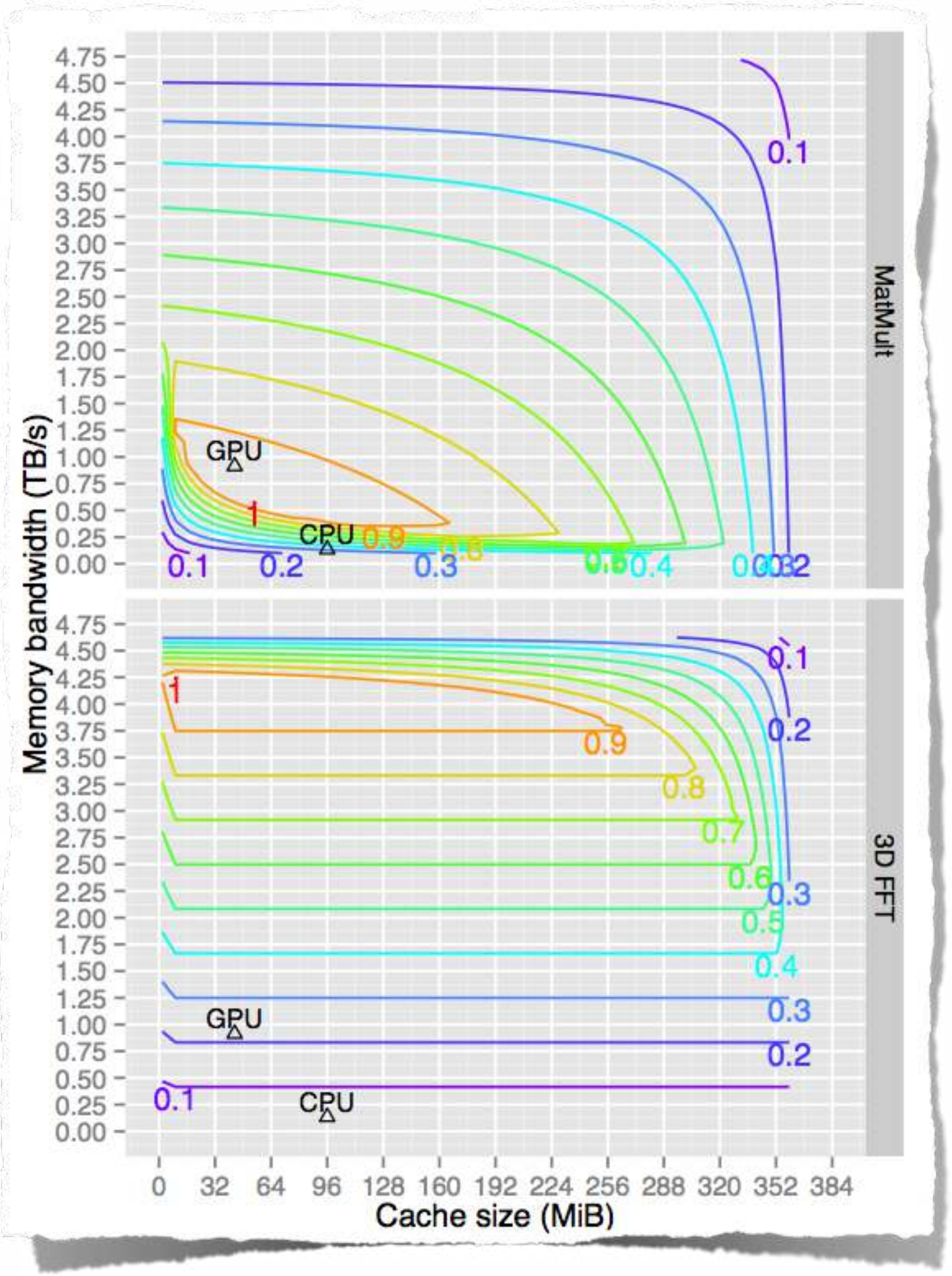
Developing new models for analyzing algorithms, software, and machines under limits on seconds, Joules, and Watts.



