Data Science for Human Well-being

Tim Althoff
Stanford University

Science Is Revolutionized By Data

Lessons from Online Social Networks

Network structure

- Small-World
  [Watts & Strogatz, 1998]
- Powerlaw topology
  [Faloutsos[3], 1999]
- Bowtie structure
  [Broder et al., 2000]

Network behavior

- Communication patterns
  [Leskovec & Horvitz, 2008]
- Information diffusion
  [Romero et al., 2011]

Lessons limited to **Online Behavior**

But how to capture offline behavior?

Wearable and Mobile Devices

69% adults own smartphones in developed countries
46% in developing economies (rapidly growing)

Wearable and mobile devices generate massive digital traces of real-world behavior and health

What did we learn from these data?

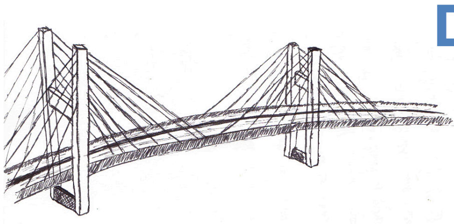
- Treasure of data with great promise
  - Data available for many years (e.g. Fitbit founded in 2007)
  - Data is regularly thrown away and overlooked

Today: How can we gain well-being insights from these data?

Physical Activity | Sleep | Mental Health

How to gain insights from these data?

Data Experts
Don't know what questions to ask & scientific impact

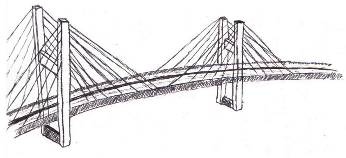


Domain Experts
Don't know data and how new methods could address their big questions

Gaining insights requires intersection of

- Knowing CS methods to extract insights from massive data
- Knowing data, its limitations, and how to address them
- Knowing big questions and how to find new ways to address them

My Research



New computational methods for digital activity traces to understand and improve human well-being

- Work with terabyte-scale data
- Conduct massive observational studies
- Generate actionable insights
- Impact health applications

Digital Activity Traces: The Data

- Multimodal data about our behaviors and health
  - Sensor data
  - Device usage data
  - Social interactions
  - Language

- Activity and health data across millions of people
  - Massive scale
  - Granular detail
  - Continuous & Long-term
  - Low cost

Impact of Digital Activity Traces: Health & Domain Experts

Limitations of health research today:

- Confined to laboratories
- Short-term (≤5 days), small scale (≤50 subjects), (binary) resolution
- Biases from self-reports/surveys (up to 700% off!)
- High cost

[Tucker et al., 2011]

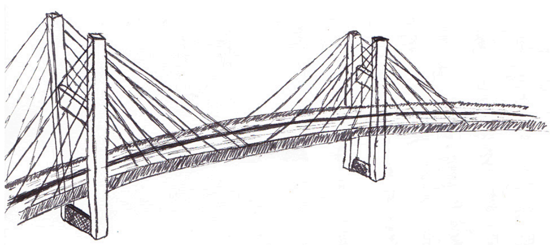
→ We know very little about our behavior & health

- How much do people exercise? What do people eat? What do they struggle with?

- **Opportunity: Improve human well-being**
  - **Advance science:** Better understanding of human behavior and health
  - **Improving healthcare:** Actionable insights

However…

...there is lack of computational models and large-scale analyses of digital activity traces for human well-being



Why is it hard to build a bridge?

Computational Challenges

Need new methods to address data limitations and model domain knowledge and questions.

1. How to integrate anecdotal and qualitative domain knowledge into computational models for empirical validation at scale
2. How to infer well-being from noisy raw data, or multimodal data sources
3. How to turn observational, biased, scientifically "weak" data into strong scientific results

Research Overview

- **Methods**
  - Data Mining
    WWW'18a, WWW'18b, WWW'18c, WWW'17a, WWW'15, KDD'15
  - Social Network Analysis
    WSDM'17, WWW'17b
  - Natural Language Processing
    TACL'16, ICWSM'14
- **Application Domains**
  - Health, Medicine and Psychology
    Nature'17, JMIR'16, NPJ DigMed'18, Pervasive Health'17

This Talk

Data Science Methods for Human Well-being

Physical Activity

1. How do patterns of activity vary around the world?
2. How can we model & predict everyday behavior?

Sleep

3. How to use search engines for sleep insights?

Mental Health

4. How to use natural language processing to improve mental health care?

Research Impact

My methods and insights are used at…

Physical Activity

UNDER ARMOUR

azumio

DEPARTMENT OF HEALTH & HUMAN SERVICES · USA

Sleep

Microsoft

bing

Mental Health

CRISIS TEXT LINE

talkspace THERAPY FOR HOW WE LIVE TODAY

Next

Data Science Methods for Human Well-being

Physical Activity

How do patterns of activity vary around the world?
2. How can we model & predict everyday behavior?

Sleep

3. How to use search engines for sleep insights?

Mental Health

4. How to use natural language processing to improve mental health care?

Althoff, Sosic, Hicks, King, Delp, Leskovec - Nature, 2017

In This Part…

1. How do patterns of activity vary globally?

[Althoff, Sosic, Hicks, King, Delp, Leskovec - Nature, 2017]

- **Macro-scale:** Leverage ubiquitous smartphone usage to study physical activity at planetary scale
- Defined & studied new measure: Activity Inequality (unevenly distributed activity)

2. How can we model everyday behavior?

[Kurashima, Althoff, Leskovec - WWW, 2018]

- **Micro-scale:** New machine learning methods to combat activity inequality by learning when to encourage individual users

Activity Tracking

Tracking actions

- Steps (automatic)
- Runs
- Walks
- Workouts
- Biking
- Weight
- Heart rate
- Food
- Drinks
- And many, many others

17

The Data

- Industry collaboration: Azumio freely shared data for open academic research

Azumio Dataset Statistics

- 5.6 million users
- Users from over 120 countries
- 791 million actions recorded
- 160 million days of steps tracking
  - >**230 billion** data points (3TB)

Challenge: How to connect data to long-standing domain questions?

