

Social Bots

Prof. Dr. Stefan Evert und Philipp Heinrich
Professur für Korpuslinguistik



@schtepf
@ausgaeraechnaet
@ccl_erlangen



Aktuelles Forschungsthema

- **Projekt Exploring the Fukushima Effect**
 - gefördert durch die *Emerging Fields Initiative* der FAU
 - Korpuslinguistik, Japanologie I, LGDV, KoWi
 - Identifikation und Analyse der Ausbreitung von **Diskursen** in der algorithmischen Öffentlichkeit
 - Korpora: Zeitungen und Twitter
- **Algorithmische Öffentlichkeit** = transnationale Semi-Öffentlichkeit der sozialen Netzwerke
 - werden durch Algorithmen gesteuert („Filterblase“)
 - komplexe Interaktion mit Öffentlichkeit der Massenmedien
 - Spielfeld für politische Propaganda & Fake News

Politische Propaganda

- Zehn Millionen **Facebook**-Nutzer haben im Vorfeld zur US-Wahl 2016 politische Anzeigen gesehen, die mutmaßlich aus Russland finanziert wurden (Quelle: [tagesschau.de](#), 07.09.2017)
- **Twitter** erklärt, eine russische Mediengruppe habe 2016 in dem Kurzmitteilungsdienst fast 2000 Anzeigen oder promotete Tweets gekauft, die zur Beeinflussung der US-Wahl gedacht gewesen sein könnten. (Quelle: [tagesschau.de](#), 29.09.2017)
- Und auch **Google** hat Anzeigen auf YouTube und bei anderen Produkten des Konzerns entdeckt. Dafür seien nach ersten Erkenntnissen Zehntausende Dollar ausgegeben worden. (Quelle: [tagesschau.de](#), 09.10.2017)



Donald J. Trump for President

Cambridge Analytica provided the Donald J. Trump for President campaign with the expertise and insights that helped win the White House, causing the most remarkable victory in modern U.S. political history.

Analyzing millions of data points, we consistently identified the most persuadable voters and the issues they cared about. We then sent targeted messages to them at key times in order to move them to action. All of this was achieved in a fraction of the time and at a much lower cost than was spent by our rivals.

With our help, your campaign can use these same intelligent targeting and sophisticated messaging techniques. Cambridge Analytica deployed three integrated teams to support the campaign: research, data science, and digital marketing.

Social Bots in sozialen Medien

- „**Gutartige**“ Social Bots erledigen automatisiert **formalisierbare Prozesse**:
 - Wetter
 - Public Relations
 - (personalisierte Werbung)
- „**Bösartige**“ Social Bots verfolgen **betrügerische Absichten**:
 - Desinformation
 - politische Propaganda
 - (tun so, als wären sie echte Menschen)

Identifikation von Social Bots

- ... **essentiell** für eine sinnvolle Interpretation von Daten aus sozialen Medien
- Mögliche Einflussnahme auf Nutzer und Data Analytics:
 - „Pushen“ von Themen durch Wiederholung
 - „Bandwaggoning“ (Verwendung hochfrequenter Hashtags)
- Annahme: Social Bots **verbreiten** Inhalte durch:
 - Retweeten von Originalen (einfach)
 - Versenden von **Nahduplicaten** (schwierig)

👉 Verteilung auf mehrere Nutzeraccounts

Nahduplicata und Normalisierung

Nutzer **topnewskoeln** am 29. Mai 2017, 17:12:30:

Allgäu Wellness Reise 2Tage 2Personen HP 4* Lindner
Park Hotel + Spa Oberstaufen <https://t.co/DXrkPUWZ6j>

Nutzer **tourismusvideo** am 30.05.2017, 15:08:53:

Allgäu Wellness Reise 2Tage 2Personen HP 4* Lindner
Park Hotel + Spa\nOberstaufen <https://t.co/dVEnKmBqMZ>

Nutzer **tourismusvideo** am 30.05.2017, 18:10:44:

Allgäu Wellness Reise 2Tage 2Personen HP 4* Lindner
Park Hotel + Spa\nOberstaufen <https://t.co/Cdwpr3TyI>



allgäuwellnessreisetagepersonenhlindnerpark
hotelspaoberstaufen

Nahduplikate und Normalisierung

Nutzer **DomeniktvNEWS** am 16.05.2017, 18:25:13:

Entwurf des SPD-Regierungsprogramms: Wie Schulz regieren will <https://t.co/4mHK6PO2Xy>

Nutzer **GTEM** am 16.05.2017, 18:30:36:

Entwurf des SPD-Regierungsprogramms: Wie Schulz regieren will <https://t.co/Q5jWYrDmrr>

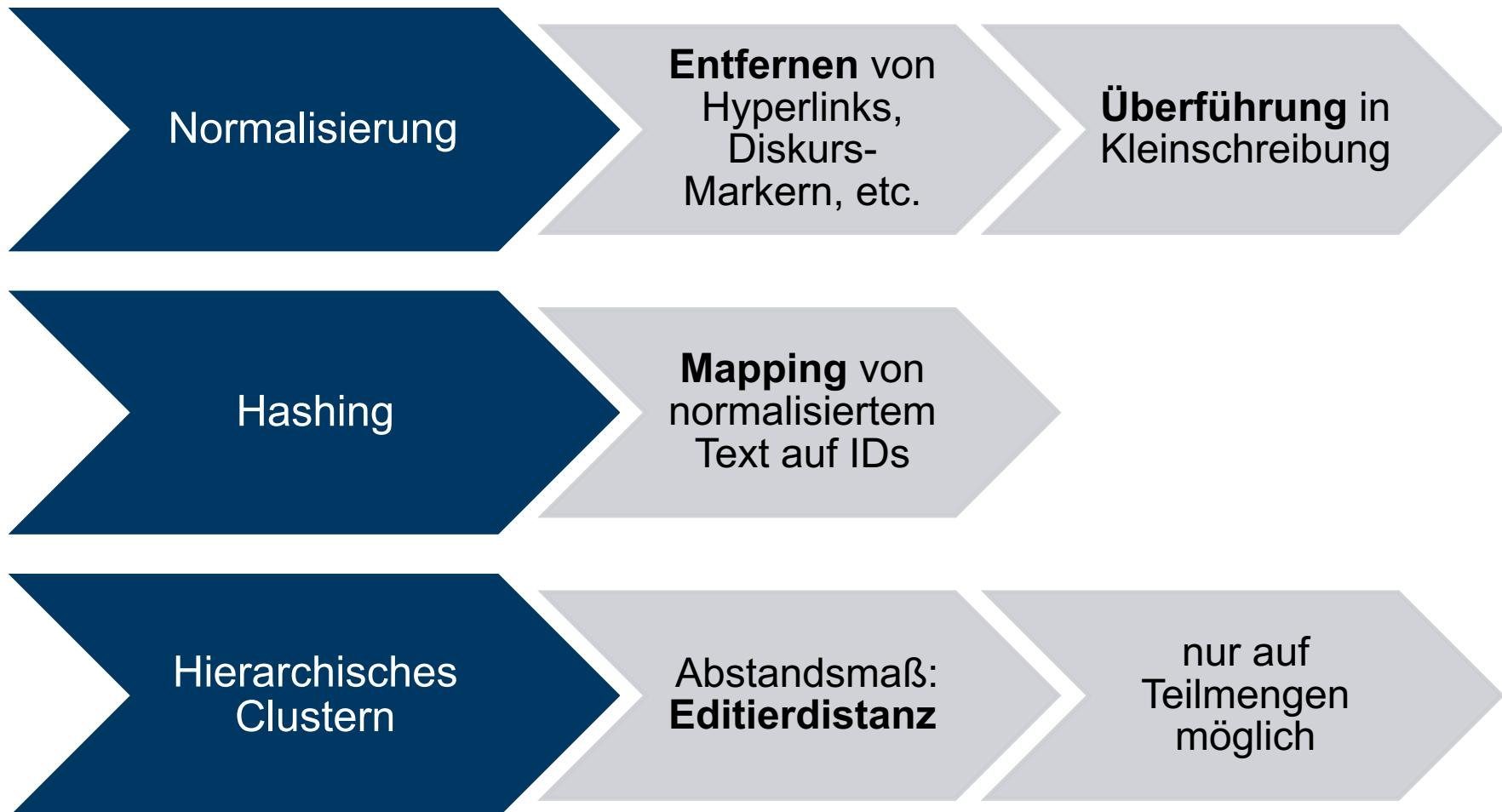
Nutzer **Augustinus_hipo** am 16.05.2017, 18:40:09:

Entwurf des SPD-Regierungsprogramms: Wie Schulz regieren will <https://t.co/Vif0oiLF4f> <https://t.co/DV8EBmiu3r>

