
User Modeling

You are what you post!

Represent users by documents
a user = a document including all she said

You are what you post!

- Modeling User Personality (Computational Social Science)
 - Personality traits in psychology
 - Big Five: extraversion, emotional stability, agreeableness, conscientiousness, and openness to experience
- Modeling User Health Profile (Computational Epidemiology)
 - Privacy of the user, Ethical principles
- Modeling Gender and Ethnicity
 - First names → gender; Last names → ethnicity
- Modeling User Location

You are what you post! *within time.*

Predicting Personal Life Events from Streaming Social Content

Maryam Khodabakhsh
Ferdowsi University of Mashhad
maryamkhodabakhsh@stu.mail.ac.ir

Fattane Zarrinkalam
Ryerson University
fzarrinkalam@ryerson.ca

Hossein Fani
University of New Brunswick
hfani@unb.ca

Ebrahim Bagheri
Ryerson University
bagheri@ryerson.ca

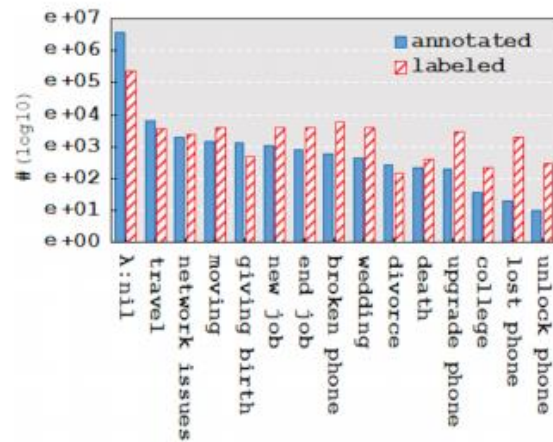


Figure 1: Distribution of personal life events by event class.

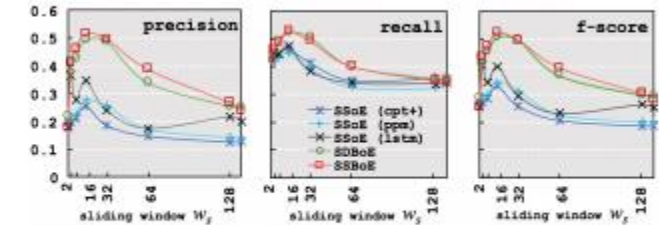
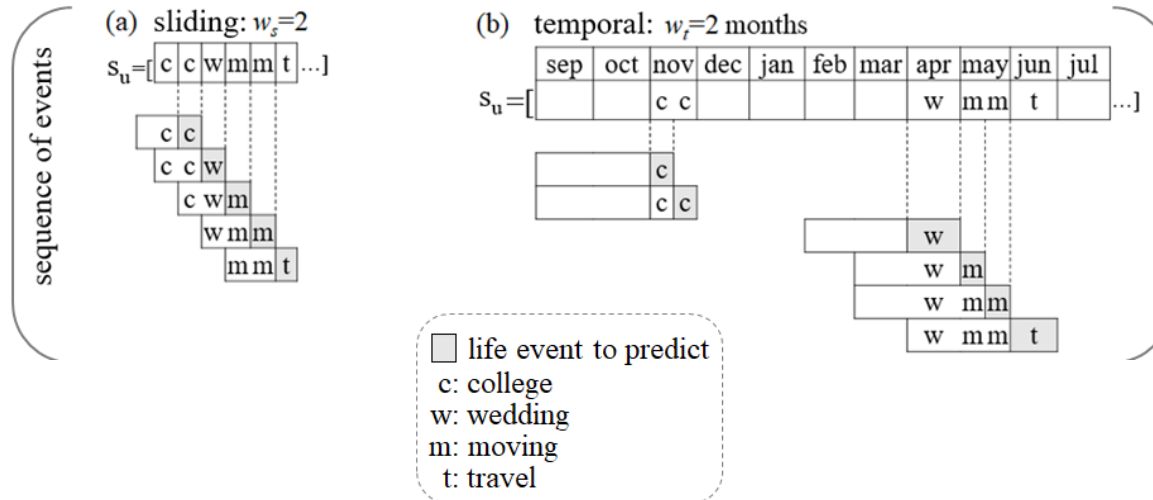


Figure 3: Comparative results of the *sliding* strategy.