18

How Physically Active Are We?

Physical activity is extremely important for health [Lee et al., 2012]. But we do not know how much physical activity people get!

According to WHO:

- 5-54% of Germans don't get enough activity
- No data for Switzerland and Israel

Health research limitations today:

- High cost, short-term, limited scale
- Biases from self-reporting

Worldwide Activity

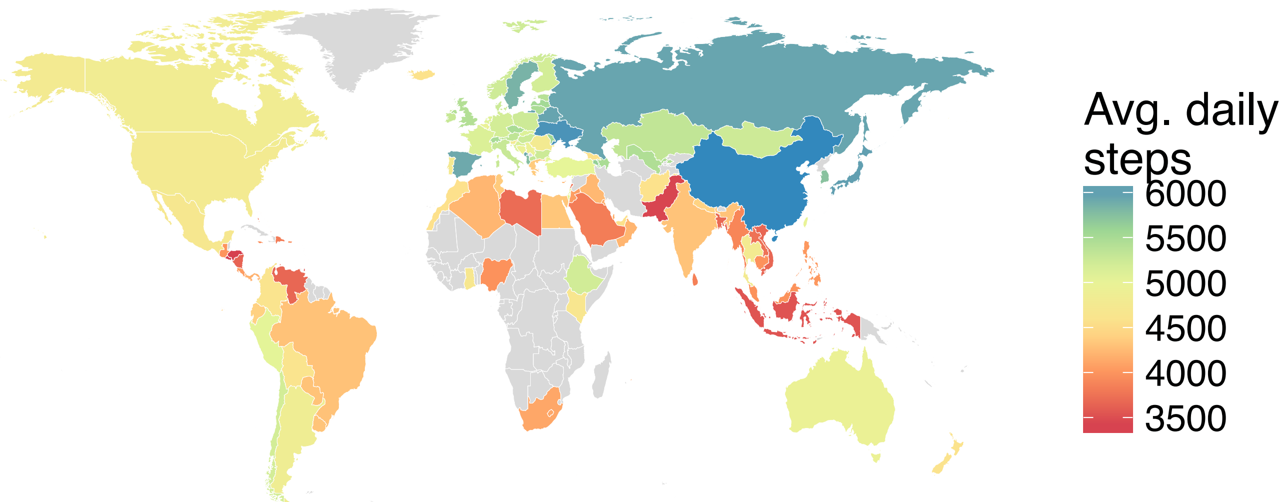

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video

Archive > Volume 547 > Issue 7663 > Letters > Article

Large-scale physical activity data reveal worldwide activity inequality

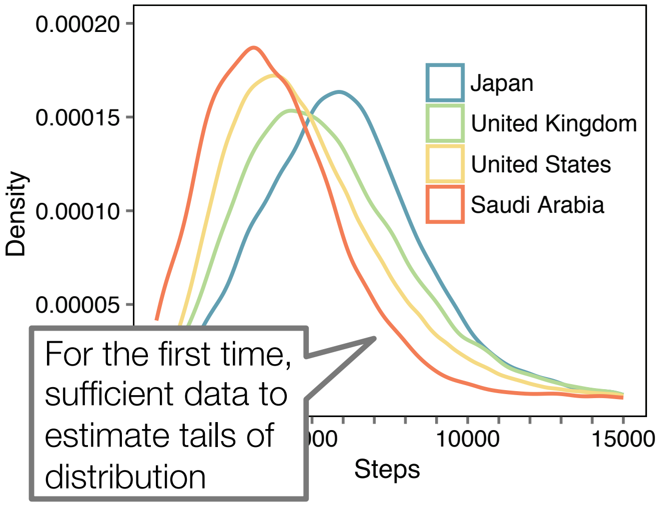
Tim Althoff, Rok Sosič, Jennifer L. Hicks, Abby C. King, Scott L. Delp & Jure Leskovec



But, how is activity distributed within the population?

Result 1: Inequality of Physical Activity

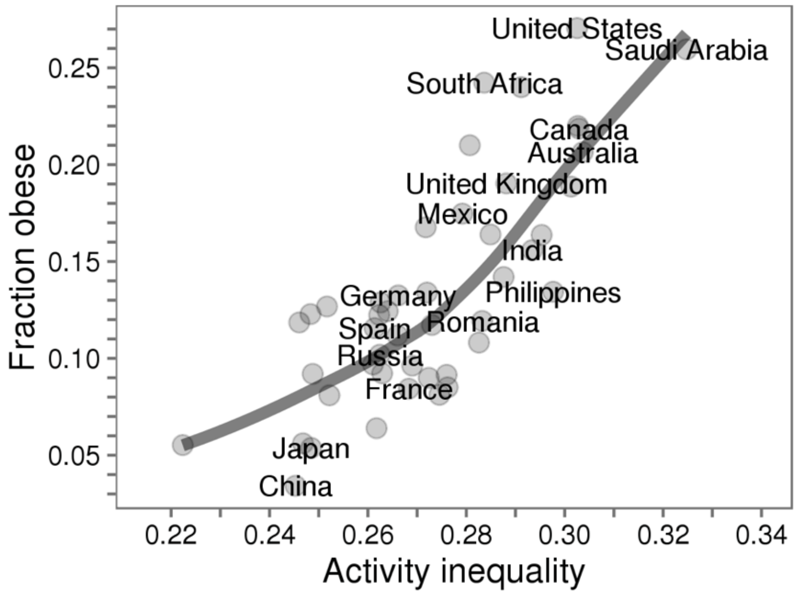
Difference in means



- **How (un)evenly is activity distributed?**
- Gini index of the activity distribution:
  - Activity rich vs. activity poor people

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|}{2\sum_{i=1}^{n}\sum_{j=1}^{n}x_j}$$

Result 2: Activity Inequality Predicts Obesity



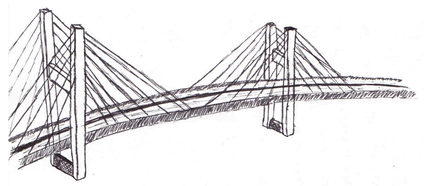
Tails/extremes matter more than the mean

R^2=0.64 (vs. 0.47 for avg. activity)

Massive digital traces uniquely enable studying tails!

