Anonymity and Identity Online
Florian Ederer, Paul Goldsmith-Pinkham & Kyle Jensen

The analysis described in this project was allowed to proceed by the Yale HRPP, protocol 2000034072.

Warning: This presentation includes numerous excerpts of offensive speech including racism and threats of sexual violence.
Description of the Project

“Economics is what economists do.”

• Questions
  - What do economists say when they are anonymous?
  - How widespread is toxic speech in economics? Who engages in such speech?

• Econ Job Market Rumors (EJMR) is a popular anonymous message board
  - 2.5 million monthly visits (SimilarWeb, 2022)

• Statistical properties of EJMR usernames reveal IPs for the majority of posts
  - Focus today will be on methods of IP address identification
  - We document widespread use of EJMR even at top universities

• We use only publicly available data (Wu, 2020)
Nonlinear least squares with fixed effects

any good place to read up on the theory? I know it can have issues like incidental parameters but I forget how to prove it formally, or any guidelines specific to NLS

Bump

This will not work. Period

If the FE are additive you to the nonlinear function problem. But if you have

Why subject yourself to all

I'm an AP who's likely to get tenure pretty soon at a top 10 economics department, based on CV and signals from seniors in my dept. My goal is to move to a good-but-clearly-lower-ranked school for a big pay raise right around tenure time. I have a US preference, but am pretty locationally flexible.

Which schools should I reach out to? Do any schools have a reputation for paying a lot to recruit seniors from higher ranked schools?
Gossip (about senior faculty moves) on EJMR

NU Macro Senior

Who did NU hire to replace MD & GL?
NU macro down to toilet?

1 YEAR AGO #2ED9 QUOTE 0 VOLOD 0 VLAD

GL to Chicago confirmed?

1 YEAR AGO #D799 QUOTE 0 VOLOD 0 VLAD

Yes

1 YEAR AGO #832A QUOTE 0 VOLOD 0 VLAD

GL to Chicago confirmed?

Yes
Racism on EJMR

If your white and you ain’t got least three kids your a traiter to your country

So tired of seeing N----S in my neighborhood!

Just run them over and make it look like an accident. Nobody cares about a de/ad N----R.

18 HOURS AGO # DB0D QUOTE 1 VOLOD 0 VLAD

17 HOURS AGO # E2DE QUOTE 1 VOLOD 0 VLAD
obviously, students at harvard know what is the easiest and fastest route to a qje. why is this surprising?

wrong the fastest route to a qje is to grift and be black
Nothing can prepare you for how awful working with women is

Astonishing. Have you told Mother?

3 MONTHS AGO # 4080 QUOTE 3 VOLOD 4 VLAD

Things were WAY better when women focused on rearing children and feeding their husbands.

3 MONTHS AGO # 453C QUOTE 8 VOLOD 6 VLAD

And slaughtering kids.

3 MONTHS AGO # 0A18 QUOTE 7 VOLOD 7 VLAD

If you run into someone in the morning, you run into some guy. If you run into people who are emotional, selfish and incompetent all day, you're the emotional, selfish and incompetent one.

3 MONTHS AGO # 3A62 QUOTE 3 VOLOD 10 VLAD

Sexism and Misogyny on EJMR

One of our female APs parties all the time. I mean, she goes to the bar pretty often. Not to embarrass her, I stopped going to a bar where she patrons often (she didn't see me). She is very revealing too. Not sure about other female APs or male APs. I don't think she will publish up to standard. But she will almost certainly get tenure if the university doesn't want a lawsuit.

4 YEARS AGO # 07BA QUOTE 0 VOLOD 2 VLAD

Hire whoever has the biggest boooobs.

all girls? no diversity?

Columbia flyouts:

9 YEARS AGO # AE17 QUOTE 4 VOLOD 10 VLAD
Racist and Sexist Discussion of Job Market Candidates

NYU Job Market Candidates 2022-2023

https://as.nyu.edu/departments/econ/job-market/candidates.html

Good luck to them!

6 MONTHS AGO # DB39 QUOTE 5 VOLOD 0 VLAD

Only 1 Asian out of 12? I hope they get hired.

6 MONTHS AGO # D7F1 QUOTE 1 VOLOD 0 VLAD

Another week cohort from NYU....

6 MONTHS AGO # 3B52 QUOTE 1 VOLOD 0 VLAD

EJMR

Are [redacted] and [redacted] in a secret same-sekhs love-hayte relationship?