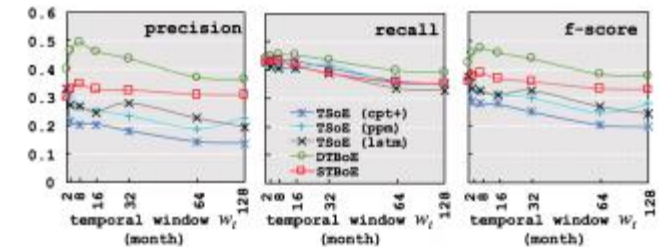


Figure 4: Comparative results of the *temporal* strategy.

You are what you post! *within time.*

User community detection via embedding of social network structure and temporal content★

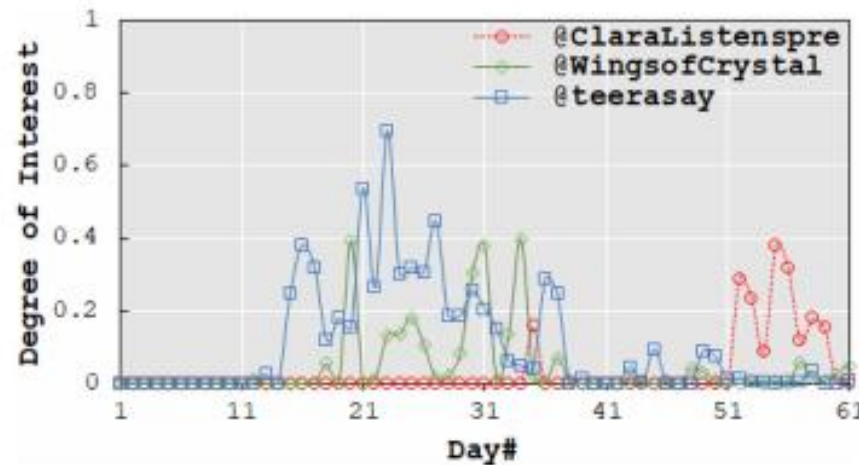


Fig. 1. Different temporal behaviour of three Twitter users with respect to the 'War in Afghanistan' topic.

All users are interested in z_{44} = 'War in Afghanistan'

You are what you post! *within time.*

User community detection via embedding of social network structure and temporal content★

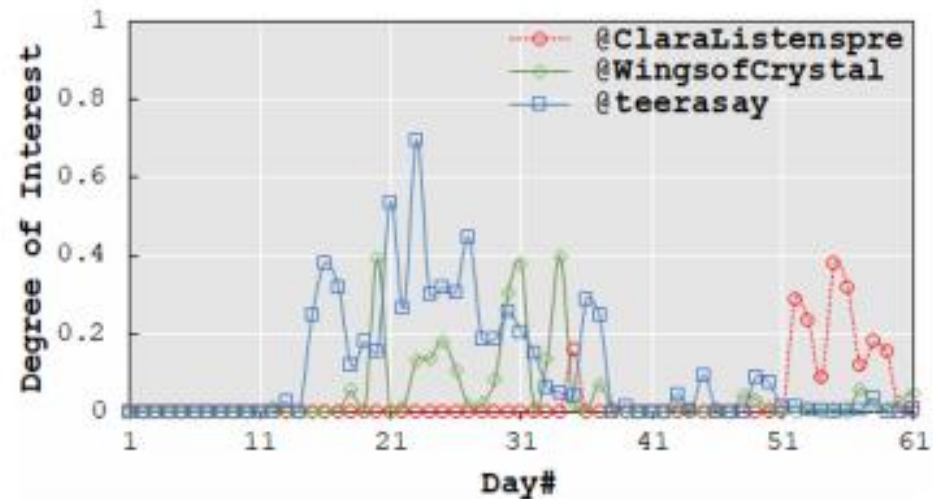
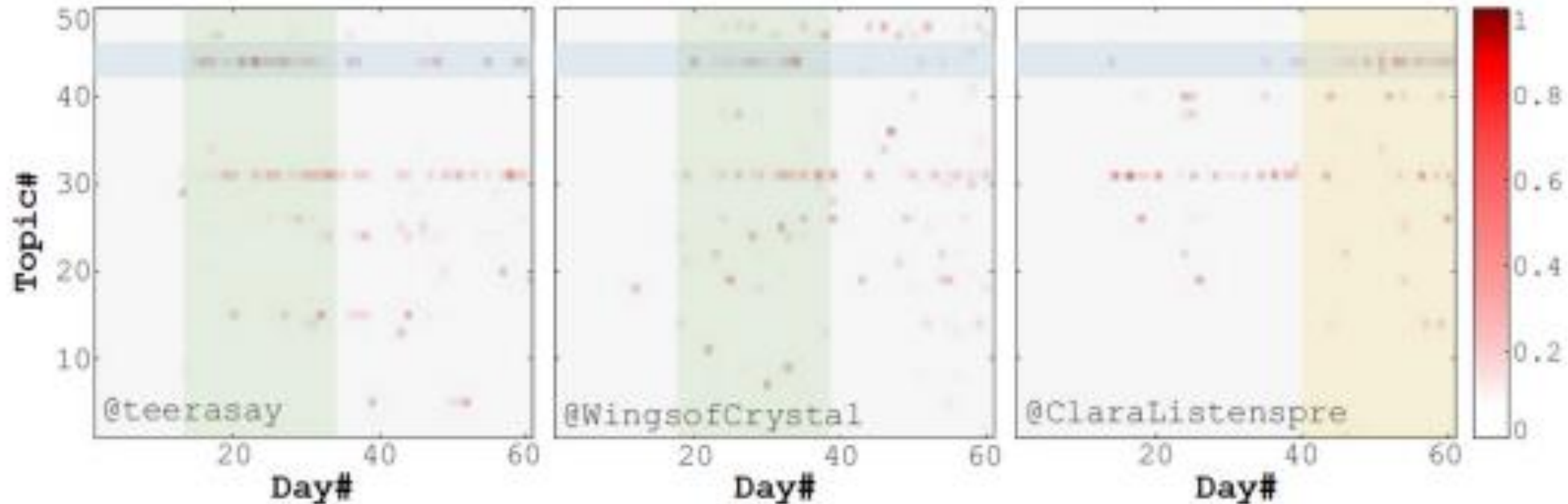