The Challenge: Convincing Domain Experts

- New concept + new instrument = skepticism
- Domain experts know that these data are ...
  - Noisy
  - Sometimes inaccurate
  - Observational
  - Biased and full of selection effects
- That is why data have been thrown out before



- Designed and conducted over 20 reweighting, resampling, stratification, and simulation experiments to demonstrate validity of results

Demonstrating Validity of Results

…in light of valid concerns

- Flawed sensor?
- But women wear phones less?
- Obesity data inaccurate?
- Biased population?
- Due to rich people?
- Missing data? Outliers?
- Inaccuracy of location inference?
- Reproducible: Released analyses and data at http://activityinequality.stanford.edu

[Nature'17]

Research Implications

- Pioneered new paradigm for monitoring populations
- Working with public health researchers on implications for obesity, policy, urban planning

How to improve health by combating activity inequality?

- **Next: Moving from macro to micro level**
  - How to **target notifications and reminders** for each individual to encourage healthy behavior?

Modeling Everyday Behavior

- Apps tracks everyday behaviors:
  drink, food, sleep, weight, heart rate, running, walking, stretching, biking, workout, …

How can we model this behavior?

[WWW’18a]

Modeling Task

- **Task: Model what action user will take next and when**

Why is this useful?

- Predictions useful as interventions if they are timely and explainable
  - Timeliness: Diet support – send diet reminder *just before* meal choice
  - Explainability: “Hey, we saw you missed your weekly run this morning. How about tomorrow morning?”

[WWW’18a]

Why Is This Task Hard?

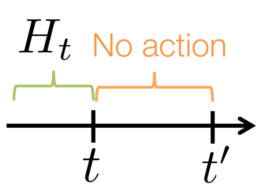
- **Human behavior is highly complex**
  - Actions vary over time
  - Interdependencies in short- & long-term
  - Creatures of habit with periodic behaviors
  - Individual preferences
- **Model requirements**
  - Predict action and continuous time
  - Need timely and explainable predictions

Model

[WWW'18a]

Background: Temporal Point Processes

- **Definition:** Random process whose realization consists of a list of discrete events localized in time $\{t_n\}_{n \in \mathbb{N}}$ with $t_n \in \mathbb{R}^+$
- **Benefits**
  - Generative process that predicts both action and time
  - Flexible through **conditional intensity function** $\lambda(t'|H_t)$ where H_t represents the history of actions until t
- Conditional density that **an event occurs** at time t'



$$f(t'|H_t) = \lambda(t'|H_t)\ \exp\left(-\int_t^{t'} \lambda(\tau|H_t)\, d\tau\right)$$

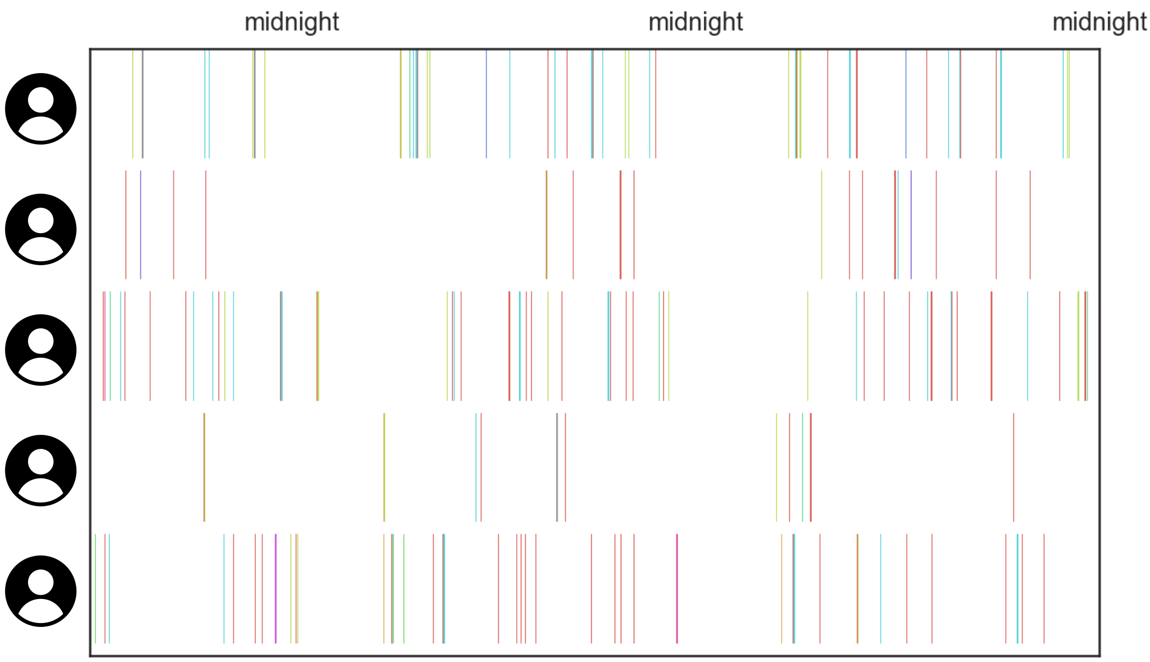
Event occurs at t' — $\lambda(t'|H_t)$

History H_t until t — H_t

No event occurred from t to t' — $\exp\left(-\int_t^{t'} \lambda(\tau|H_t)\, d\tau\right)$

[WWW'18a]