A theoretical framework for algorithm-architecture co-design

Kenneth Czechowski, Richard Vuduc
School of Computational Science and Engineering
Georgia Institute of Technology, Atlanta, Georgia
{kentcz,richie}@gatech.edu

bstract—We consider the problem of how to en- 28, 44], as well as the classical theory o



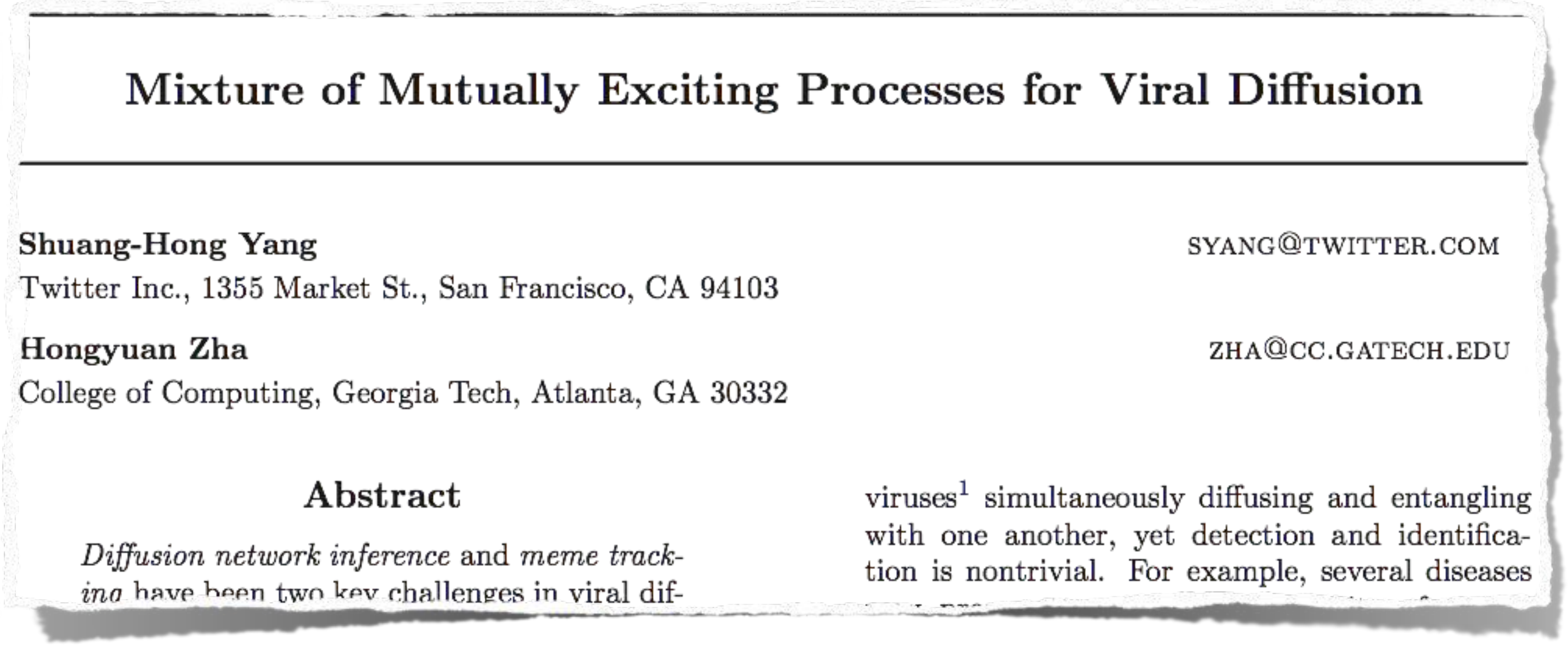
Follow along at bit.ly/GTCSE2020

Hongyuan **Zha**

Former yahoo from *Yahoo!* (or rather, Inktomi)

Computational math & machine learning

Bridging scientific computing and machine learning to solve problems in web search, text mining, and network analysis.