Fig. 1. Different temporal behaviour of three Twitter users with respect to the 'War in Afghanistan' topic.

All users are interested in z_{44} = 'War in Afghanistan'
but not aligned in time!

You are what you post! *within time.*

User community detection via embedding of social network structure and temporal content☆



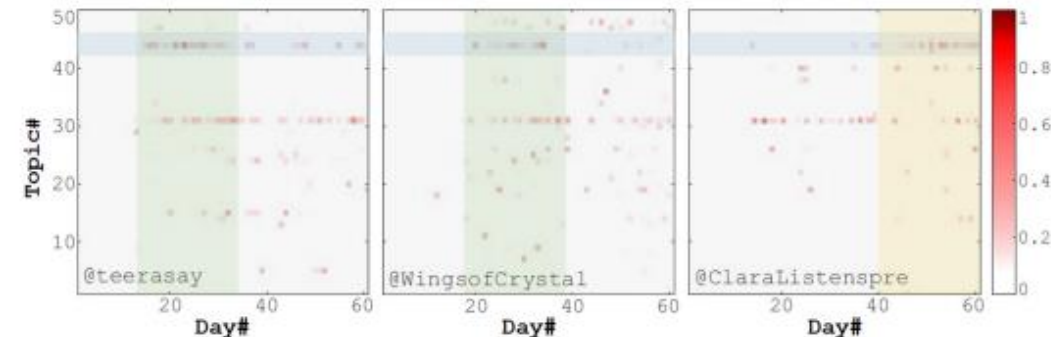
All users are interested in z_{44} = 'War in Afghanistan'
but not aligned in time!

Diachronically Like-minded User Community Detection

- User Clustering
 - Timeseries (Image) Clustering

User \leftrightarrow Documents \rightarrow User Vector \leftrightarrow Document Vector

- How to include time?