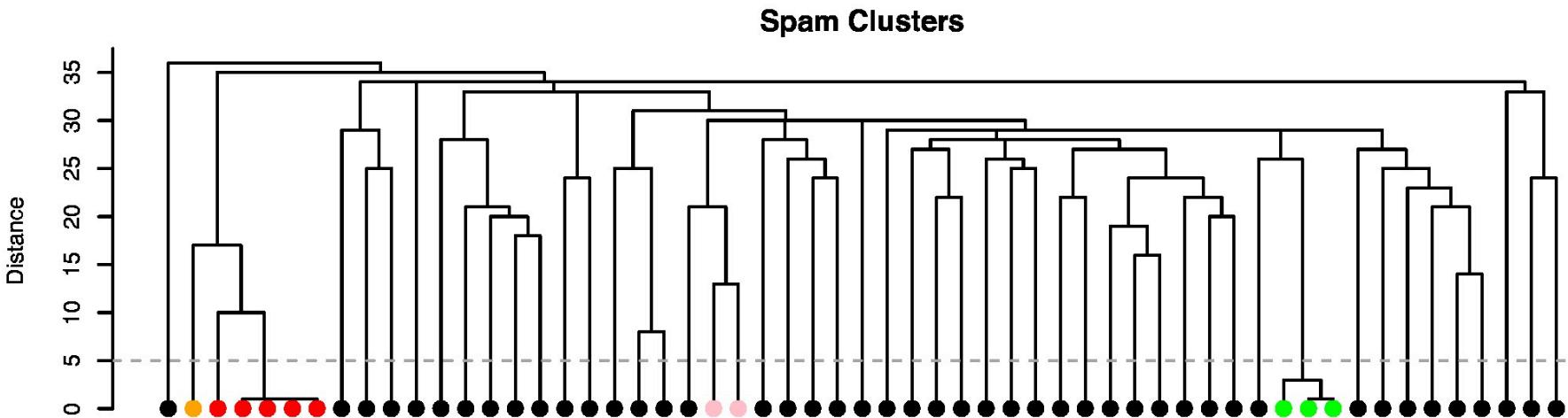


entwurfdesspdregierungsprogramms
wieschulzregierenwill

Nahduplicaterkennung



Hierarchische Cluster von Tweets eines identifizierten Social Bot



Twitter-Korpora

- **Datenakquise**
 - Stream-API (1%-Stichprobe aller Tweets)
 - Keyword-Filter
 - Einkauf (GNIP)
- **Datensätze für empirische Analysen**
 - Bundestagswahl 2013 (932.996 Originale)
 - davon 15,4% Duplikate
 - Shūgiin-Wahl 2014 (243.914 Originale; Gnip: 1.491.974)
 - davon 61,2% Duplikate
 - Bundestagswahl 2017 (5.717.127 Originale)
 - davon 19,7% Duplikate
 - Shūgiin-Wahl 2017 (408.439 Originale)
 - davon 30,6% Duplikate

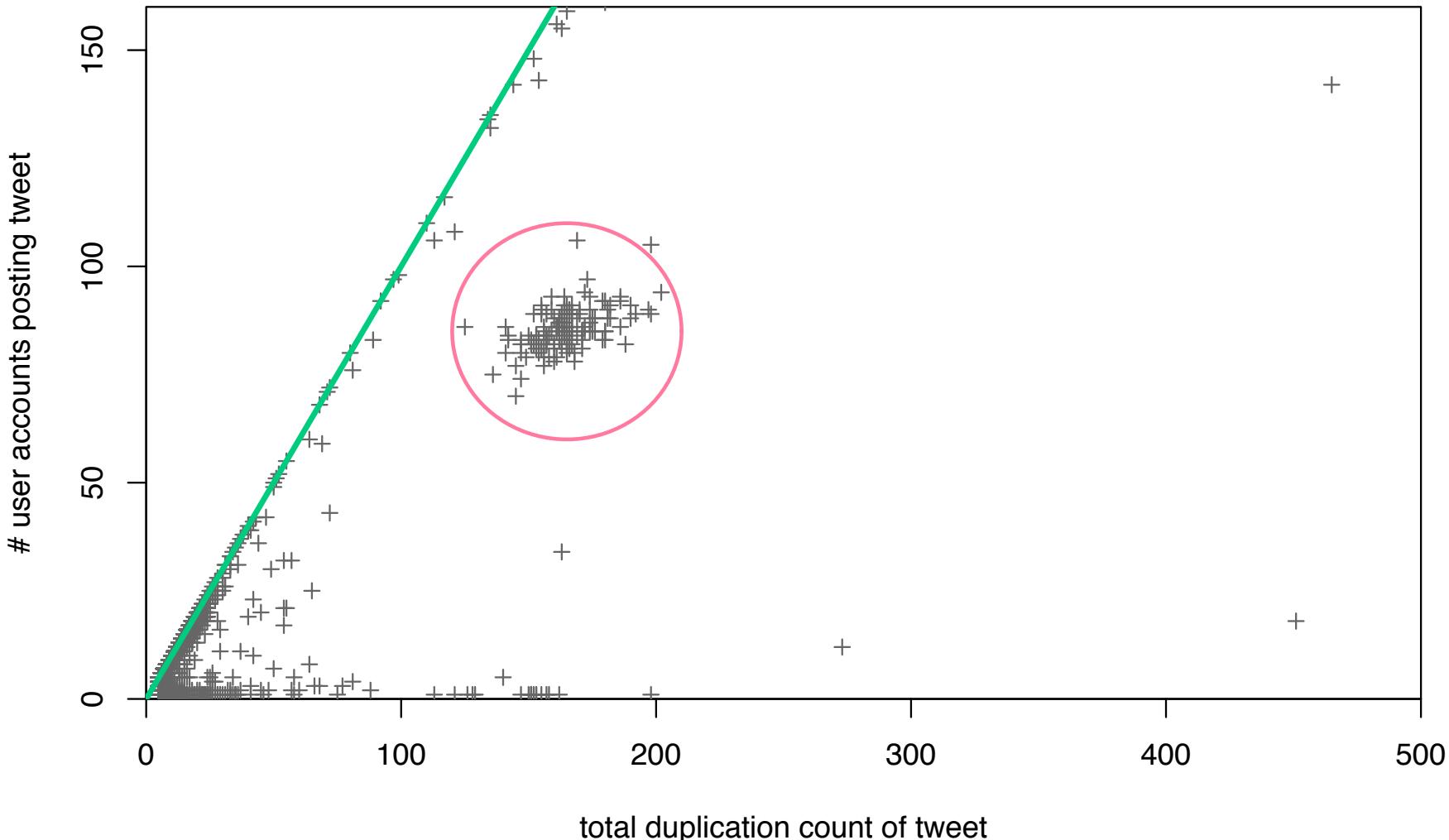
Hypothese

- Social Bots erstellen **Nahduplicaten**
 - freier Parameter: Anzahl Nutzeraccounts
 - Ähnlichkeit der Tweets reflektiert Intelligenz der Bots
- **Footprint** eines Botnetzes:
für jeden *unique tweet* = Gruppe von Nahduplicaten

$$\frac{\text{Anzahl Nutzeraccounts}}{\text{Anzahl Kopien}}$$

BT 2013: Visualisierung

Bundestagswahl 2013



Grüne Linie: „Share me on Twitter“

- #Merkel via @SPIEGELONLINE
- #SPD via @SPIEGELONLINE
- All registered officially a #CSU student !
- Das #TVDuell gewinnt @PeerSteinbrueck #Steinbrück via @focus_topnews
- Das #TVDuell gewinnt @PeerSteinbrueck #Steinbrück via @spiegel
- Geht wählen !!! #btw13
- Gewählt . #btw13
- Ich habe ein @YouTube-Video positiv bewertet: URL Unnütze Fakten über Wahlen ! #btw13 - WTF #31
- ...