Mixture of Mutually Exciting Processes for Viral Diffusion

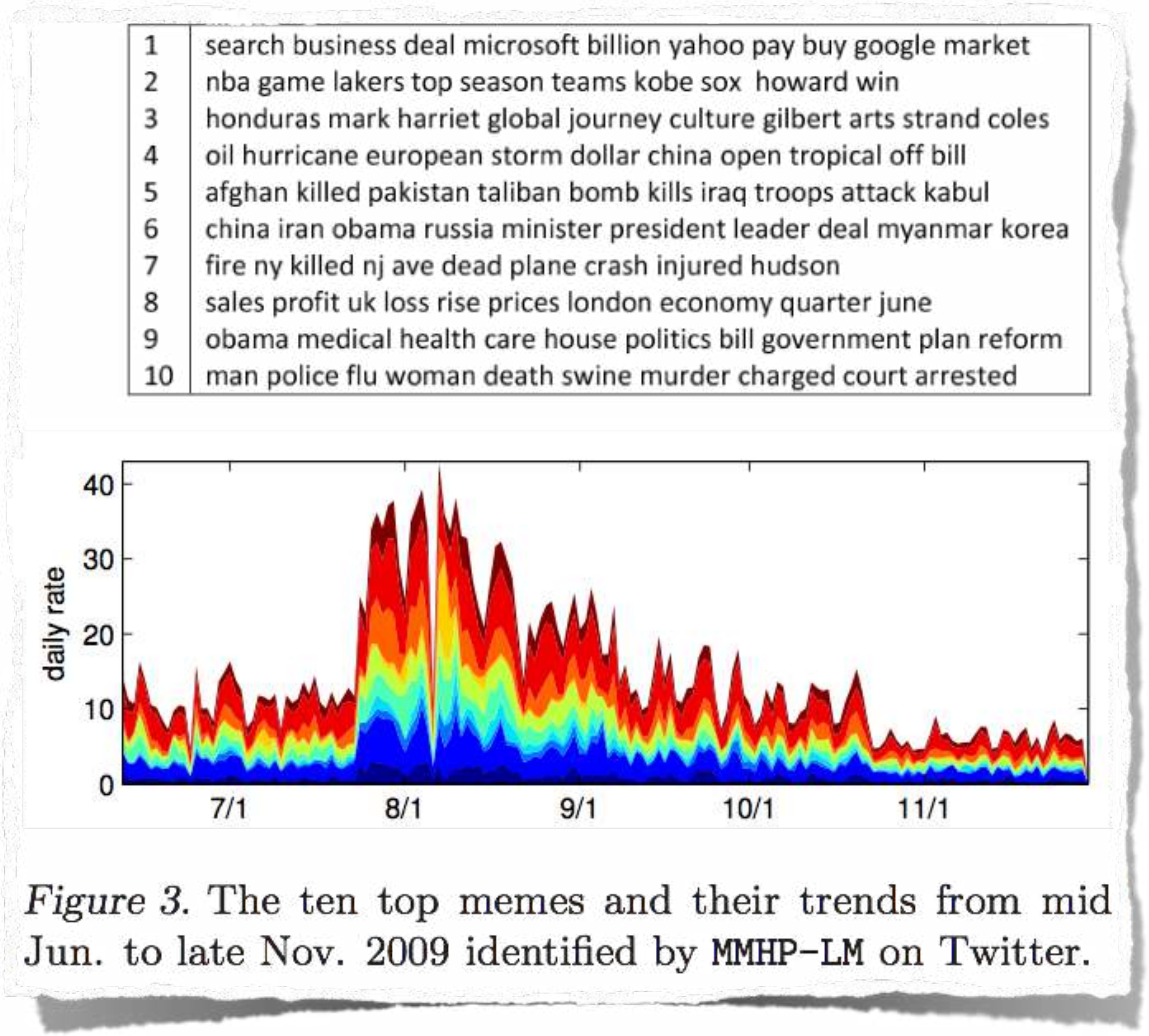
Shuang-Hong Yang SYANG@TWITTER.COM
Twitter Inc., 1355 Market St., San Francisco, CA 94103

Hongyuan Zha ZHA@CC.GATECH.EDU
College of Computing, Georgia Tech, Atlanta, GA 30332

