disinformation campaigns have principal-agent problems

Identify groups of accounts that coordinate messages, not individual accounts with bot-like behavior

Political astroturfing around the world

Franziska B. Keller[1] ✉ fbkeller@ust.hk
David Schoch[2] ✉ david.schoch@manchester.ac.uk
Sebastian Stier[3] ✉ sebastian.stier@gesis.org
JungHwan Yang[4] ✉ junghwan@illinois.edu

[1] *Hong Kong University of Science and Technology*
[2] *The University of Manchester*
[3] *GESIS–Leibniz Institute for the Social Sciences*
[4] *University of Illinois at Urbana-Champaign*

Political astroturfing
Centrally coordinated *disinformation campaign* in which participants pretend to be ordinary citizens acting independently as part of a genuine *grassroots movement*.
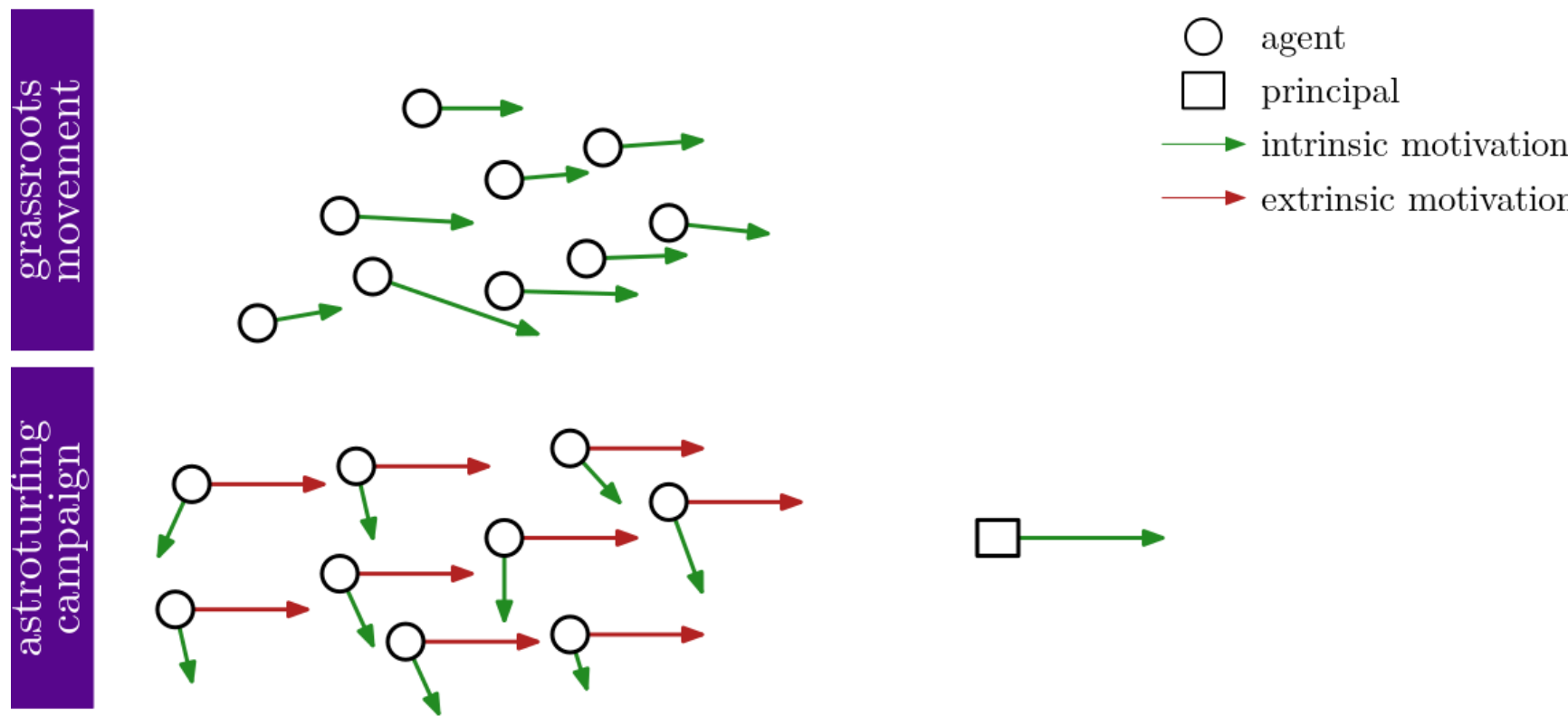
Data

Twitter is making archives of tweets and accounts that are believed to be part of state-backed disinformation campaigns publicly available.

	users	tweets	period
Catalonia	130	10,423	11/2011-04/2019
China	940	3,614,906	11/2010-07/2019
Germany	111	102,655	10/2015-10/2017
Iran	6,687	9,162,169	11/2010-04/2019
Russia	1,039	3,952,477	11/2010-06/2018
South Korea	813	193,773	06/2012-01/2013
USA	2,800	5,367,472	11/2010-11/2018
Venezuela	1,951	9,935,649	11/2010-12/2018

Principal-Agent Problem

Campaign organizers (Principal):
Hires agents who should produce many genuine-sounding messages.