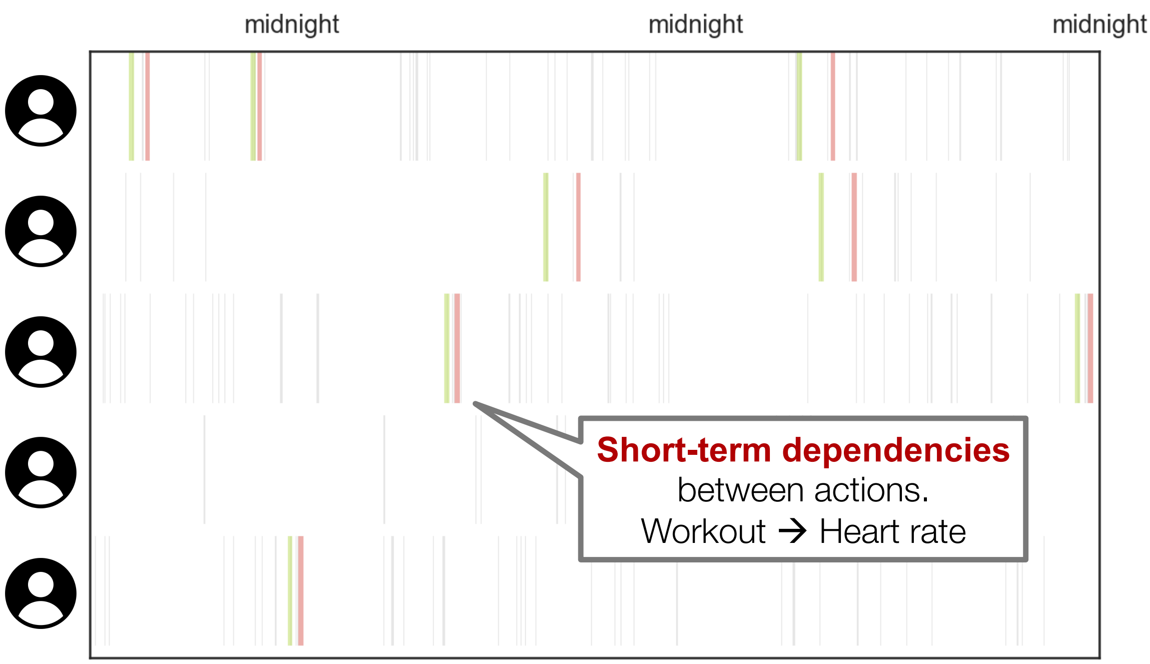
Real Activity Data



Color denotes activity type

[WWW'18a]

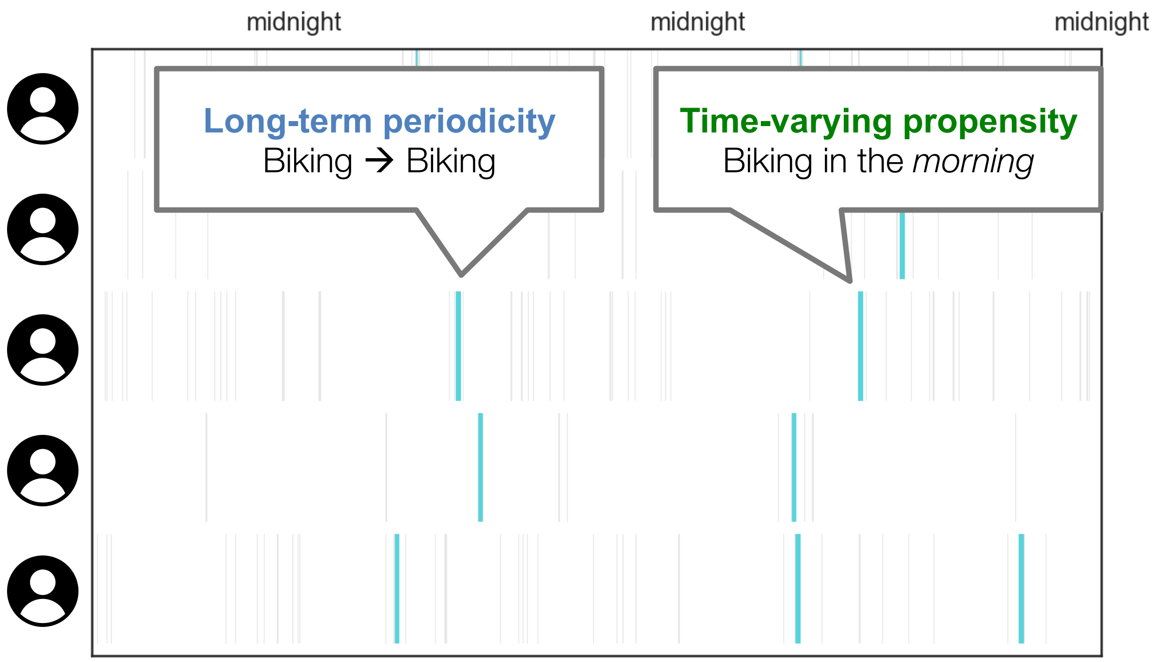
Real Activity Data



Color denotes activity type

[WWW'18a]