Let's move our eyes from [redacted] to USC Zhangettes on the job market.

They look so good omg. I am coming. I am licking the screen

4 MONTHS AGO # 3FA8 QUOTE 1 VOLOD 0 VLAD

Mode, what are you? You deleted my comment telling you to delete this.
Who writes such things?

**Economist e96c**

*people from Hong Kong are the worst cheaters and rent seekers.*

I lower the grades of all Mainland Chinese students. Not TW or HKers. I look at the names and they way they are romanized. Yale System Cantonese name e.g. Wong Tai Gok = Hong Konger → B+ becomes A-. Wade Giles e.g. Wang Tong-Hui = Taibianese → B+ becomes A+. Pinyin e.g. Wang Xiaodiao = Mainlander → B+ becomes C-. I think that’s only fair and compensates for the likelihood that they cheated their way in. Also, I just don’t like them.

You’re working too hard. If their English is good, they are from Hong Kong. If their English is just okay, they are from Taiwan. If the only reason they passed the TOEFL is because they had another person take their exam and they speak no English at all, they are a mainlander.

**Economist a483**

*Hell Merkler!*

If it is black guys on white girl it is a cultural diversity training, proper to show on TV and in cinema.

If it is white guys on black girl it is a hate crime, white guys go to prison, black girl get multimillion legal settlement.

**Economist a863**

*It's about ching chong taking bubba's job and bubba putting on a white pointy hood in response.*

Rapefugees Welcome!!!

- Merkel

**Economist 577b**

*Trenchant insight, bro*

I just notices there is practically no black nudity on TV anymore, unless it is BBC, while there are plenty of white p*ssy everywhere wide open.

**Economist 5b12**

*Chile gets warm. And bitches are fugly.*

If you don't like hot weather you should consider Moscow, Russia.

*after all the people giving him blow jobs on EMRs you would have expected great things.*
Who writes such things?

IP addresses at Harvard, Stanford, Yale, Chicago ...
Who writes such things?

**Economist e96c**
people from Hong Kong are the worst cheaters and rent seekers.

I lower the grades of all Mainland Chinese students. Not TW or HKers.
I look at the names and they way they are romanized.
Yale System Cantonese name e.g. Wong Tai Gok = Hong Konger --> B+ becomes A-
Wade Giles e.g. Wang Teng-Hui = Taiwanese --> B+ becomes A-
Pinyin e.g. Wang Xiaoliao = Mainlander --> B+ becomes C-
I think that’s only fair and compensates for the likelihood that they cheated their way in. Also, I just don’t like them.

You’re working too hard. If their English is good, they are from Hong Kong. If their English is just okay, they are from Taiwan. If the only reason they passed the TOEFL is because they had another person take their exam and they speak no English at all, they are a mainlander.

**Economist d644**

It’s about ching chong taking bubba’s job and bubba putting on a white pointy hood in response.

**Economist a483**
Rape refugees Welcome!!!!
- Merkel

I just notices there is practically no black nudity on TV anymore, unless it is BBC, while there are plenty of white pussy everywhere wide open.

**Economist 577b**

I hate hot weather. South Korea looks nice - no hotter than 23 degrees C. Chile is good too. I don’t want to work in some miserable hot, sticky jungle or desert place.

Chile gets warm. And bitches are fugly.
If you don’t like hot weather you should consider Moscow, Russia.

**Economist 6b12**

after all the people giving him blow jobs on EMRs you would have expected great things.

IP addresses at Harvard, Stanford, Yale, Chicago ... and the NBER HQ at 1050 Mass Ave
Username Allocation on EJMR

• Each topic is assigned a **topic id**
  - https://www.econjobrumors.com/topic/right-vs-left-wing-dictatorships
  - https://www.econjobrumors.com/topic/1127272

“slug”

“topic id”

An incrementing counter common to WordPress sites
Username Allocation on EJMR

• Each topic is assigned a topic id
  - https://www.econjobrumors.com/topic/right-vs-left-wing-dictatorships
  - https://www.econjobrumors.com/topic/1127272

• Each post is assigned a username, which is topic-specific

Economist
824e

OK

59 MINUTES AGO # QUOTE 1 VOLOD 0 VLAD
Username Allocation on EJMR

- Each topic is assigned a **topic id**
  - https://www.econjobrumors.com/topic/right-vs-left-wing-dictatorships
  - https://www.econjobrumors.com/topic/1127272