Abstract

Diffusion network inference and *meme tracking* have been two key challenges in viral dif-

viruses[1] simultaneously diffusing and entangling with one another, yet detection and identification is nontrivial. For example, several diseases



1	search business deal microsoft billion yahoo pay buy google market
2	nba game lakers top season teams kobe sox howard win
3	honduras mark harriet global journey culture gilbert arts strand coles
4	oil hurricane european storm dollar china open tropical off bill
5	afghan killed pakistan taliban bomb kills iraq troops attack kabul
6	china iran obama russia minister president leader deal myanmar korea
7	fire ny killed nj ave dead plane crash injured hudson
8	sales profit uk loss rise prices london economy quarter june
9	obama medical health care house politics bill government plan reform
10	man police flu woman death swine murder charged court arrested

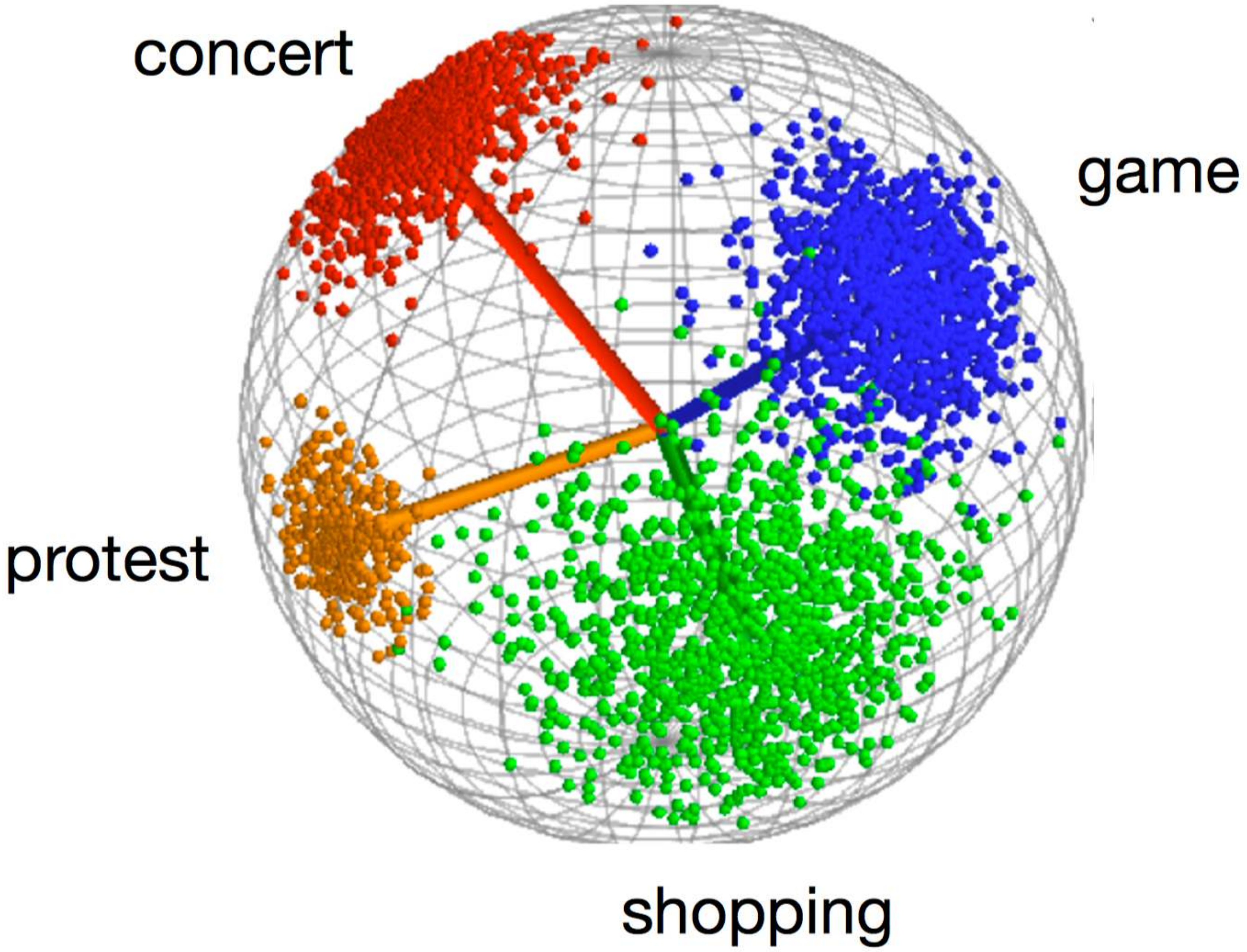
Figure 3. The ten top memes and their trends from mid Jun. to late Nov. 2009 identified by `MMHP-LM` on Twitter.