Roter Kreis: Catbots



CatsForFunn
@CatsForFunn11

Follow

#spd #openruhr



6:41 PM - 15 Jun 2013



CatsForFun
@CatsForFun2

Follow

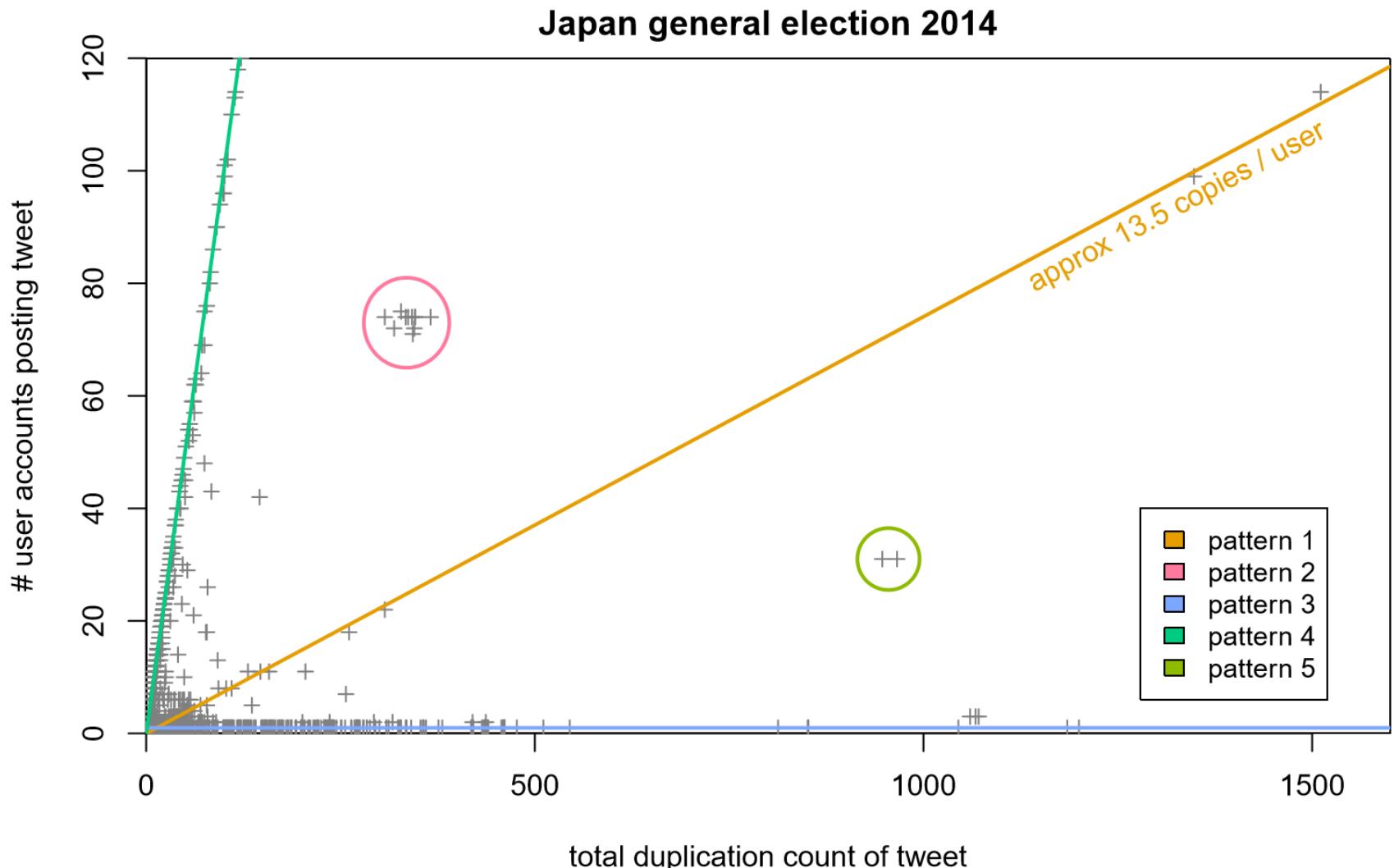
#openruhr #linke



9:29 PM - 15 Jun 2013



GE2014: Visualisierung



GE2014: Inhaltliche Analyse

- Muster 1 (gelbe Linie):
 - Abe-Unterstützungsbotnetz
 - Nutzeraccounts: @ABE_CRUSADER; @JNSC_ABEMAMOROU; @GOSPEL_ABERULER
- Muster 2 und 5 (Kreise):
 - Nutzeraccounts: @END_OF_JAP, @Wazaru_Messatsuas, @HorobiroWazaru, @KiesareWazaru, @WazaruBokumetsu
 - rechtes Propagandabotnetz?
 - Inhaltsanalyse der häufigsten Nahduplicata: Pro-Abe-Propaganda

GE2014: Inhaltliche Analyse

- Muster 3 (horizontale blaue Linie):
 - Hashtag „sex. Belästigung“ und „Auflösung“ in 9.065 Tweets von ein und demselben Fake-Account der LDP-Lokalpolitikerin Ayaka Shiomura, die im Wahljahr sexistischen Beschimpfungen im Lokal-parlament ausgesetzt war;
 - 99,4% davon sind Kopien (bzw. Nahduplicata) von 57 Original-Tweets, die jedoch so gut wie nie retweetet wurden;
 - dennoch lässt sich aufgrund des Inhalts und der verlinkten Videos eindeutig auf rechte Propaganda schließen, die versucht, *trending hashtags* im Wahljahr zu instrumentalisieren.
- Muster 4 (steile grüne Linie):
 - Share-Button des Onlinemagazins *Politis*

Pattern 3



The image shows a Twitter profile page. At the top, there's a large profile picture of a woman with short brown hair and bangs. Below it is a smaller thumbnail of the same woman. The profile name is "塩村あやかの本音" and the handle is "@Stupid00002". The bio text is in Japanese:

塩村文夏は旅館の女将になりたかったし、政治家は遊びでやつてます。放射能汚染した東京をぶつ壊わしたい。オリンピックなんか失敗すればいい。備掃社の塩村一ってダレよ? 本ア
力:twitter.com/shiomura

Below the bio is a blue button with the text "Tweet to 塩村あやかの本音" and a pen icon.

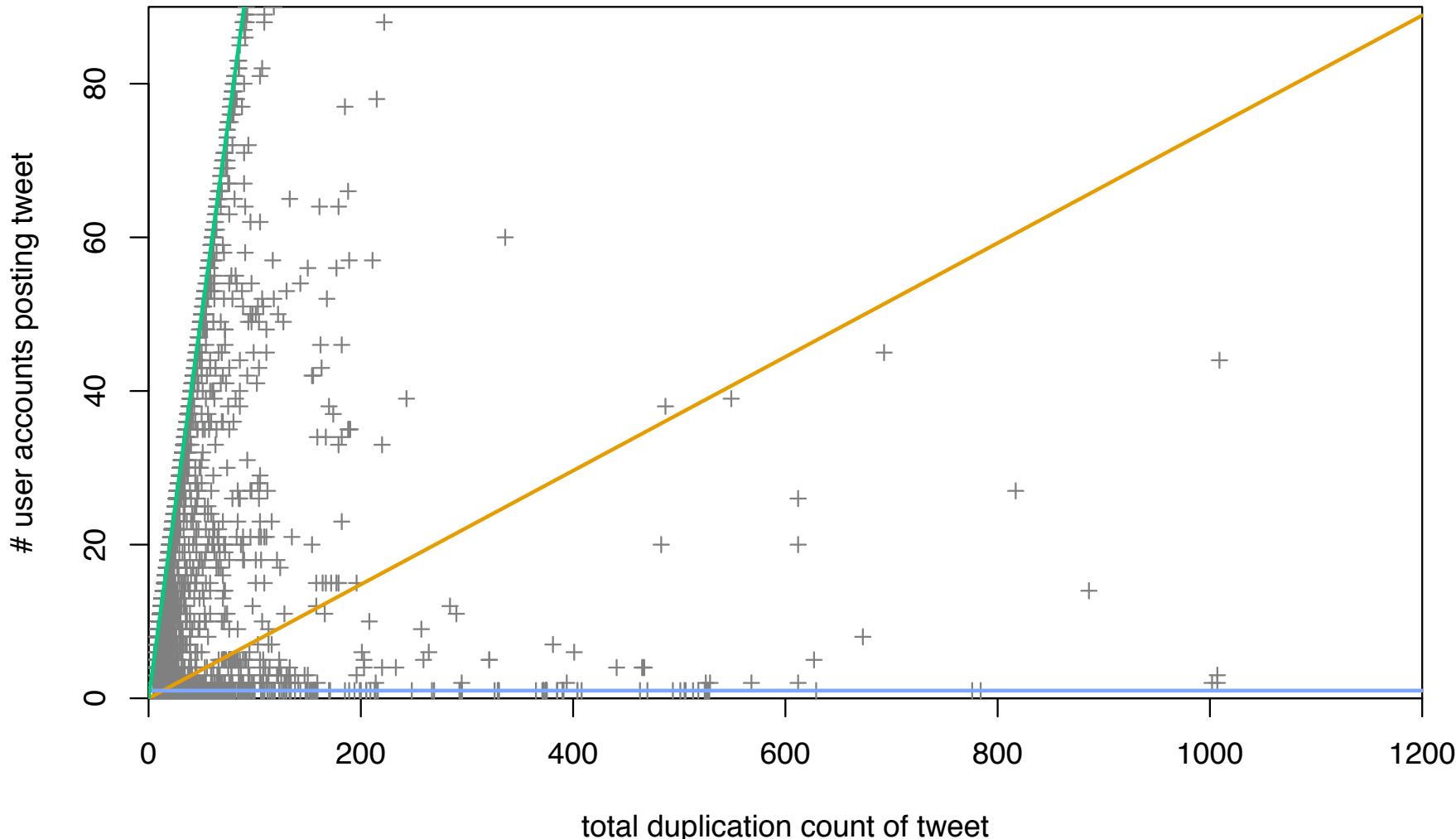
The main stats at the top are: TWEETS 305K, FOLLOWING 275, FOLLOWERS 311, and FAVORITES 5. There are also "Follow" and gear icons.