- Each post is assigned a **username**, which is topic-specific

- **Username** was characters 10-13 of the SHA-1 hash of **topic ID** and **IPv4 address**
  - Hash **did not use a salt** (a random secret) and was in plain sight for over a decade

- E.g. **824e** — a 4-digit hexadecimal

<table>
<thead>
<tr>
<th>Topic ID:</th>
<th>1127272</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv4 Address:</td>
<td>130.132.173.94</td>
</tr>
<tr>
<td>SHA-1 Input:</td>
<td>1127272130.132.173.94</td>
</tr>
<tr>
<td>SHA-1 Output:</td>
<td>eae3d51bc824e273e203b9fbfb608828431a6d48</td>
</tr>
<tr>
<td>Username:</td>
<td>824e</td>
</tr>
</tbody>
</table>
From usernames to IP addresses

• SHA-1 hash is one-way. It cannot be reversed.

• But we can try every possible IP address to see if some of them match:

\[
\begin{align*}
11272720.0.0.0 & \rightarrow 59d5e2cc45a94a9b27a1c5e69644783d9f16726a \\
11272720.0.0.1 & \rightarrow d77bff2cfa90e8e2053b9b5d28847a1cfc919510 \\
& \cdots \cdots \\
1127272130.132.173.94 & \rightarrow eae3d51bc824e273e203b9fbfb608828431a6d48 \\
& \cdots \cdots \\
1127272255.255.255.255 & \rightarrow 0ef63f5c3d79e2fce6a2270dc52f88118a2a3949
\end{align*}
\]

• Record all IP addresses where observed topic-username matches the hash
  - Feasible on GPUs
  - 695,364 topics \times 2^{32} possible IPv4 addresses \approx 3 \text{ quadrillion hashes}
  - Returns a set of matching IPs per topic-username: 65,536 in expectation
First post:

- **IP**: 130.132.173.94
- **Topic ID**: 1127272
- **Username**: 824e

Second post:

- **IP**: 130.132.173.94
- **Topic ID**: 1127329
- **Username**: 607e

---

**Economist**

- **824e**
- 59 minutes ago # QUOTE 1 VOLOD B VLAD

**Economist**

- **607e**
- 59 minutes ago # QUOTE 1 VOLOD B VLAD

---

**Matching IPs for 1127272/824e**

- 0.1.96.120
- 0.3.5.107
- 0.4.159.248
- 0.6.29.0
- 0.2.213.60
- 0.4.41.216
- 0.5.109.255
- 0.7.104.143
- 255.248.137.217
- 255.249.166.138
- 255.250.174.16
- 255.253.207.68
- 255.254.220.185

65,668 IPs

**Matching IPs for 1127329/607e**

- 0.0.229.36
- 0.2.255.120
- 0.4.27.95
- 0.7.209.187
- 0.0.229.36
- 0.2.255.120
- 0.4.27.95
- 0.7.209.187
- 255.246.223.98
- 255.247.105.179
- 255.247.157.134
- 255.252.29.118
- 255.255.56.111

65,907 IPs

**Which IPs occur in both sets?**

- 130.132.173.94
First post: 

<table>
<thead>
<tr>
<th>IP</th>
<th>130.132.173.94</th>
</tr>
</thead>
<tbody>
<tr>
<td>topic ID</td>
<td>1127272</td>
</tr>
<tr>
<td>username</td>
<td>824e</td>
</tr>
</tbody>
</table>

My IP: 130.132.173.94

Second post: 

<table>
<thead>
<tr>
<th>IP</th>
<th>130.132.173.94</th>
</tr>
</thead>
<tbody>
<tr>
<td>topic ID</td>
<td>1127329</td>
</tr>
<tr>
<td>username</td>
<td>607e</td>
</tr>
</tbody>
</table>

Matching IPs for 1127272/824e

- 65,668 IPs

  - 0.1.96.120
  - 0.3.5.107
  - 0.4.159.248
  - 0.6.29.0
  - 0.2.213.60
  - 0.4.41.216
  - 0.5.109.255
  - 0.7.104.143
  - 255.248.137.217
  - 255.249.166.138
  - 255.250.174.16
  - 255.253.207.68
  - 255.254.220.185

Matching IPs for 1127329/607e

- 65,907 IPs

  - 0.0.229.36
  - 0.1.4.220
  - 0.2.255.120
  - 0.4.27.95
  - 0.7.209.187
  - 0.8.254.155
  - 255.246.223.98
  - 255.247.105.179
  - 255.247.157.134
  - 255.252.29.118
  - 255.255.56.111

Which IPs occur in both sets?