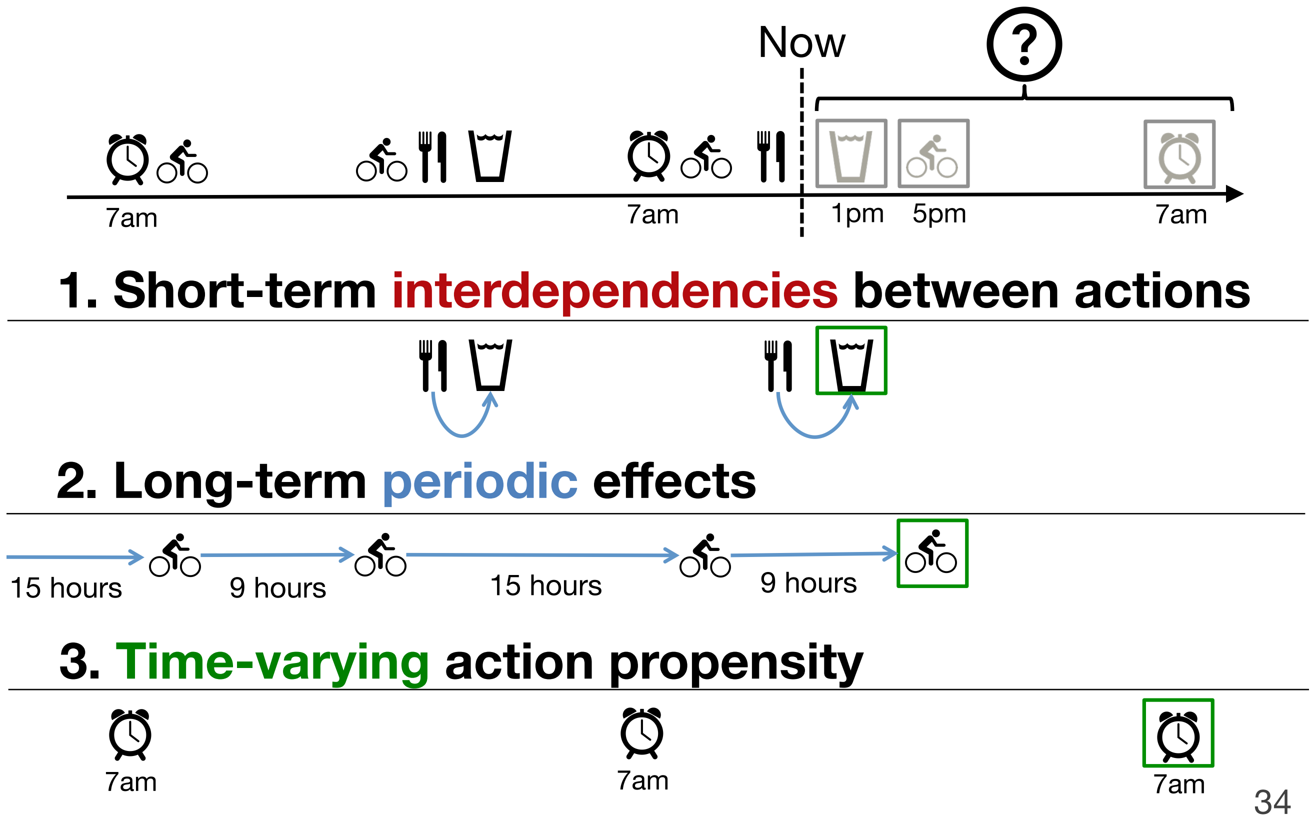
Real Activity Data



Color denotes activity type

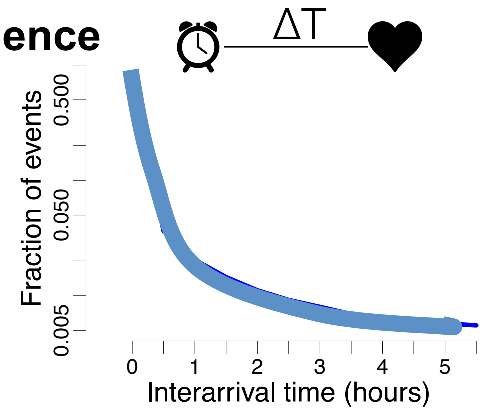
[WWW'18a]

My Approach: Three Components



1. Short-term Interdependency

Empirical Evidence



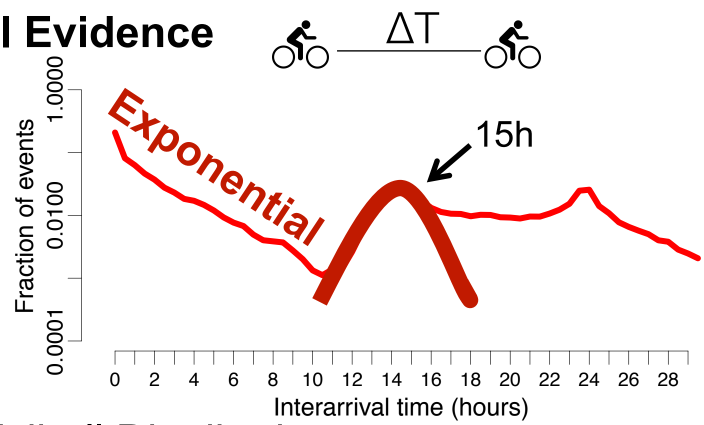
Model: Exponential Distribution

$$ShortTerm_u(t,a) = \sum_{(t',a') \in H_{ut}} \theta_{a'a} \omega_{a'a} \exp(-\omega_{a'a} \Delta_{t't})$$

- Exponential $\omega_{a'a} > 0$ Rate parameter – Shape
- Importance
- Sum over all previous events

2. Long-term Periodicity

Empirical Evidence



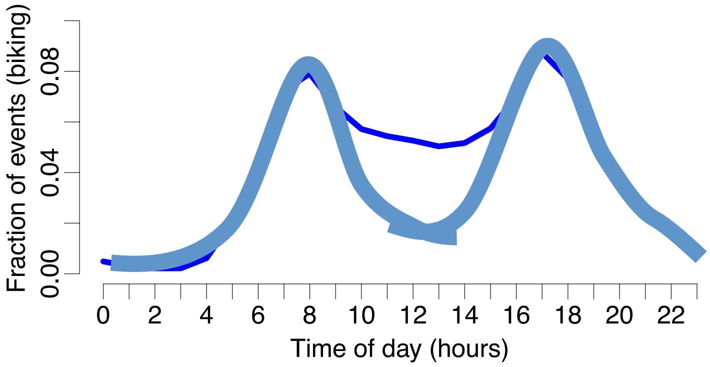
Model: Weibull Distribution

$$LongTerm_u(t,a) = \sum_{t' \in H^a_{ut}} \phi_{c_{t'}a} \gamma_{c_{t'}a} \kappa_{c_{t'}a} \Delta_{t't}^{\kappa_{c_{t'}a}-1} \exp(-\gamma_{c_{t'}a} \Delta_{t't}^{\kappa_{c_{t'}a}})$$