Diachronically Like-minded User Community Detection

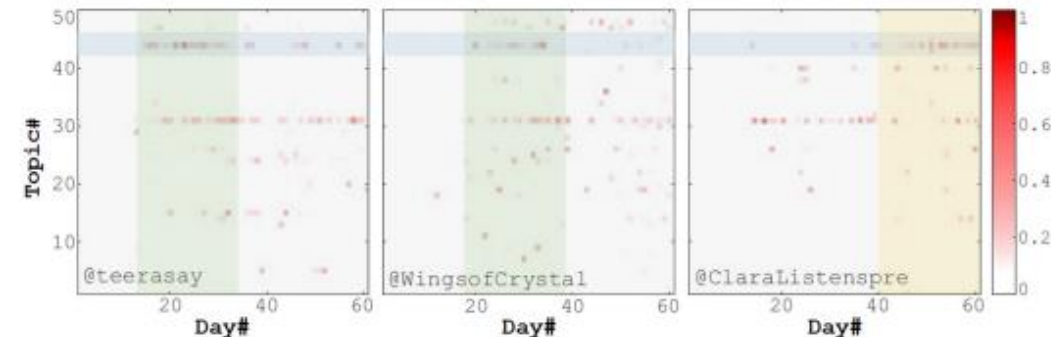
- User Clustering
 - User Vector Representation

User \leftrightarrow Documents \rightarrow User Vector \leftrightarrow Document Vector

- How to include time?

User at time $t \leftrightarrow$ A document that has all she said at time t

User = $[Doc_0, Doc_1, \dots, Doc_T]$



Diachronically Like-minded User Community Detection

(a)

		t_{20}	t_{21}	t_{22}	t_{23}	t_{24}	t_{25}	t_{26}	t_{27}	t_{28}	t_{29}	t_{30}
u_1 @teerasay	z_{40}			0.2					0.1			
	z_{41}											
u_2 @WingsofCry	z_{40}	0.4		0.1	0.6					0.2	0.2	0.3
	z_{41}			0.1						0.2		
u_3 @ClaraListe	z_{40}							0.4	0.2	0.8		
	z_{41}											
	z_{42}											
	z_{43}							0.1	0.3	0.8		
	z_{44}											
	z_{45}											

$$\text{User} = [\text{Doc}_0, \text{Doc}_1, \dots, \text{Doc}_T]$$



LDA



$$\text{User} = [[z^{(0)}_{1:K}], [z^{(1)}_{1:K}], \dots, [z^{(T)}_{1:K}]]$$

(a)

		t_{20}	t_{21}	t_{22}	t_{23}	t_{24}	t_{25}	t_{26}	t_{27}	t_{28}	t_{29}	t_{30}
u_1 @teerasay	z_{40}			0.2					0.1			
	z_{41}											
	z_{42}											
	z_{43}	0.2	0.2	0.3								
	z_{44}	0.2	0.5	0.3	0.7	0.4	0.3	0.4	0.5	0.2	0.2	0.3
	z_{45}							0.2				
u_2 @WingsofCry	z_{40}	0.4		0.1	0.6				0.2	0.2	0.2	0.3
	z_{41}			0.1					0.2			
	z_{42}								0.2			
	z_{43}	0.3		0.1		0.9	0.5		0.2		0.4	
	z_{44}	0.4		0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.8	0.3
	z_{45}				0.5	0.2	0.5	0.2	0.2	0.4	0.2	
u_3 @ClaraListe	z_{40}					0.4	0.2	0.8				
	z_{41}											
	z_{42}											
	z_{43}					0.1	0.3	0.6				
	z_{44}											
	z_{45}											

(b)

		t_{50}	t_{51}	t_{52}	t_{53}	t_{54}	t_{55}	t_{56}	t_{57}	t_{58}	t_{59}	t_{60}
u_3 @ClaraListe	z_{40}		0.1	0.2	0.1	0.2			0.1		0.2	
	z_{41}				0.1							
	z_{42}		0.1			0.1						
	z_{43}		0.2		0.1							
	z_{44}			0.3	0.2	0.1			0.1		0.1	
	z_{45}		0.1	0.1	0.1						0.1	

Diachronically Like-minded User Community Detection

(a)

		t_{20}	t_{21}	t_{22}	t_{23}	t_{24}	t_{25}	t_{26}	t_{27}	t_{28}	t_{29}	t_{30}
u_1 @teerasay	z_{40}			0.2					0.1			
	z_{41}											
	z_{42}				0.1	0.6				0.2	0.2	0.3
	z_{43}				0.1					0.2		
	z_{44}											
	z_{45}											
u_2 @WingsofCry	z_{40}	0.4										
	z_{41}											
	z_{42}											
	z_{43}											
	z_{44}											
	z_{45}											
u_3 @ClaraListe	z_{40}							0.4	0.2	0.8		
	z_{41}											
	z_{42}											
	z_{43}											
	z_{44}											
	z_{45}											