The timeline shows two tweets from the user:

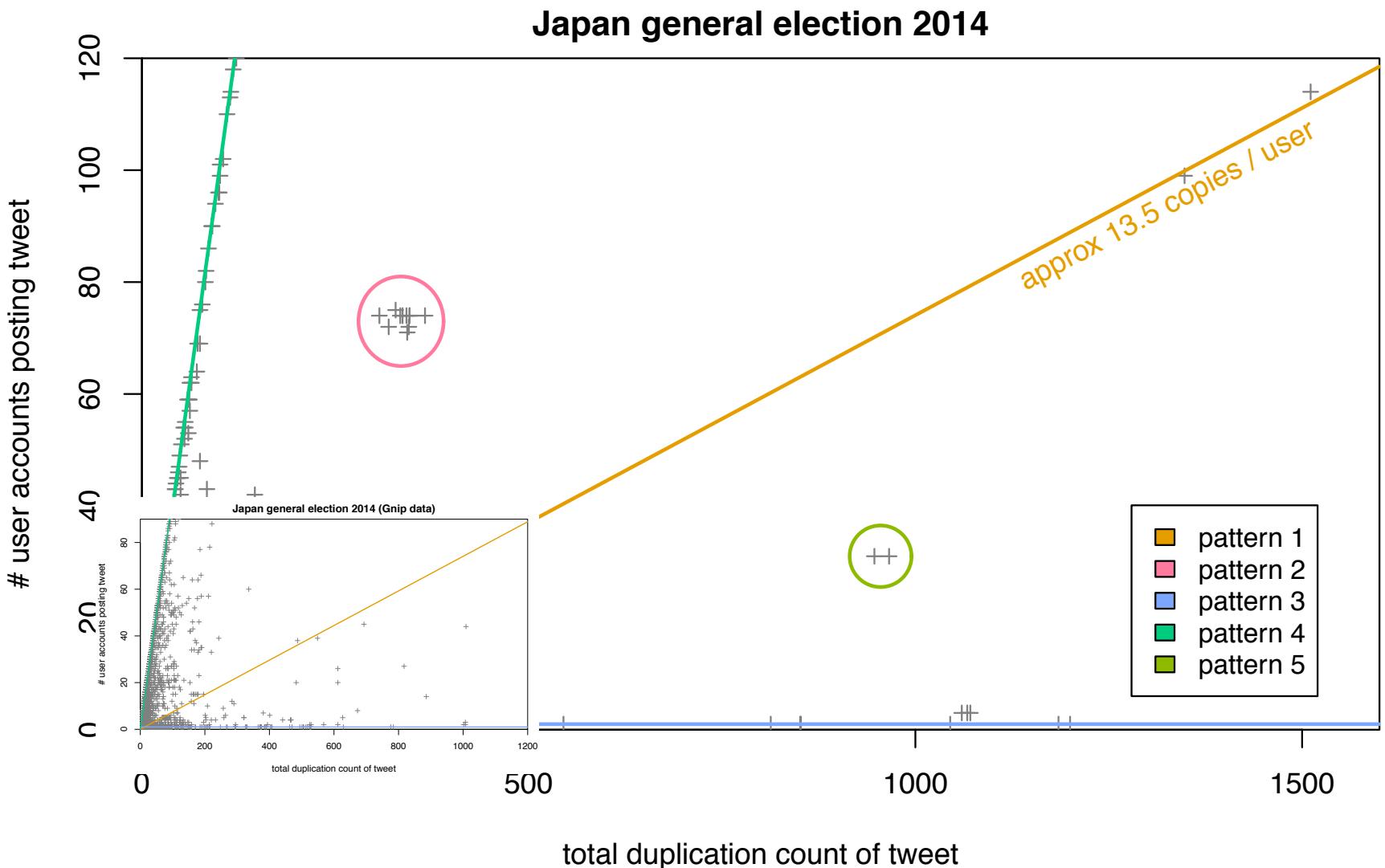
- Tweet 1:** 塩村あやかの本音 @Stupid00002 · Jun 17
#自民党 #セクハラ #解散 投票率5割なら舛添氏、7割なら“劇的な結果”
- PRESIDENT Online dlvr.it/BDI6Sb #Tokyo Radioactivity
[View summary](#)
- Tweet 2:** 塩村あやかの本音 @Stupid00002 · Jun 17
#自民党 #セクハラ #解散 安倍首相、解散決断で「民主党をぶっ潰す」
- PRESIDENT Online dlvr.it/BDI64p #Tokyo Radioactivity
[View summary](#)

Validierung: vollständige Daten von Gnip

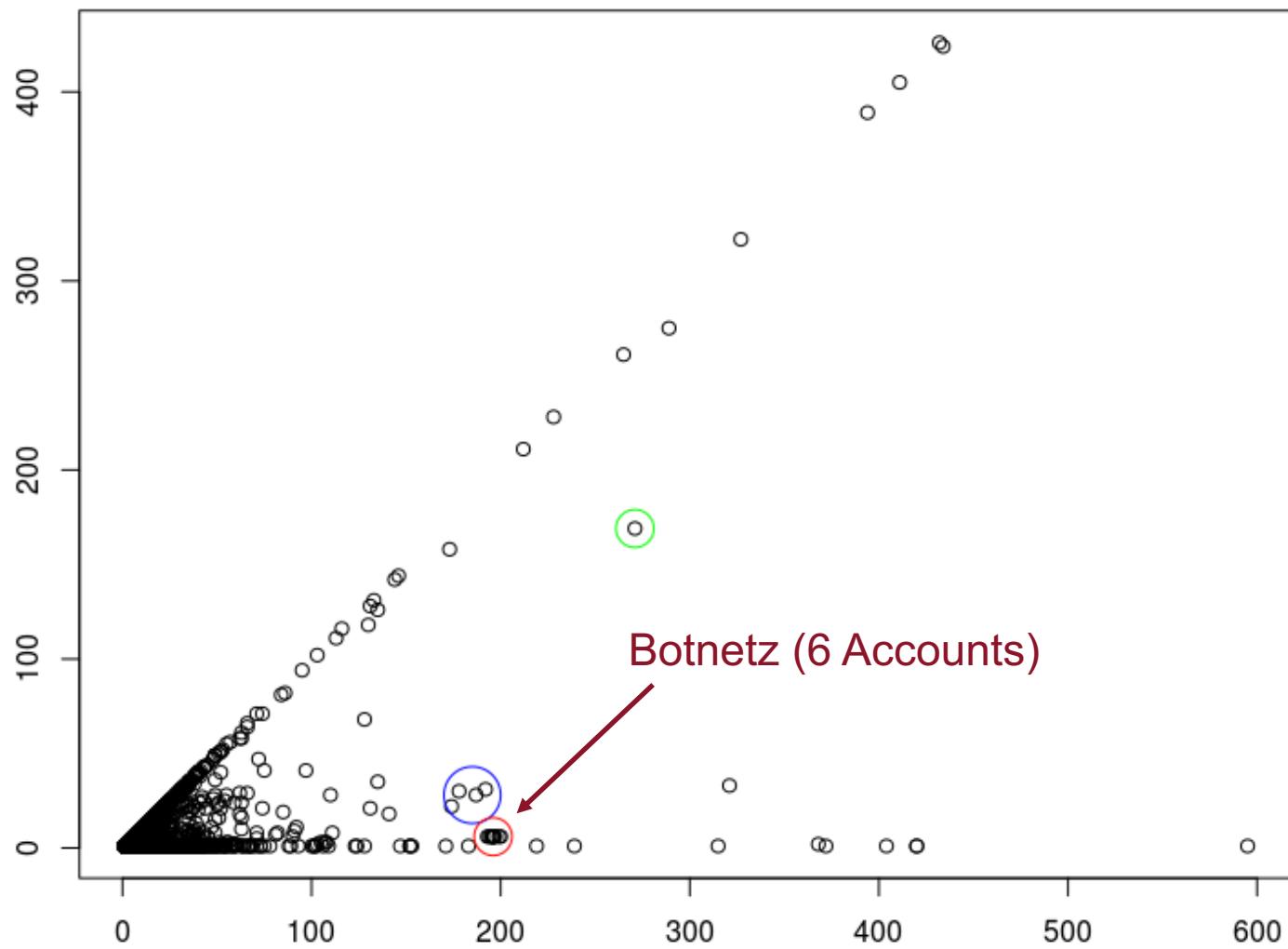
Japan general election 2014 (Gnip data)