Just one: 130.132.173.94
Identifying true IP is possible because

1. True IP is always present in matching set
2. Other “noise” IPs are i.i.d. uniformly over IP space due to SHA-1 avalanche property

True IPs will show up much more often than noise IPs!

Statistical properties

- Probability of a “noise” IP appearing in a topic follows a hypergeometric distribution
- Number of times a “noise” IP appears in a week follows a Poisson binomial distribution
What is the IP address for this post with topic id = 175901 and username = 6b42?
An Attribution Example

- We start with $2^{32}$ possible IP addresses.
- The hash inversion narrows it down to 65,385 matching IP addresses for this post.
- How many of these 65,385 IPs explain other topic-usernames in a 7-day window?

<table>
<thead>
<tr>
<th># other explained $(t, u)$</th>
<th>p-value</th>
<th># IP addresses</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.00000000</td>
<td>58,367</td>
</tr>
</tbody>
</table>
An Attribution Example

- We start with $2^{32}$ possible IP addresses.
- The hash inversion narrows it down to 65,385 matching IP addresses for this post.
- How many of these 65,385 IPs explain other topic-usernames in a 7-day window?

<table>
<thead>
<tr>
<th># other explained $(t, u)$</th>
<th>p-value</th>
<th># IP addresses</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.000000000</td>
<td>58,367</td>
</tr>
<tr>
<td>1</td>
<td>0.10694695</td>
<td>6,627</td>
</tr>
</tbody>
</table>

about 6,605 expected by chance
An Attribution Example

- We start with $2^{32}$ possible IP addresses.
- The hash inversion narrows it down to 65,385 matching IP addresses for this post.
- How many of these 65,385 IPs explain other topic-usernames in a 7-day window?

<table>
<thead>
<tr>
<th># other explained $(t, u)$</th>
<th>p-value</th>
<th># IP addresses</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.000000000</td>
<td>58,367</td>
</tr>
<tr>
<td>1</td>
<td>0.10694695</td>
<td>6,627</td>
</tr>
<tr>
<td>2</td>
<td>0.00592438</td>
<td>378</td>
</tr>
</tbody>
</table>

about 373 expected by chance
An Attribution Example

- We start with $2^{32}$ possible IP addresses.
- The hash inversion narrows it down to 65,385 matching IP addresses for this post.
- How many of these 65,385 IPs explain other topic-usernames in a 7-day window?

<table>
<thead>
<tr>
<th># other explained $(t, u)$</th>
<th>p-value</th>
<th># IP addresses</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.00000000</td>
<td>58,367</td>
</tr>
<tr>
<td>1</td>
<td>0.10694695</td>
<td>6,627</td>
</tr>
<tr>
<td>2</td>
<td>0.00592438</td>
<td>378</td>
</tr>
<tr>
<td>3</td>
<td>0.00022049</td>
<td>12</td>
</tr>
</tbody>
</table>

about 12 expected by chance
An Attribution Example

- We start with $2^{32}$ possible IP addresses.
- The hash inversion narrows it down to 65,385 matching IP addresses for this post.
- How many of these 65,385 IPs explain other topic-usernames in a 7-day window?

<table>
<thead>
<tr>
<th># other explained $(t, u)$</th>
<th>p-value</th>
<th># IP addresses</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.000000000</td>
<td>58,367</td>
</tr>
<tr>
<td>1</td>
<td>0.10694695</td>
<td>6,627</td>
</tr>
<tr>
<td>2</td>
<td>0.00592438</td>
<td>378</td>
</tr>
<tr>
<td>3</td>
<td>0.00022049</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>$6.167 \times 10^{-6}$</td>
<td>0</td>
</tr>
</tbody>
</table>

About 0.39 expected by chance.
An Attribution Example

- We start with $2^{32}$ possible IP addresses.
- The hash inversion narrows it down to 65,385 matching IP addresses for this post.
- How many of these 65,385 IPs explain other topic-usernames in a 7-day window?