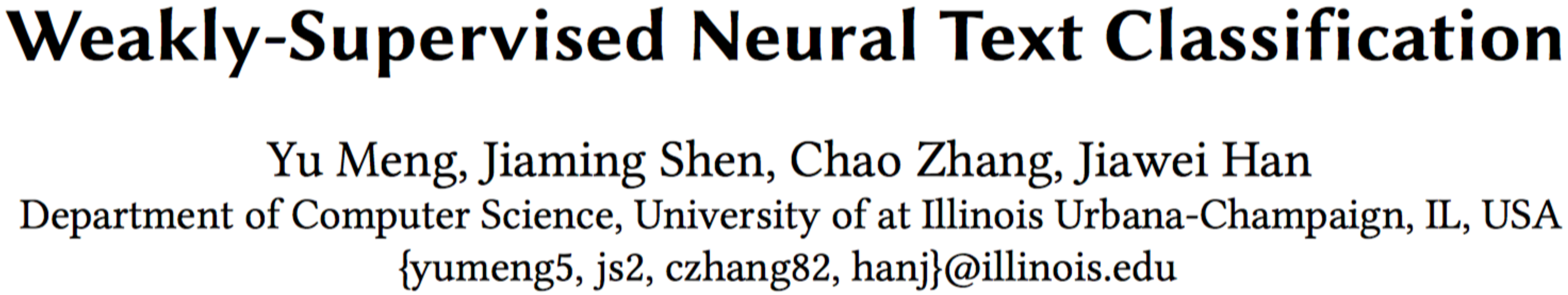
Chao **Zhang**

PhD (2018), UIUC

Data mining and machine learning

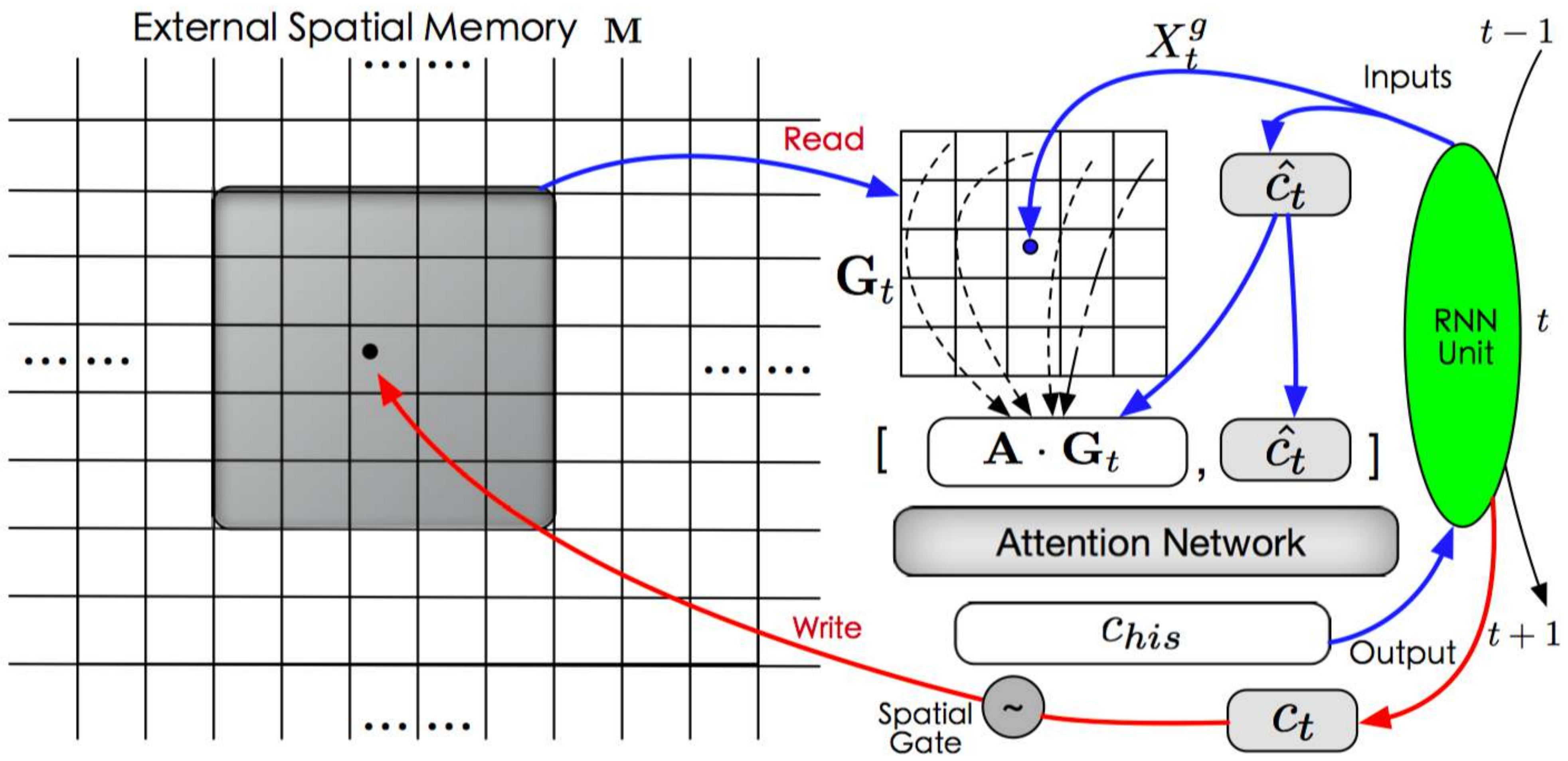
Developing label-efficient and robust machine learning algorithms for task support and decision making, with a focus on text data analysis and spatiotemporal data mining.