$$\text{User} = [\text{Doc}_0, \text{Doc}_1, \dots, \text{Doc}_T]$$



LDA



$$\text{User} = [[z^{(0)}_{1:k}], [z^{(1)}_{1:k}], \dots, [z^{(T)}_{1:k}]]$$

(a)

		t_{20}	t_{21}	t_{22}	t_{23}	t_{24}	t_{25}	t_{26}	t_{27}	t_{28}	t_{29}	t_{30}
u_1 @teerasay	z_{40}			0.2					0.1			
	z_{41}											
	z_{42}											
	z_{43}	0.2	0.2	0.3								
	z_{44}	0.2	0.5	0.3	0.7	0.4	0.3	0.4	0.5	0.2	0.2	0.3
	z_{45}							0.2				
u_2 @WingsofCry	z_{40}	0.4		0.1	0.6				0.2	0.2	0.2	0.3
	z_{41}			0.1					0.2			
	z_{42}								0.2			
	z_{43}	0.3		0.1		0.9	0.5		0.2		0.4	
	z_{44}	0.4		0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.8	0.3
	z_{45}				0.5	0.2	0.5	0.2	0.2	0.4	0.2	
u_3 @ClaraListe	z_{40}					0.4	0.2	0.8				
	z_{41}											
	z_{42}											
	z_{43}											
	z_{44}											
	z_{45}											

(b)

		t_{50}	t_{51}	t_{52}	t_{53}	t_{54}	t_{55}	t_{56}	t_{57}	t_{58}	t_{59}	t_{60}
u_3 @ClaraListe	z_{40}		0.1	0.2	0.1	0.2			0.1		0.2	
	z_{41}				0.1							
	z_{42}		0.1			0.1						
	z_{43}		0.2		0.1							
	z_{44}			0.3	0.2	0.1			0.1		0.1	
	z_{45}		0.1	0.1	0.1						0.1	

Two users are similar if they share more cells!

each cell = $1 \times 1 \times 1$ cube = $\{u_i\} \times \{z_j\} \times \{t_k\}$

Shared cell = $n \times m \times k$ cube

e.g., $\{u_1 u_2\} \times \{z_{44}\} \times \{t_{22} t_{23} \dots t_{30}\}$

Diachronically Like-minded User Community Detection

(a)

		t ₂₀	t ₂₁	t ₂₂	t ₂₃	t ₂₄	t ₂₅	t ₂₆	t ₂₇	t ₂₈	t ₂₉	t ₃₀
@teerasay u ₁	z ₄₀			0.2					0.1			
	z ₄₁											
	z ₄₂											
	z ₄₃											
	z ₄₄											
	z ₄₅											
@WingsofCry u ₂	z ₄₀	0.4		0.1	0.6					0.2	0.2	0.3
	z ₄₁			0.1						0.2		
	z ₄₂											
	z ₄₃											
	z ₄₄											
	z ₄₅											
@ClaraListe u ₃	z ₄₀								0.4	0.2	0.8	
	z ₄₁											
	z ₄₂											
	z ₄₃								0.1	0.3	0.8	
	z ₄₄											
	z ₄₅											

Region of Like-mindedness (RoL) iff

$$y_t^u[z] \approx y_t^v[z]$$

Two users are similar if they share more cells!

each cell = 1×1×1 cube = {u_i} × {z_j} × {t_k}

Shared cell = n×m×k cube

e.g., {u₁u₂} × {z₄₄} × {t₂₂ t₂₃ ... t₃₀}

(a)