<table>
<thead>
<tr>
<th># other explained $(t, u)$</th>
<th>p-value</th>
<th># IP addresses</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.00000000</td>
<td>58,367</td>
</tr>
<tr>
<td>1</td>
<td>0.10694695</td>
<td>6,627</td>
</tr>
<tr>
<td>2</td>
<td>0.00592438</td>
<td>378</td>
</tr>
<tr>
<td>3</td>
<td>0.00022049</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>$6.167 \times 10^{-6}$</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>$1.380 \times 10^{-7}$</td>
<td>0</td>
</tr>
</tbody>
</table>

about 0.0089 expected by chance
An Attribution Example

- We start with $2^{32}$ possible IP addresses.
- The hash inversion narrows it down to 65,385 matching IP addresses for this post.
- How many of these 65,385 IPs explain other topic-usernames in a 7-day window?

<table>
<thead>
<tr>
<th># other explained $(t, u)$</th>
<th>p-value</th>
<th># IP addresses</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.00000000</td>
<td>58,367</td>
</tr>
<tr>
<td>1</td>
<td>0.10694695</td>
<td>6,627</td>
</tr>
<tr>
<td>2</td>
<td>0.00592438</td>
<td>378</td>
</tr>
<tr>
<td>3</td>
<td>0.00022049</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>$6.167 \times 10^{-6}$</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>$1.380 \times 10^{-7}$</td>
<td>0</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
An Attribution Example

- We start with $2^{32}$ possible IP addresses.
- The hash inversion narrows it down to 65,385 matching IP addresses for this post.
- How many of these 65,385 IPs explain other topic-usernames in a 7-day window?

<table>
<thead>
<tr>
<th># other explained $(t, u)$</th>
<th>p-value</th>
<th># IP addresses</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.000000000</td>
<td>58,367</td>
</tr>
<tr>
<td>1</td>
<td>0.10694695</td>
<td>6,627</td>
</tr>
<tr>
<td>2</td>
<td>0.00592438</td>
<td>378</td>
</tr>
<tr>
<td>3</td>
<td>0.00022049</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>$6.167 \times 10^{-6}$</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>$1.380 \times 10^{-7}$</td>
<td>0</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>102</td>
<td>$2.253 \times 10^{-258}$</td>
<td>0</td>
</tr>
<tr>
<td>103</td>
<td>$2.072 \times 10^{-267}$</td>
<td>1</td>
</tr>
</tbody>
</table>

Very much not expected by chance
IP Attribution as an Optimization Problem

- Find a small set of these active IPs that explains the observed data.
  - Subject to locality and significance constraints
  - *Definitely has errors.* These are estimable and in the paper.

- Simple intuition of this optimization approach
  - 65k IP addresses can explain any given post.
  - One IP explains many other posts that week.
  - What is the likely origin IP of this post?
  - It’s probably this highly explanatory IP.

This is essentially an argument that number of users $\ll 4.3B$ possible IP addresses
Attribution with Correct and Incorrect Hashes

- Using any incorrect substring of the hash only generates “noise” IPs
- Determine $p$-value threshold $p^*$ by comparing distributions of minimum $p$-values
- Only assign IPs to posts with minimum $p$-value $< p^* \approx 10^{-11}$

![Graph showing the share of observations vs. P-value of Poisson-Binomial Test](image)

We attribute 67.9% of posts.
Robustness Checks for Attribution

• Incorrect hashing set ✓
  - Using the (incorrect) position-11 hashing set, none of the roughly 7.1 million posts observed on EJMR are attributed an IP address.

• Bogon addresses ✓
  - There are nearly 600 million bogon addresses which occupy 13.8% of the entire IPv4 address space. These cannot post to EJMR but could be attributed to posts by mistake.
  - Our process attributes zero posts to bogon IP addresses.

• Time pattern of posters ✓
  - IP addresses post during the standard work and day time hours of their geolocation.

• Language of posters ✓
  - The dominant non-English language of the country of origin of the IP address is the country’s native language.
From Methods to Results

What tools does our analysis give us?

1. Panel dataset linking poster IPs across topics and time
2. Geolocation of posters (high quality at state and country level, can go down to city)
3. Information on ISPs (universities, organizations, corporations, hotels)

What questions are we asking?

- Who are the posters on this site?
  - “It’s just people at lower-ranked universities.”
  - “It’s just grad students.”
- Is the toxicity of the site widespread?
  - “It’s just a few bad apples.”
- Are there two dialogues (toxic and professional)? Are they separate? Within people? Within topic?
- Is there actually valuable inside information?
- Does EJMR make people more toxic?
- Are hundreds of thousands of visitors just paying attention to what a few people post?
Attribution and Geolocation

• 7.1 million posts in total → 4.7 million posts (66.1%) with attributed IP address

• 47,630 distinct IP addresses attributed to posts
  - These are the most frequent posters, but ...
  - ... there are many more infrequent posters, and ...
  - ... there are even more viewers (ratio of roughly 70 views to 1 post).