- Weibull $\kappa_{c_{t'}a} > 0$ Shape $\gamma_{c_{t'}a} > 0$ Scale
- Importance
- All previous events of same type

3. Time-varying Action Propensity

Empirical Evidence



Model: Mixture of Gaussians

$$Time_u(t,a) = \sum_{z \in \boldsymbol{Z}} \frac{\beta_{az}}{\sqrt{2\pi\sigma_{az}^2}} \exp\left(-\frac{(l_t - \mu_{az})^2}{2\sigma_{az}^2}\right)$$

Gaussian

Importance: How likely does Gaussian trigger event?

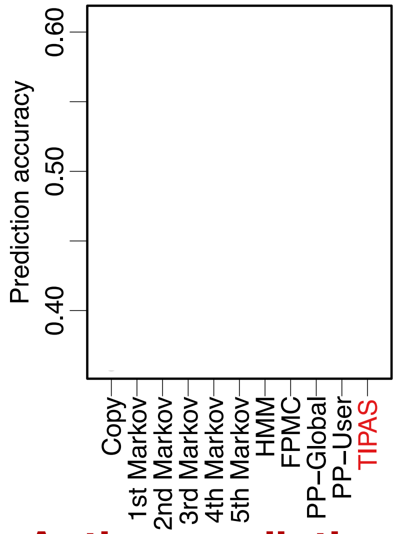
Model Inference

$$\lambda_u(t,a) = \boxed{\alpha_{ua}} + Time_u(t,a) + ShortTerm_u(t,a) + LongTerm_u(t,a)$$

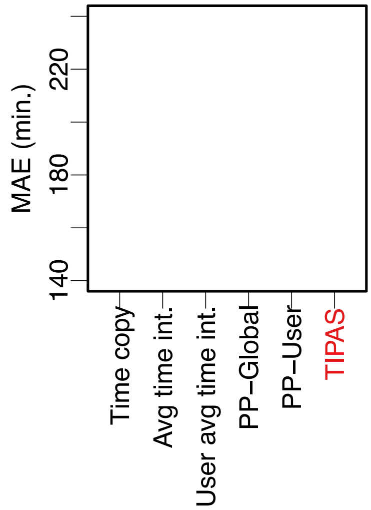
Personalization factor

- Learn parameters via Expectation-Maximization algorithm

Prediction Results



Action prediction (10)

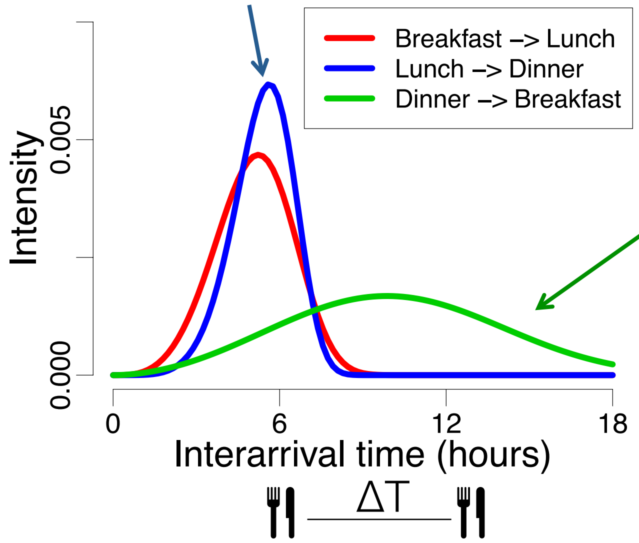


Time prediction

Model Explainability

- Few model parameters (~500)
- Can visualize inferred distributions to see what TIPAS model learned from data

Earlier lunches mean earlier dinners! (~5h period)



Not obvious! Does not hold for dinners!

Important for interventions (e.g. diet reminder)

[WWW'18a]

Modeling Summary

- Generative model that encodes empirical insights on human behavior
  - Takes previous actions into account (early lunch)
  - Models interdependencies between actions
- Predictions enable personalized health interventions
  - Timely and explainable predictions tell us when & how to notify users

Code and data available at http://snap.stanford.edu/tipas/

Next

Data Science Methods for Human Well-being

Physical Activity

1. How do patterns of activity vary around the world?
2. How can we model & predict everyday behavior?

Sleep

3. How to use search engines for sleep insights?

Mental Health

4. How to use natural language processing to improve mental health care?

Althoff, Horvitz, White, Zeitzer – WWW, 2017
Althoff, Horvitz, White – NPJ Digital Medicine, 2018

In This Part…

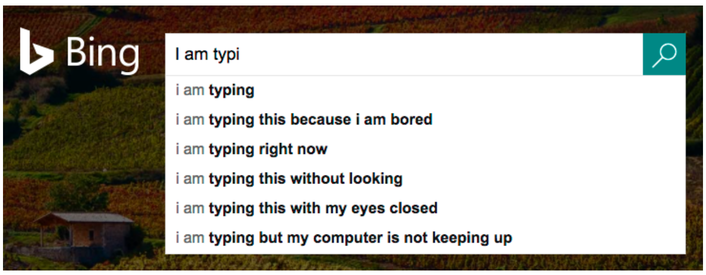
- **Q:** How does sleep affect cognitive performance?
- **Bridge: Search logs studied for a decade, domain experts never thought of looking there**
  - First-ever combination of web search and wearable data
  - Statistical model encoding biological domain knowledge

bing +

[WWW'17a]