Validierung: vollständige Daten von Gnip

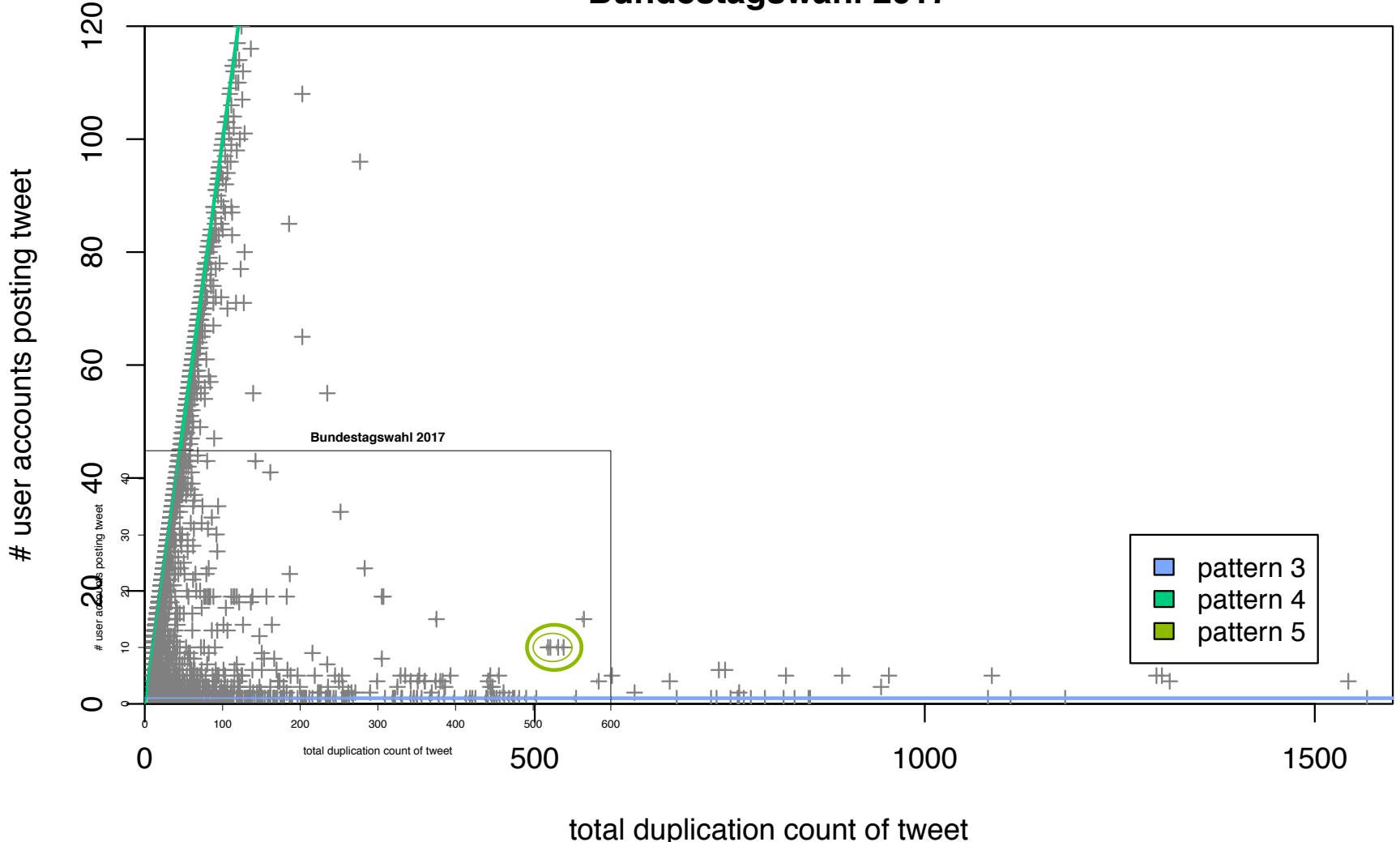


GE 2017 (Shūgiin-Wahl in Japan)