• Vast majority from countries with top research institutions in economics & finance
  - US (61.9%)
  - Canada (8.3%), United Kingdom (5.5%)
  - Australia (2.4%), Germany (2.2%), Hong Kong (1.9%), Italy (1.6%), France (1.5%)
  - Remaining share of geolocated posts (13.6%) from rest of the world
Time Pattern of Posts

- 7.1 million posts since December 2010
  - Average 70,000 monthly posts & 1,100 monthly unique IPs as of 2022

- Steady increase over time, but large traffic increase during COVID-19
  - Primarily driven by tripling in U.S.
  - Other countries experience more temporary increases
  - Very large rise in off-topic forum posts

- Strong cyclicality of job market posts
  - Disruption of job market cyclicality starting in 2021
Posts by City

- Majority of posts come from large cities in the US
  - Chicago, New York, Philadelphia

- Some fraction from large cities outside the US
  - Hong Kong, London, Montreal, Toronto

- Smaller US cities with leading research institutions
  - Cambridge, Berkeley
University IP Addresses on EJMR

- 10.9% of allocated posts originate from IP addresses of universities or research institutions

- Contributors are using university networks to post on EJMR

- Posts come from top US universities (top 25 econ departments)
Are these the elites of economics?

Number of Weekly Sonesta Posts concentrate in July at the Royal Sonesta

NBER SI was online-only due to Covid-19.
<table>
<thead>
<tr>
<th>University ISP</th>
<th>Harvard</th>
<th>MIT</th>
<th>Stanford</th>
<th>Berkeley</th>
<th>UChicago</th>
<th>Yale</th>
<th>NYU</th>
<th>NWU</th>
<th>Columbia</th>
<th>UPenn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harvard</td>
<td>7.9</td>
<td>9</td>
<td>5.2</td>
<td>1.4</td>
<td>3.7</td>
<td>2</td>
<td>2.1</td>
<td>0.9</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td>MIT</td>
<td>4.7</td>
<td>9.8</td>
<td>6</td>
<td>0.9</td>
<td>2.6</td>
<td>0.4</td>
<td>2.6</td>
<td>1.7</td>
<td>2.1</td>
<td>1.3</td>
</tr>
<tr>
<td>Stanford</td>
<td>4.4</td>
<td>6.4</td>
<td>7.4</td>
<td>1.7</td>
<td>4.5</td>
<td>1.2</td>
<td>1.9</td>
<td>1.4</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>UC Berkeley</td>
<td>1.6</td>
<td>3.7</td>
<td>1.7</td>
<td>4.2</td>
<td>1.8</td>
<td>1.2</td>
<td>1.8</td>
<td>0.8</td>
<td>0.6</td>
<td>0.9</td>
</tr>
<tr>
<td>UChicago</td>
<td>2.1</td>
<td>4.8</td>
<td>1.4</td>
<td>0.7</td>
<td>8.3</td>
<td>0.8</td>
<td>1.6</td>
<td>0.7</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>Yale</td>
<td>1.5</td>
<td>3.8</td>
<td>0.9</td>
<td>0.7</td>
<td>1.8</td>
<td>3.4</td>
<td>1.3</td>
<td>0.5</td>
<td>0.4</td>
<td>1</td>
</tr>
<tr>
<td>NYU</td>
<td>2.5</td>
<td>4.6</td>
<td>3.1</td>
<td>0.7</td>
<td>1.9</td>
<td>1.1</td>
<td>5.8</td>
<td>1</td>
<td>1.1</td>
<td>2.4</td>
</tr>
<tr>
<td>Northwestern</td>
<td>2.5</td>
<td>4.1</td>
<td>1.8</td>
<td>1.1</td>
<td>2.3</td>
<td>1.6</td>
<td>2.7</td>
<td>3.5</td>
<td>0.8</td>
<td>1.1</td>
</tr>
<tr>
<td>Columbia</td>
<td>3</td>
<td>4.9</td>
<td>2.3</td>
<td>1.3</td>
<td>2.8</td>
<td>1.6</td>
<td>3.1</td>
<td>1</td>
<td>5</td>
<td>2.4</td>
</tr>
<tr>
<td>UPenn</td>
<td>2.2</td>
<td>3.5</td>
<td>1.8</td>
<td>0.9</td>
<td>2.4</td>
<td>1</td>
<td>3.1</td>
<td>1.1</td>
<td>0.3</td>
<td>5.1</td>
</tr>
<tr>
<td>Others</td>
<td>1.1</td>
<td>3.8</td>
<td>0.5</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
<td>1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Everybody on EJMR loves talking about themselves ... and about MIT.
Distribution in Toxic Speech across IPs

Average IP has 15% of posts labeled toxic/hate speech/misogynistic.
Distribution in Participation in Toxic Conversation

Average IP has 53.68% of posts in topics with at least one post labeled toxic/hate speech/misogynistic.