Key Insight: Cognitive Performance through Search Engine Interactions

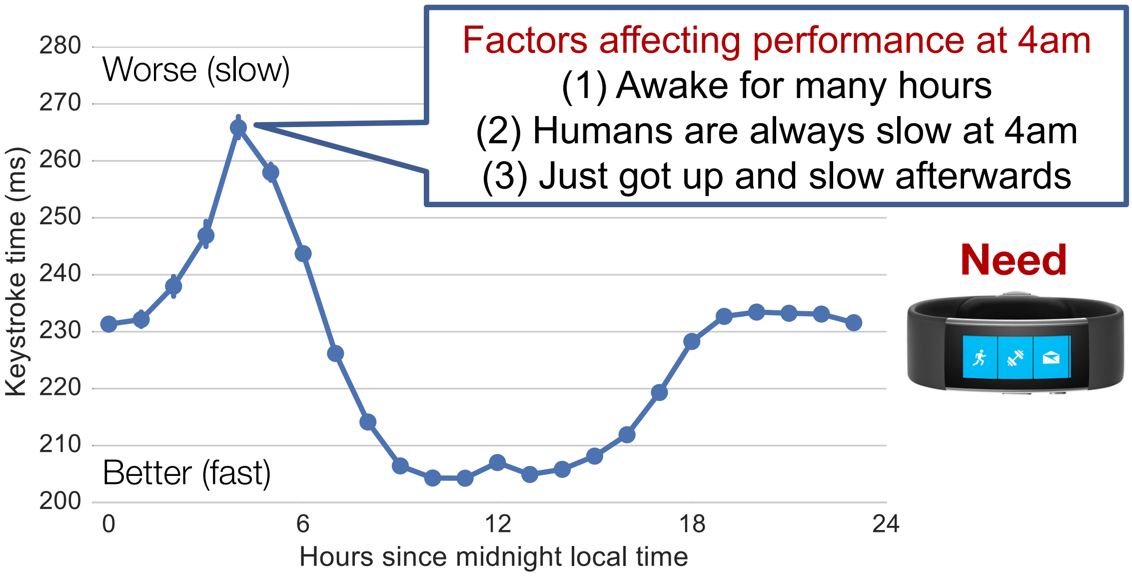
- Search engines are used repeatedly every day, awake or sleepy, by billions of people
- **Key insight:** Reframe everyday interactions with web search engine as series of performance tasks
  - Query typing speed (or click on search result)



fa
fac
face
…

Δt("c") = 237ms
Δt("e") = 219ms

Result: Real-World Performance Variation



Need

- Performance far from constant (31% variation)

How can we distinguish these three factors?

Modeling Challenges

How to disentangle the three effects?

- Many factors, highly correlated
- Current approach: Forced desynchrony protocol in sleep lab & active sleep deprivation at tiny scale

My approach

- Leverage existing variation of real-world interactions with web search engines across millions of people
- Develop statistical model to disentangle effects

Biologically-inspired Statistical Model

- Bridge: Generative model encoding multiple biological processes to disentangle effects (domain knowledge)
- **Generalized Additive Model** [Hastie & Tibshirani, 1990]

$$y \sim \boldsymbol{N}(\mu(x), \sigma^2)$$

Keystroke timing

Keystroke features

Gaussian noise

Intercept

Time of day

$$\mu(x) = \boxed{\alpha} + \boxed{f(\mathrm{key}(x))} + \boxed{g(\mathrm{timeofday}(x))} + \boxed{h(\mathrm{timeawake}(x))}$$

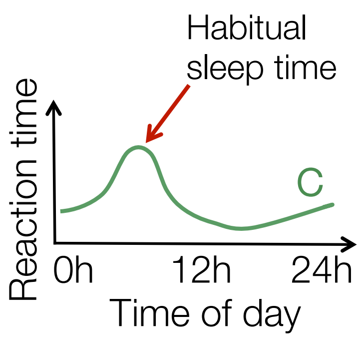
Keystroke
(control for key pressed: "A", "a", "@", …)

Time since wakeup
(wearable sleep measurement)

Model: Why Time of Day?

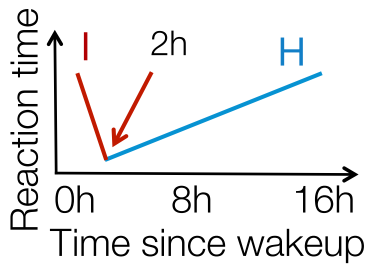
- Lab studies: Several biological processes drive performance variation
  1. Circadian rhythm (C): behavior-independent, near 24h oscillations that is time-dependent

→ model time of day $\boxed{g(\mathrm{timeofday}(x))}$



Model: Why Time Since Wakeup?

- Two additional biological processes impact performance
  2. Homeostatic sleep drive (H): the longer awake, the more tired you become
  3. Sleep inertia (I): performance impairment experienced immediately after waking up

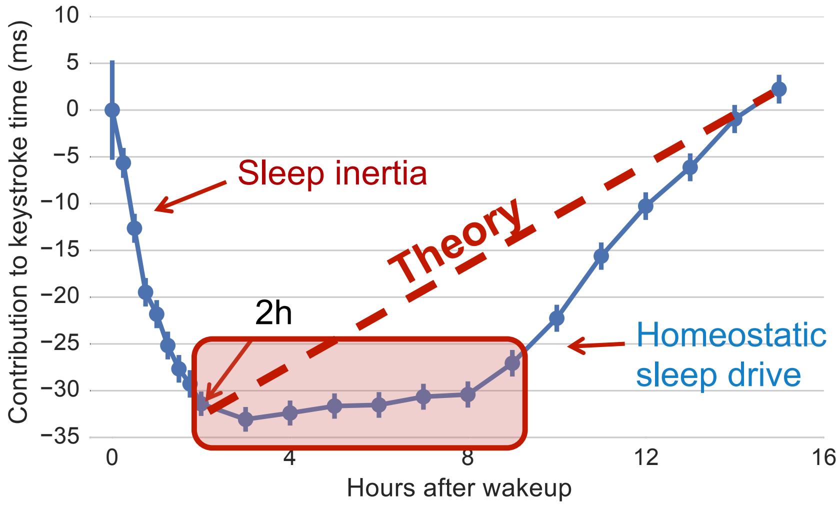


Model: Parameter Learning

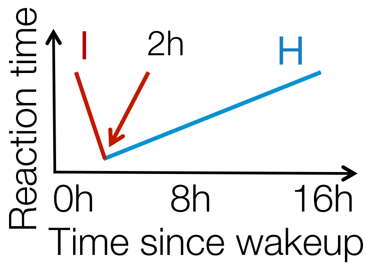
$$\mu(x) = \boxed{\alpha} + \boxed{f(\mathrm{key}(x))} + \boxed{g(\mathrm{timeofday}(x))} + \boxed{h(\mathrm{timeawake}(x))}$$