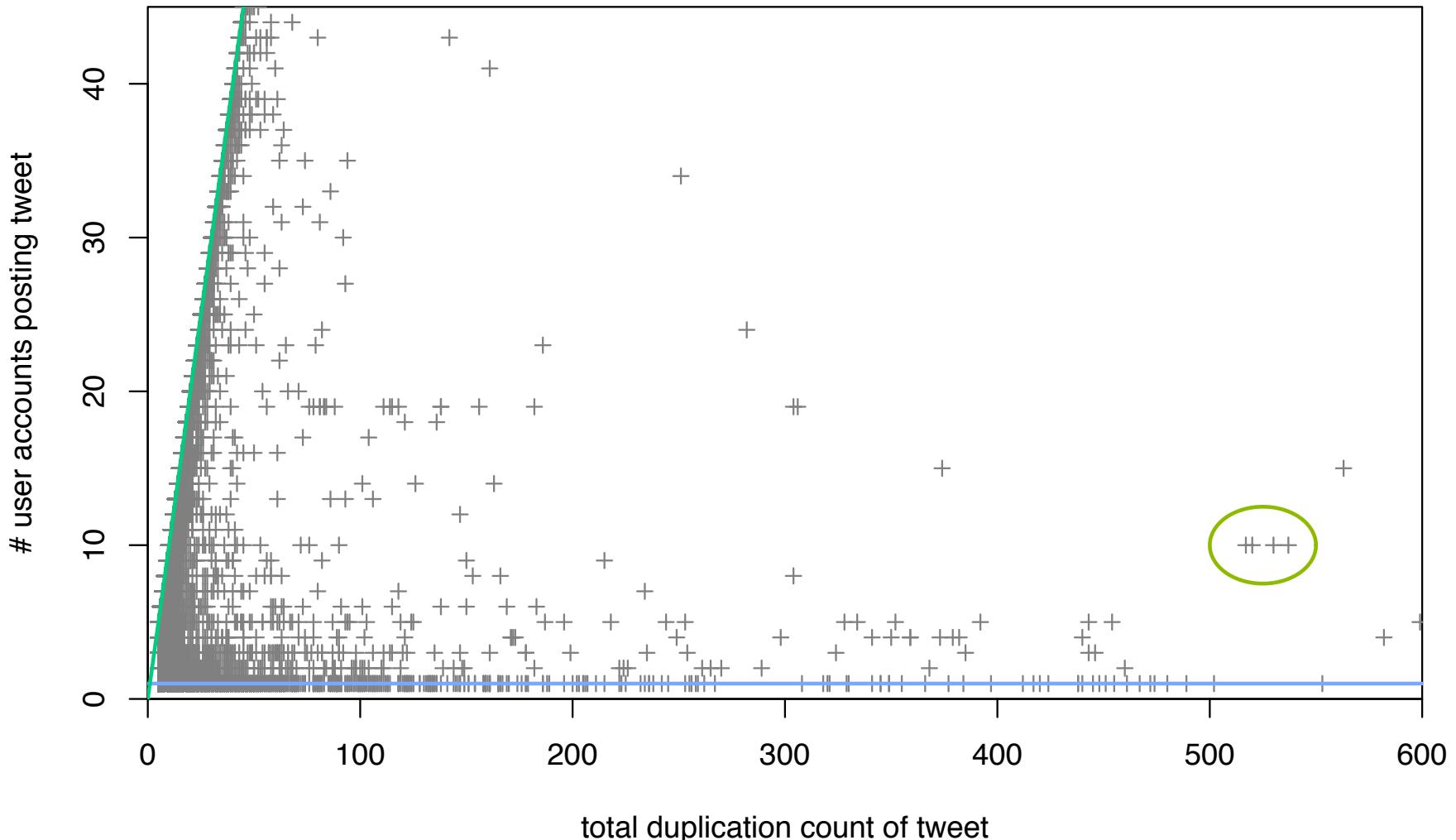
Bundestagswahl 2017

Bundestagswahl 2017



Bundestagswahl 2017

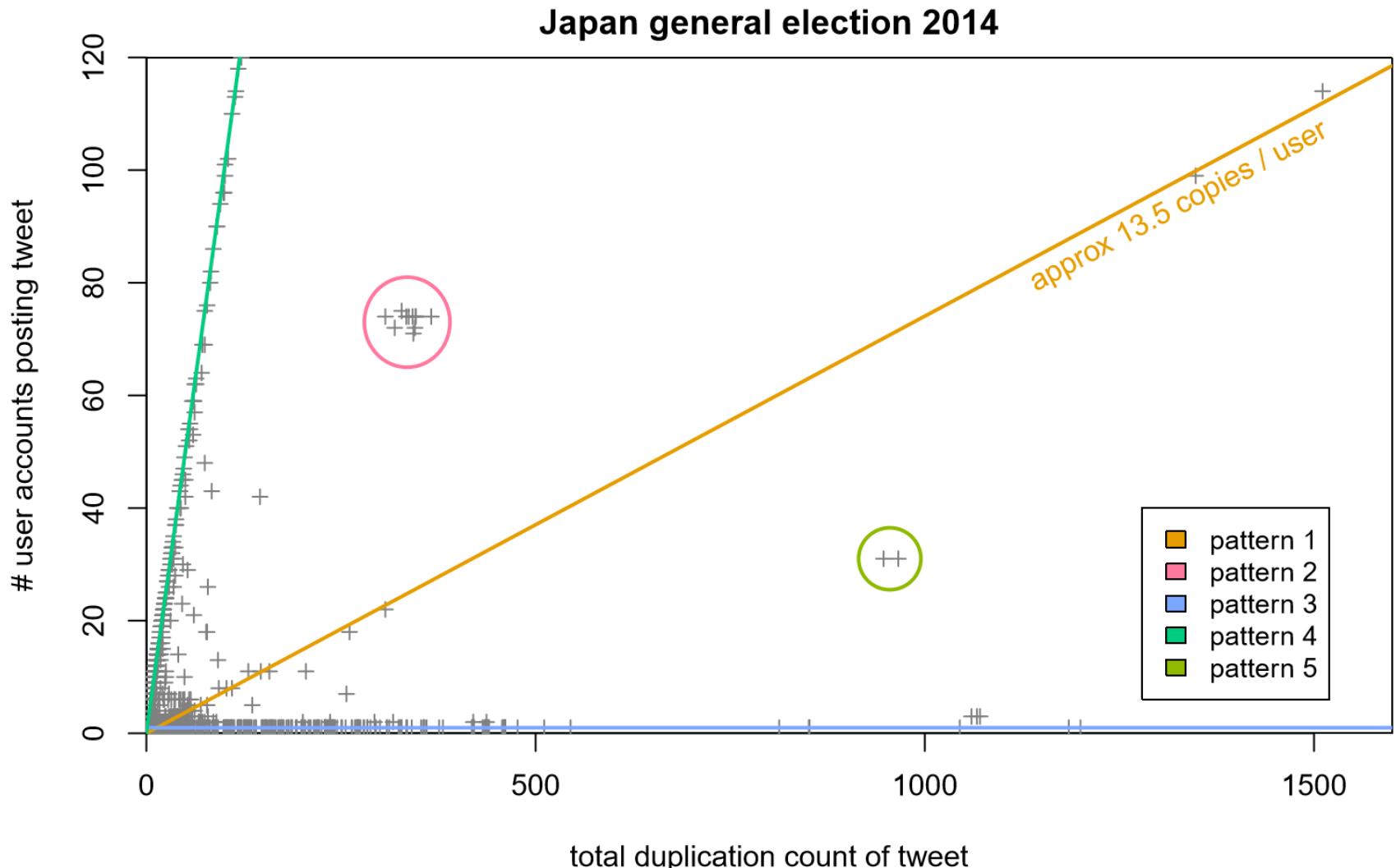
Bundestagswahl 2017



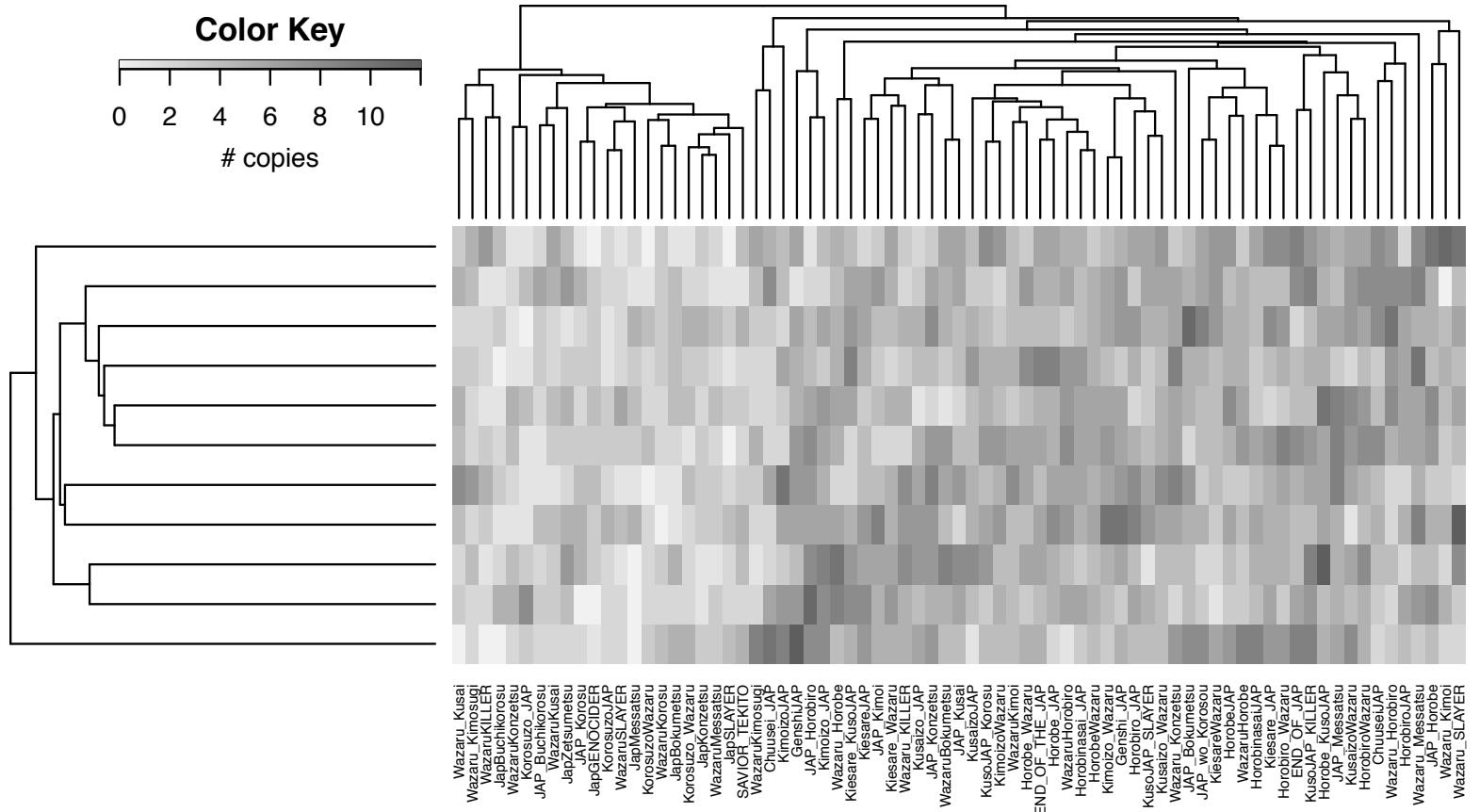
Sind Social Bots jetzt unsichtbar?

- Bisher: sehr „grobe“ Visualisierung
 - es wird nicht berücksichtigt, welche User-Accounts die auffälligen Tweets verschickt haben
 - d.h. die eigentlichen Bot-Accounts werden nicht als solche erkannt
- Verfeinerung: **Kookkurrenz-Matrix**
 - Zeilen = User-Accounts
 - Spalten = unterschiedliche Tweets
 - Sortierung mit *bond energy algorithm*
- Vermutung: Bot-Netze als „Blöcke“ erkennbar
 - alle Bots in einem Netz posten dieselben Tweets
 - legitime Nutzer posten diese „Spam“-Tweets nicht