Share of Attributed Posts Posting in Topics with Posts Labeled Toxic/Hate Speech/Misogynistic

Density

IPs with 10 or more posts attributed
Average University IP has 12.92% of posts labeled toxic/hate speech/misogynistic
Average Non-University IP has 15.17% of posts labeled toxic/hate speech/misogynistic

Density
University IP
Non-University IP
IPs with 10 or more posts attributed
Average University IP has 53.12% of posts in topics with at least one post labeled toxic/hate speech/misogynistic.

Average Non–University IP has 53.77% of posts in topics with at least one post labeled toxic/hate speech/misogynistic.

Share of Attributed Posts Posting in Topics with Posts Labeled Toxic/Hate Speech/Misogynistic

IPs with 10 or more posts attributed
Toxic Speech on EJMR by University ISPs

Hong Kong University of Science and Technology
Yale University
University of Washington
University of Chicago
University of Notre Dame
James Madison University
Columbia University
Stanford University
The Pennsylvania State University
University of Hawaii
Sam Houston State University
Universidad de Piura
Northern Arizona University
Chapman University
University of Arkansas at Little Rock
Hong Kong University of Science and Technology
Yale University
University of Washington
University of Chicago
University of Notre Dame

Note: Only universities with more than 100 posts are shown.
Anonymity and Identity Online
Florian Ederer, Paul Goldsmith-Pinkham & Kyle Jensen

We look forward to feedback and comments.

- florian.ederer@gmail.com
- paul.goldsmith-pinkham@yale.edu
- kyle.jensen@yale.edu

- nodes = EJMR IP addresses
- color = mean year of activity for IP
- edges = user interaction (Adamic/Adar)
Appendix
From usernames to IP addresses (in practice)

• Hash inversion is conceptually simple but intolerably slow in practice.
  - 695,364 topics × 2^{32} possible IPv4 addresses ≈ 3 quadrillion hashes
  - Check which hashes correspond to observed topic-username combinations
  - Easy to write code in Python, but would take over 60 years on a modern CPU

• Computation is tractable with graphical processing units (GPUs).
  - Basically the same as Bitcoin mining
  - Hash inversion is highly parallelizable
  - 230 hours of total computing time on Nvidia A100 devices each with 6,912 cores

• Hash inversion is only feasible because there is no salt.
From usernames and topics to IP addresses

```
IPv4 Address
Topic ID
Random Salt
```

```
"Hash"
SHA-1
pos 10-13
Username
```

```
Username
Topic ID
```

```
"Inverse Hash"
Set of ≈ 65k matching IPs
```

Known many-to-one function

\[
g(Topic\ ID, IPv4\ Address) = Username
\]

Set of matching IPs

\[
g^{-1}(\text{Topic ID}, Username) = \text{Set of matching IPs}
\]

Inverse function returns a set

IPv4 Address ∈ Set of matching IPs
Finding Active IP Addresses

• Statistical properties
  - The noise IPs that match a topic-username \((t, u)\) are **uniformly** distributed across the IPv4 space.
  - The probability that a noise IP \(a\) is observed in any particular set \(A_{(t,u)}\) containing IPs that match \((t, u)\) is a **hypergeometric** distribution which depends on the number of \(u\) in \(t\).
  - The number of times \(n_a\) that a noise IP \(a\) is observed across all matching IP sets \(A_{(t,u)}\) follows a **Poisson binomial** distribution.

• Our approach
  - Null hypothesis that observed counts of an IP \(a\) are generated purely by noise
  - Calculate probability under the null hypothesis that an IP \(a\) would be observed \(n_a\) times *by chance* (i.e., calculate p-values for each \(a\))
  - Rejection of null hypothesis \(\rightarrow\) IP \(a\) is active
Attributing Active IP Addresses to Posts

• IP address Attribution as an optimization problem
  - Find a small covering set of active IPs for the observed posts.
  - Compute p-values for each \( a \) in the set \( A_{(t,u)} \) for all \((t, u)\).
  - For each \((t, u)\), identify the IP with the lowest p-value.
  - If its p-value < \( p^* \), attribute to that IP. Otherwise, leave \((t, u)\) unattributed.