Astroturfing agents:
Not intrinsically motivated and try to minimize work.

Legend: ○ agent · □ principal · → intrinsic motivation · → extrinsic motivation · grassroots movement · astroturfing campaign

Co-tweeting
Copy&pasting tweets across several accounts within a short time period.

Co-retweeting
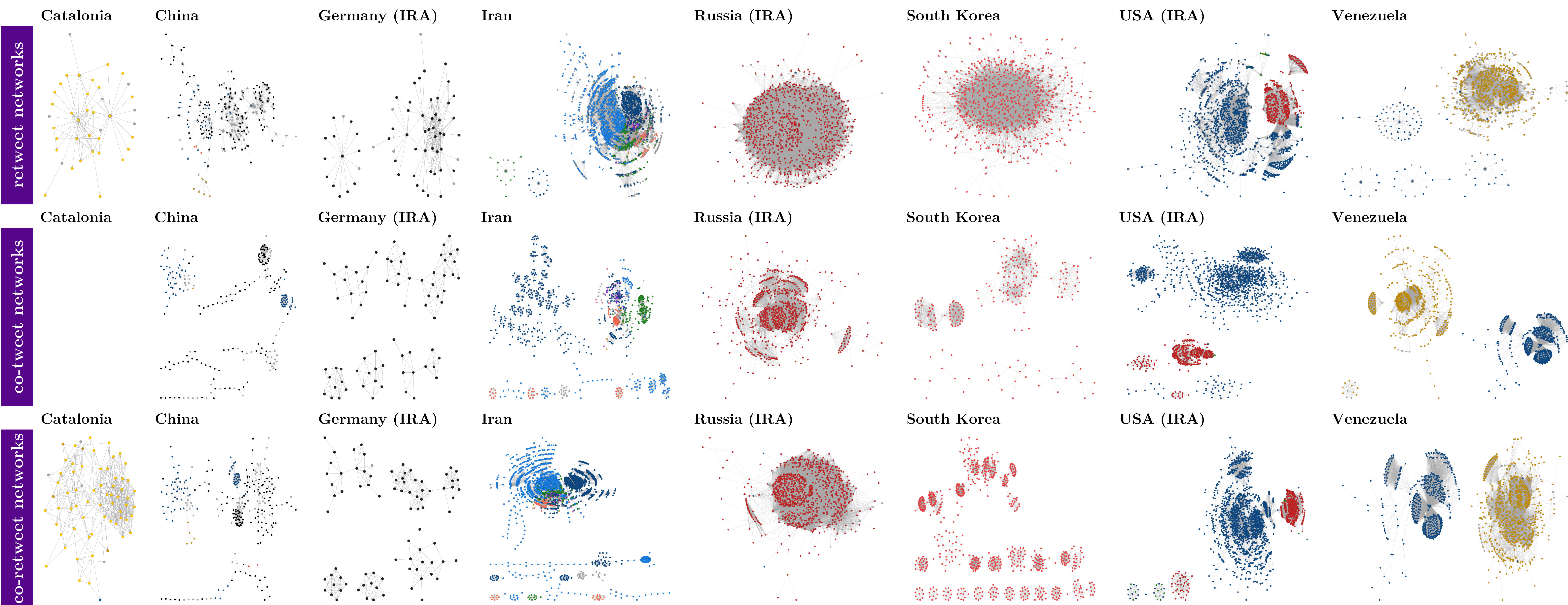Focusing on retweeting (also each other) within a short time period.

Only work when supervised
Tweeting mostly during office hours on workdays.

🚫 **What about bot detection?**
not every bot is an astroturfer
coordination remains hidden

Future work
Compare results to random samples of regular users.

Extend exact co-tweeting with text similarity methods.

Coordination patterns

retweet networks: Catalonia · China · Germany (IRA) · Iran · Russia (IRA) · South Korea · USA (IRA) · Venezuela

co-tweet networks: Catalonia · China · Germany (IRA) · Iran · Russia (IRA) · South Korea · USA (IRA) · Venezuela

co-retweet networks: Catalonia · China · Germany (IRA) · Iran · Russia (IRA) · South Korea · USA (IRA) · Venezuela

References
"Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign" *Political Communication* (forthcoming)