GE2014: Visualisierung

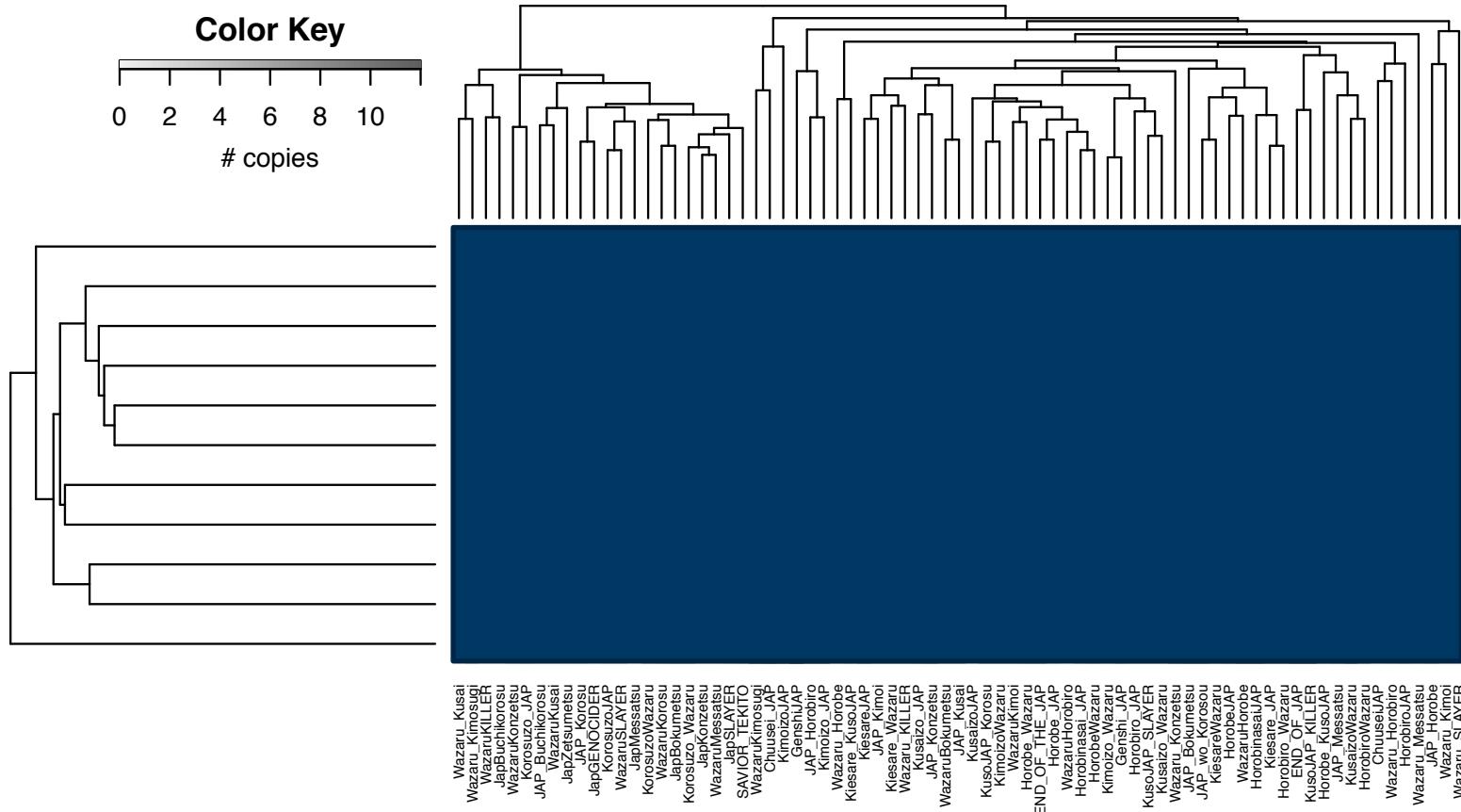


GE 2014: Pattern 2



Poisson-Regression:
keine signifikante Abweichung von gleichmäßiger Verteilung

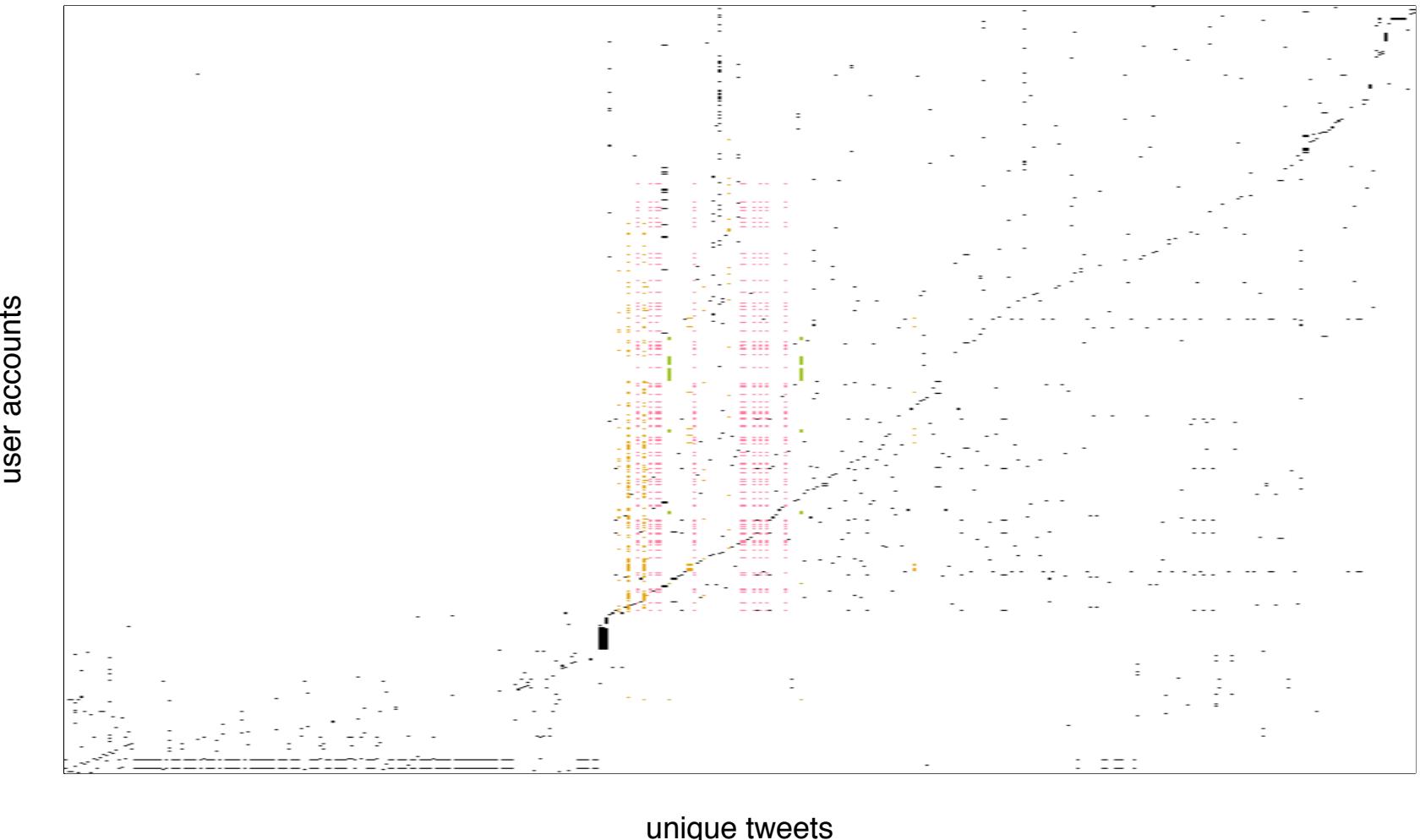
GE 2014: Pattern 2



Binäre Kookkurrenz-Matrix: Welcher Account hat welchen Tweet verschickt (unabhängig von Häufigkeit)?