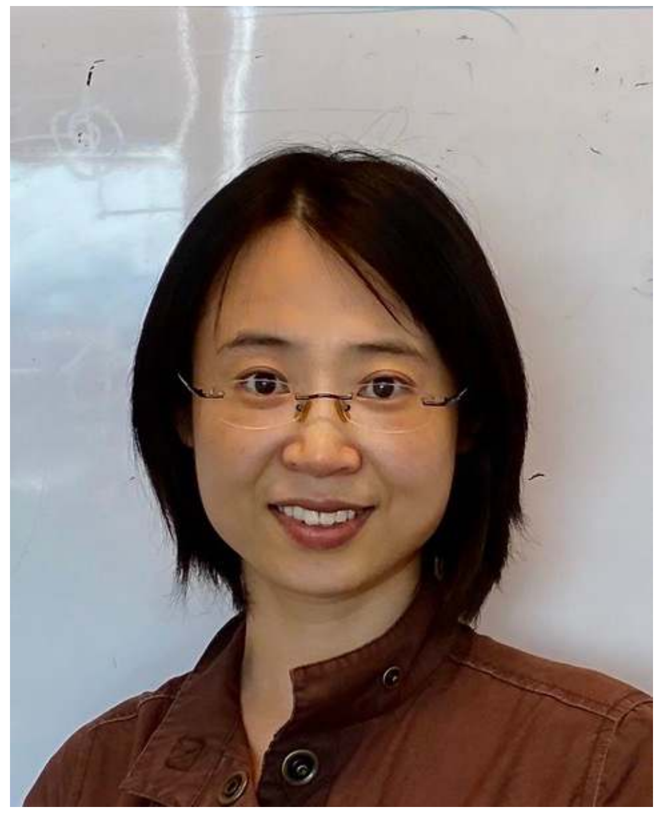


Weakly-Supervised Neural Text Classification

Yu Meng, Jiaming Shen, Chao Zhang, Jiawei Han
Department of Computer Science, University of at Illinois Urbana-Champaign, IL, USA
{yumeng5, js2, czhang82, hanj}@illinois.edu