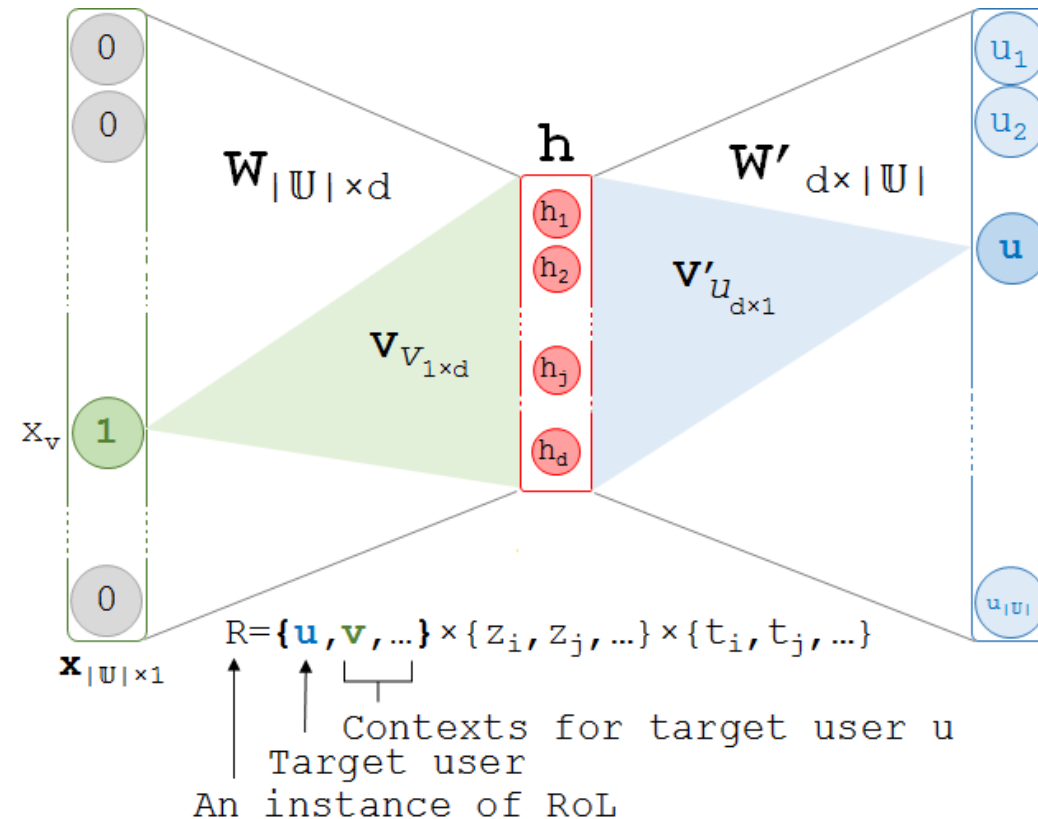
		t ₂₀	t ₂₁	t ₂₂	t ₂₃	t ₂₄	t ₂₅	t ₂₆	t ₂₇	t ₂₈	t ₂₉	t ₃₀
@teerasay u ₁	z ₄₀			0.2					0.1			
	z ₄₁											
	z ₄₂											
	z ₄₃	0.2	0.2	0.3								
	z ₄₄	0.2	0.5	0.3	0.7	0.4	0.3	0.4	0.5	0.2	0.2	0.3
	z ₄₅							0.2				
@WingsofCry u ₂	z ₄₀	0.4		0.1	0.6				0.2	0.2	0.2	0.3
	z ₄₁			0.1					0.2			
	z ₄₂								0.2			
	z ₄₃	0.3		0.1		0.9	0.5		0.2		0.4	
	z ₄₄	0.4		0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.8	0.3
	z ₄₅				0.5	0.2	0.5	0.2	0.2	0.4	0.2	
@ClaraListe u ₃	z ₄₀					0.4	0.2	0.8				
	z ₄₁											
	z ₄₂											
	z ₄₃											
	z ₄₄											
	z ₄₅											

(b)

		t ₅₀	t ₅₁	t ₅₂	t ₅₃	t ₅₄	t ₅₅	t ₅₆	t ₅₇	t ₅₈	t ₅₉	t ₆₀
@ClaraListe u ₃	z ₄₀		0.1	0.2	0.1	0.2			0.1		0.2	
	z ₄₁				0.1							
	z ₄₂		0.1			0.1						
	z ₄₃		0.2		0.1							
	z ₄₄			0.3	0.2	0.1			0.1		0.1	
	z ₄₅		0.1	0.1	0.1						0.1	