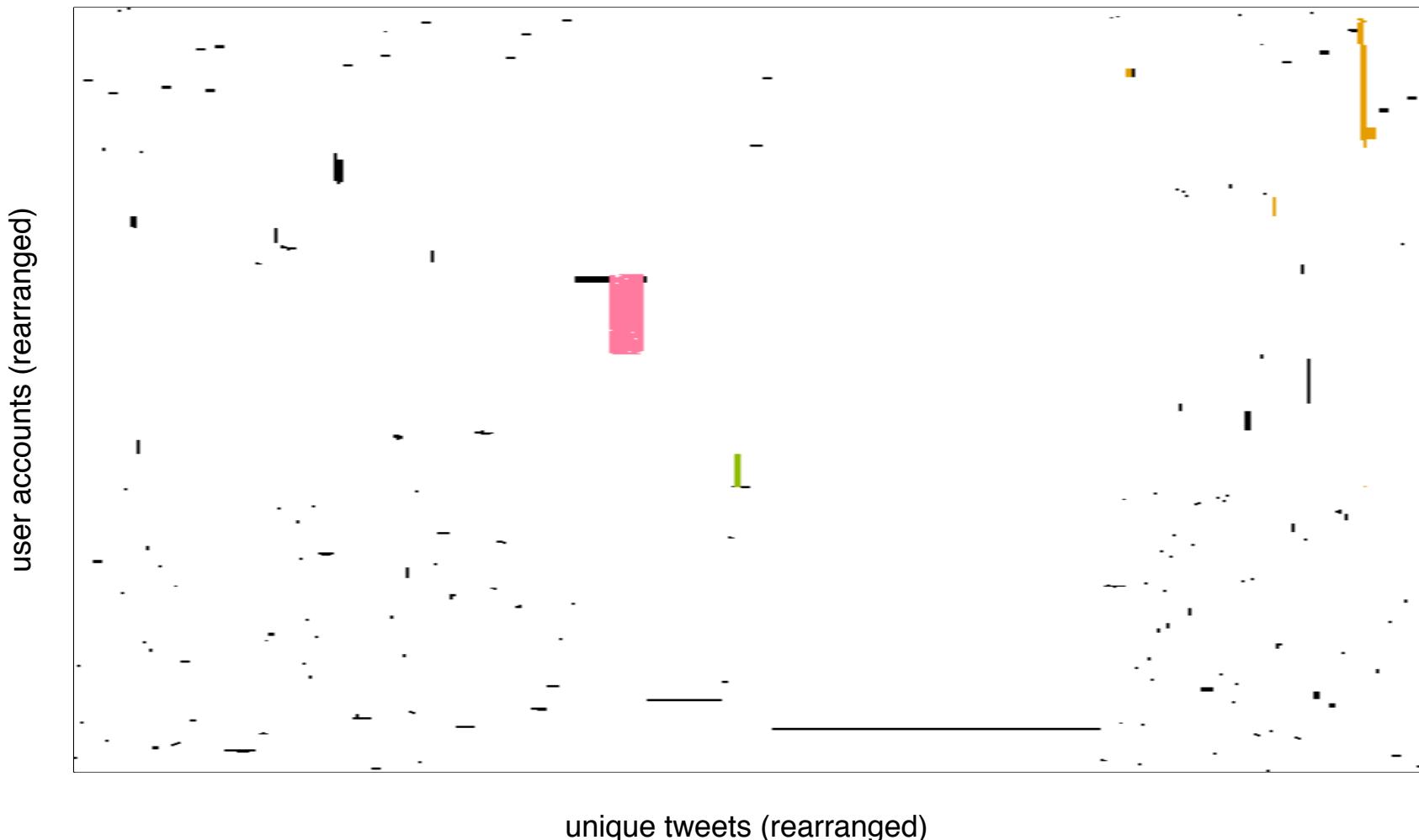
GE2014: Kookkurrenz-Matrix

Japan GE 2014: co-occurrence matrix



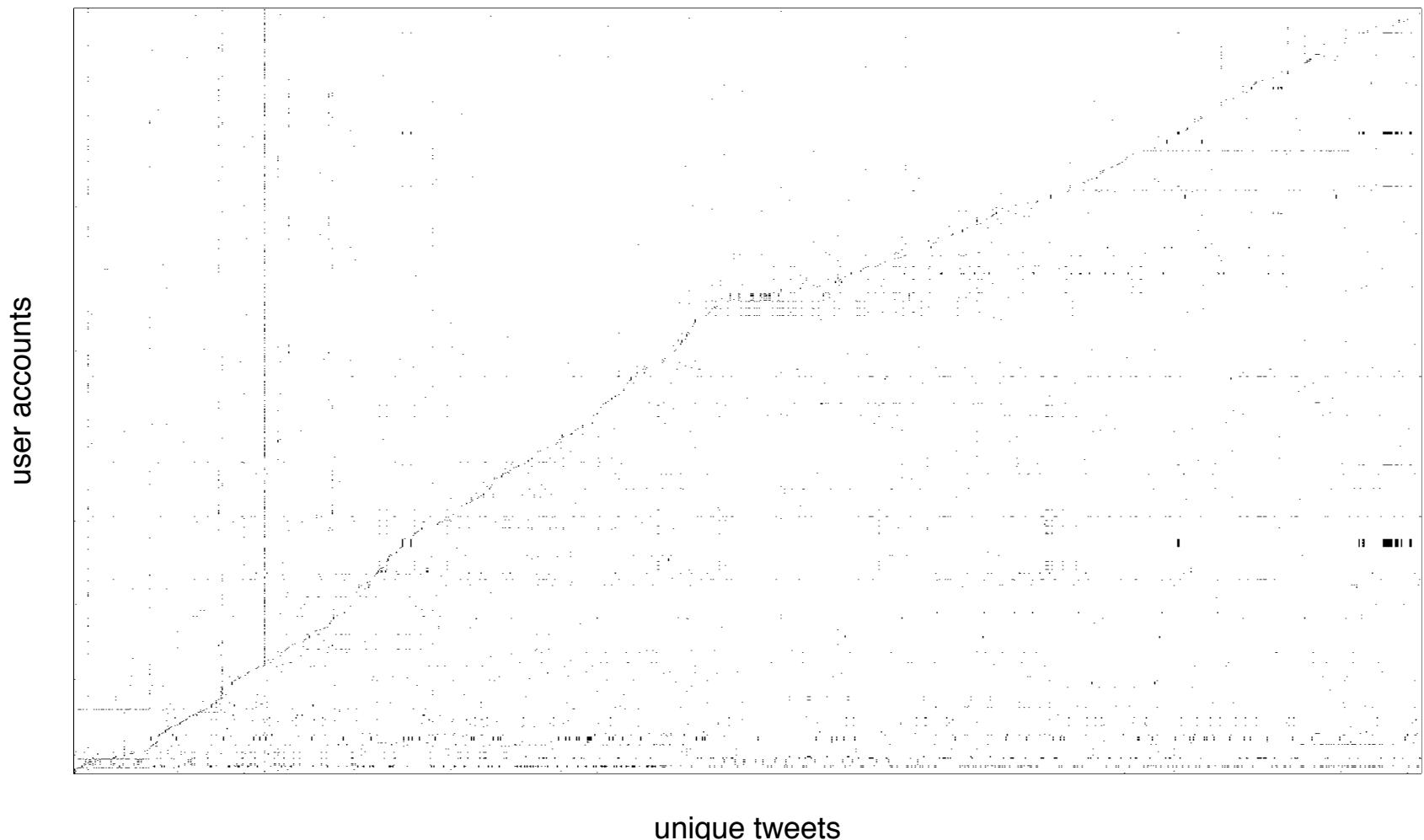
GE2014: Kookkurrenz-Matrix

Japan GE 2014: co-occurrence matrix



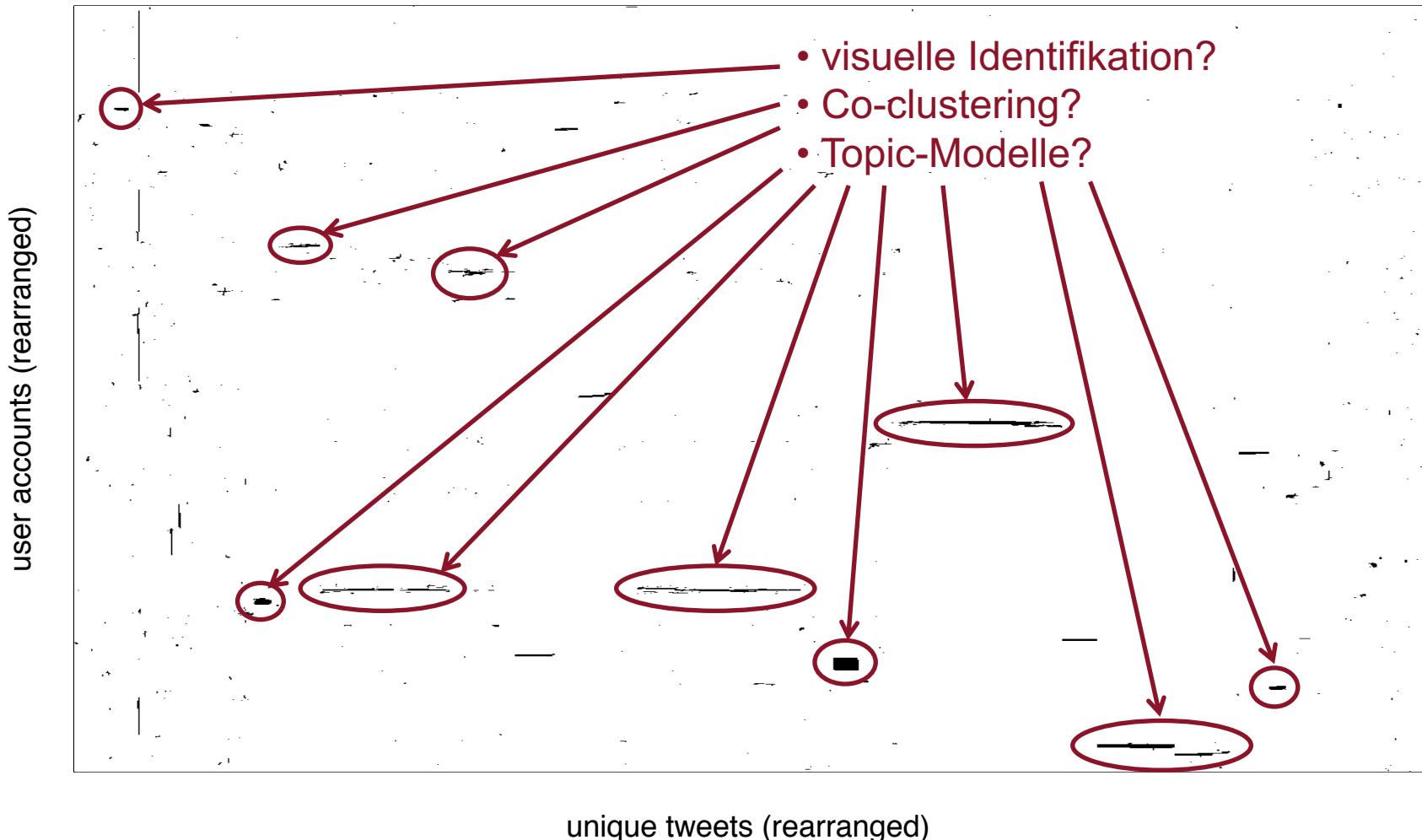
Bundestagswahl 2017: Kookkurrenz-Matrix

Bundestagswahl 2017: co-occurrence matrix



Bundestagswahl 2017: Kookkurrenz-Matrix

Bundestagswahl 2017: co-occurrence matrix



Zusammenfassung und Ausblick

- Automatisierte „Social Bots“ spielen eine immer größere Rolle in sozialen Medien
- Untersuchungsgegenstand hier:
Politische Einflussnahme durch Botnetze in Twitter
- Ansatz: Erkennen von Nahduplicaten
- Statistisch auffällige Muster → Botnetze
- Sozialwissenschaftliche Untersuchung der Mechanismen und Ziele dieser Botnetze
- Nächster Schritt: verfeinerte Ansätze zur Erkennung moderner Botnetze mit inhaltlicher Analyse

Vielen Dank für Ihre Aufmerksamkeit!



@schtepf
@ausgaeraechnaet
@ccl_erlangen