Follow along at bit.ly/GTCSE2020

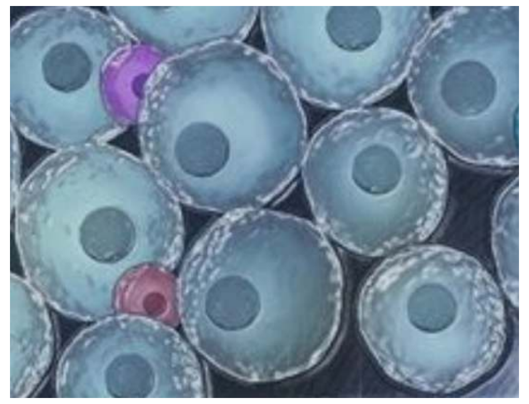


Xiuwei Zhang

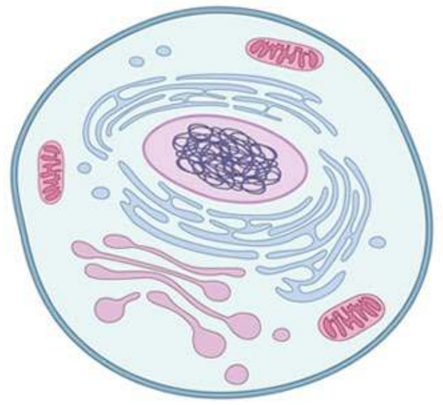
Computational Biologist

Models and algorithms for single cell data

Developing computational tools to study mechanisms in cell development and differentiation

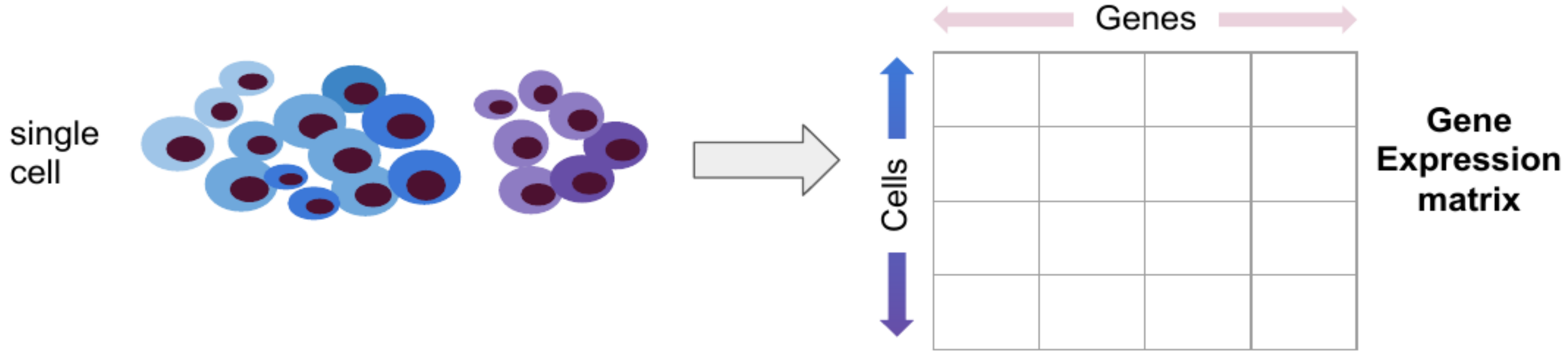


Every cell is unique;



We aim at understanding the function of each cell using data measured from multiple modalities of a cell

The types of data are usually in the form of high-dimensional matrices



Our computational tasks:

- Cluster the cells to find new cell types while integrating multiple data types
- Infer causality relationships between genes to understand mechanisms that control gene functions
- Develop *in silico* simulators which simulate realistic data as benchmark systems to evaluate new methods