• Simple intuition of this optimization approach
  - 65k IP addresses can explain any given post, but imagine one of these IPs also explains many posts in other topics around the same time.
  - What is the likely origin IP of this post? It’s probably this highly explanatory IP.
Algorithmic IP address attribution approach

• Key assumptions
  - Sparsity of IP posters + uniformity of hash over full IPv4 space

• Potential issues with current approach
  1. Multiple-hypothesis testing problem / inference on winners
  2. A given IP will show up \( N \pi_0 \) times randomly, even under the null hypothesis.

• Solutions
  1. Choose conservative p-value thresholds \( p^* \approx 10^{-11} \) based on known null distribution
  2. Window the data in relatively short time intervals (7 days, 31 days, 91 days)

• Work in progress
  - Generative model to construct probability statements for each post and IP combination
Choosing the p-value threshold $p^*$

- Use a wrong substring of SHA-1 hash to construct a pure noise baseline
  - Repeat the entire hash inversion with incorrect hash positions.
  - Compute p-values and attribute post to IP with lowest p-value.
  - Calculate $p^*$ such that we would obtain zero attributions of posts to an IP

- Window-specific p-value thresholds $p^*$

$$
\begin{align*}
  p_{7d}^* &= 1.37 \times 10^{-10} \\
  p_{31d}^* &= 2.51 \times 10^{-11} \\
  p_{91d}^* &= 1.39 \times 10^{-11}
\end{align*}
$$

- With these $p^*$, the number of IP addresses that never posted to EJMR but that we mistakenly attribute to any of the roughly 7 million posts is, in expectation, less than one.
Summary of Hash Inversion and IP Attribution Steps

1. Create **topic**-specific **usernames** from hashes of all possible **topic-IP** pairs

2. For each **topic-username** find the set of **matching IPs**

3. Evaluate which **matching IPs** occur “much too often” than expected by random chance (“active IPs”) in a short time window (7 days)

4. Attribute **active matching IP** with lowest p-value < $p^*$ to post or leave post unattributed
Detecting Hash Changes

- Hash changed on July 8, 2013
- Average minimum p-values are much lower for the correct hash
- Average minimum p-values of incorrect hashes closely track each other
United States and Other Major Countries

Date of Posting

Monthly Posts

Grouped Forum
- Economics
- Job Market Rumors
- Off-Topic/Non-Econ
66.1% of posts come from 47,630 IPs.

These power posters fit stretched exponential.

Long tail of occasional posters is “unobserved.”

582,541 IPs are predicted to have posted on EJMR at least once.
Linguistic Analysis

- EJMR posters use l33tspeak and obfuscation to escape automatic EJMR moderation
  - “Hey a$$h01e, I left you a message earlier too. I will be there in Boston to FIEK and RAEP you, so cover your $hitty a$$ and your mouth now.” (2014-12-26)
  - “Mold-fa//g//g//ot, I will split your a//s/s in two with my HUMONGOUS super HARD shalong. You will be squealing like the little beia/tch that you are.” (2020-01-28)
  - “those d4mn j3ws had no morals either.” (2022-08-13)

- Deobfuscation process
  - Collect high-frequency non-English words in English posts to deobfuscate some of the most commonly obfuscated terms
  - Remove common symbol-based obfuscations
  - Transform leetspeak to its canonical form

- Run each post through a number of transformer-based machine learning models for toxicity, sentiment, and misogyny
  - Cross-validate with data by Wu (2020)
  - Note: potential classification error, but unlikely to be correlated across IPs
Median IP has 6.45% of posts labeled hate speech/misogynistic
Distribution in Participation in Hateful/Misogynistic Conversation

Average IP has 28.34% of posts in topics with at least one post labeled hate speech/misogynistic.
Do frequent IP posters post more toxic speech? Not really.

\[ y = 0.092 + 0.014 \times \log_{10}(N) \]

Note: Only IPs with more than 100 posts are shown.
Toxic Speech on EJMR by University ISPs

\[ y = 0.066 + 0.014 \times \log_{10}(N) \]

Note: Only IPs with more than 100 posts are shown.