- No assumptions about functional form!
- Convex optimization problem (~1000 parameters, ~75M observations)

Result: Time Since Wakeup



$$\boxed{h(\text{timeawake}(x))}$$



- **Validation:** Model identifies homeostatic sleep drive and sleep inertia consistent with lab-based studies
- **New insights:** It was impossible to measure cognitive performance at scale and outside lab. Now we can!

Research Impact

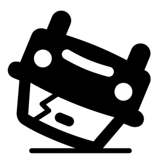
New science

[Althoff, Horvitz, White, Zeitzer – WWW, 2017]

1. Used my method to estimate impact of **sleep deprivation** on real-world performance
   - Largest-ever study by 400x

Reducing vehicle accidents

[Althoff, Horvitz, White – NPJ Digital Medicine, 2018]



2. Used my method at US population scale to predict vehicle **accident risk**
   - 16 billion keystrokes across ~2700 US counties
   - Technology could help **reduce vehicle accidents**

Next

Data Science Methods for Human Well-being

Physical Activity

1. How do patterns of activity vary around the world?
2. How can we model & predict everyday behavior?

Sleep

3. How to use search engines for sleep insights?

Mental Health

4. How to use natural language processing to improve mental health care?

Althoff*, Clark*, Leskovec - TACL, 2016

NLP for Mental Health

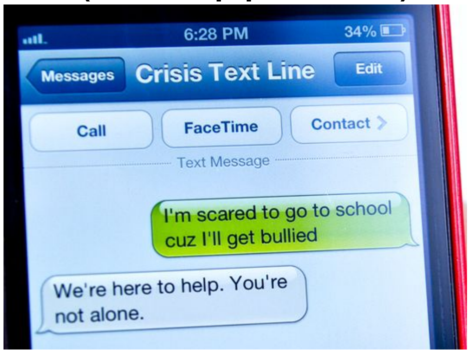
- **Question:** How to talk to someone to help them feel better?

- Mobile devices enable counseling conversations wherever you are
  - Massive scale: >56M messages to date
  - Daily(!) active rescues for danger of suicide

CRISIS TEXT LINE |™

Leveraging Data to Improve Treatment

- Text-based counseling enables quantitative study of conversation strategies (IRB approved)
  - Full conversation transcripts
  - Conversation outcomes



- *Helps answer important questions*
  - *Why are some counselors much better than others?*

Data-driven Conversation Strategies

Developed computational models and provided quantitative evidence for five conversation strategies:

1. Adaptability: Language model comparison
   - Best counselors adapt to conversation
2. Dealing with ambiguity: Clustering
   - Best counselors react differently to identical situations
3. Creativity: Subspace analysis
   - Best counselors use less generic/templated language
4. Making progress: HMM extension
   - Best counselors understand problem quickly & solve
5. Change in perspective: Coordination analysis
   - Best counselors change people's perspective

Althoff*, Clark*, Leskovec - TACL, 2016

[TACL’16]

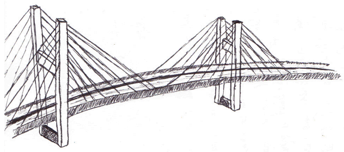
Mental Health: Impact

- Insights concretely improved counseling training

CRISIS TEXT LINE ™

Summary

Talk Summary



- Digital traces capture behavior and health at scale
- New methods needed to unlock insights
- Developed new methods in Data Mining, Social Network Analysis, Natural Language Processing
  - Concrete impact on understanding of human well-being
  - My methods and insights have been used at Microsoft, Under Armour, Crisis Text Line, and many other orgs.

Acknowledgements





Collaborators & Colleagues

Rok Sosic, David Hallac, Will Hamilton, Hima Lakkaraju, Jennifer Hicks, Abby King, Adam Miner, Takeshi Kurashima, Emma Pierson, Zhiyuan Lin, Michele Catasta, Srijan Kumar, Marinka Zitnik, Ashton Anderson, Austin Benson, Caroline Lo, Robert West, Justin Cheng, Xiang Ren, David Jurgens, Mitchell Gordon, Boris Ivanovic, Hamed Nilforoshan, Jamie Zeitzer, Kevin Clark, Andrej Krevl, Peter Kacin, Adrijan Bradaschia, Joy Ku, Jessica Selinger, Ina Fiterau, Jason Fries, Jenna Hua, Eric J Daza, Steven Bell, Hector Garcia-Molina, Jeff Ullman, Yesenia Gallegos, Marianne Siroker, Pranav Jindal, Ali Shameli, Amin Saberi, Ryen White, Cristian Danescu-Niculescu-Mizil, Xin Luna Dong, Kevin Murphy, Safa Alai, Van Dang, Wei Zhang, the SNAP crew, the InfoLab, the Mobilize Center, et al.

Family & Friends

Thank you!

Tim Althoff
althoff@cs.stanford.edu

Talk Overview

Data Science Methods for Human Well-being

Physical Activity
1. How do patterns of activity vary around the world?
2. How can we model & predict everyday behavior?

Sleep
3. How to use search engines for sleep insights?

Mental Health
4. How to use natural language processing to improve mental health care?