Diachronically Like-minded User Community Detection

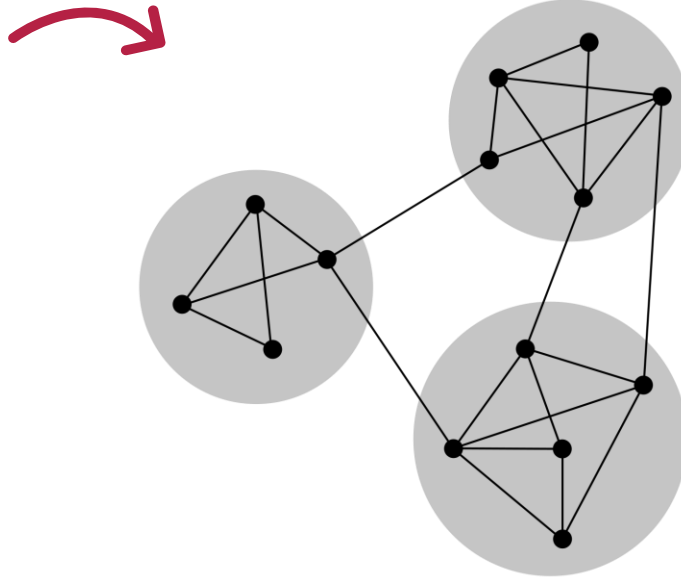
- User Clustering
 - ~~Timeseries (Image) Clustering~~
 - User2Vec: User Vector Representation: Two Similar Users \rightarrow Similar Vectors



Diachronically Like-minded User Community Detection

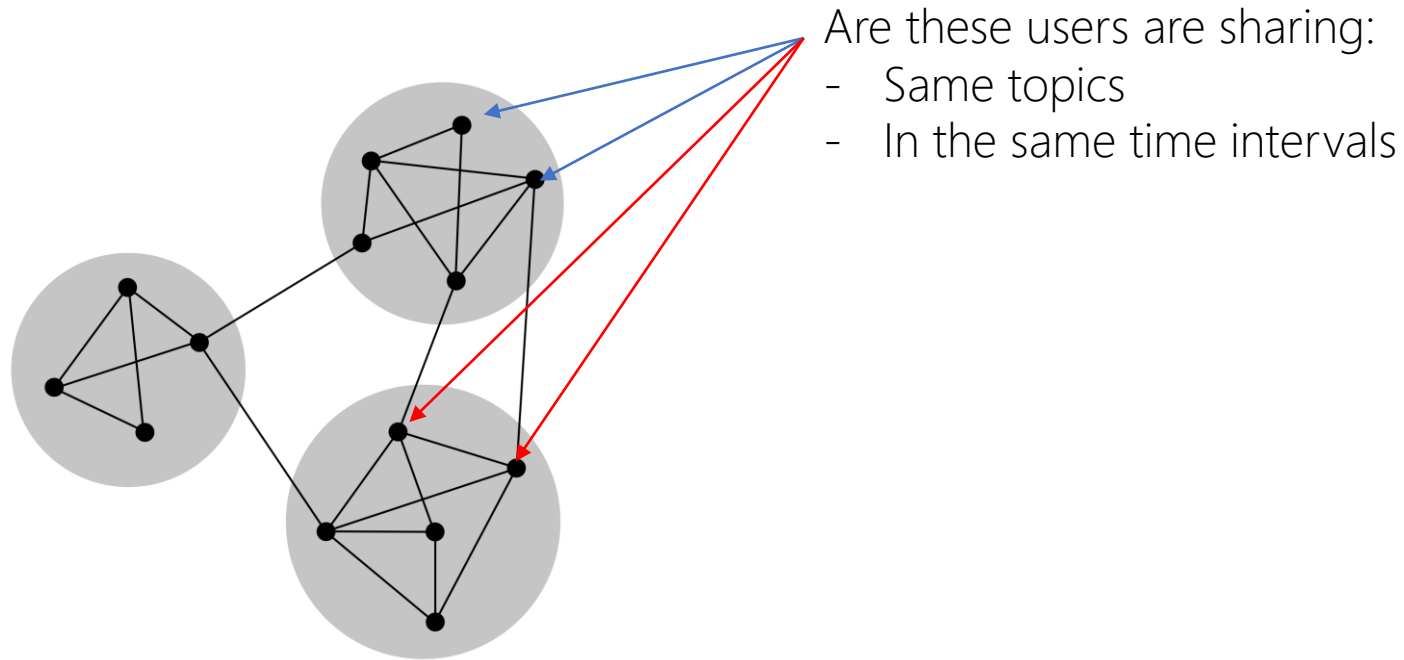
Louvain Method (Blondel et al. JSTAT 2008)

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & & & A_{2n} \\ \vdots & & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}$$



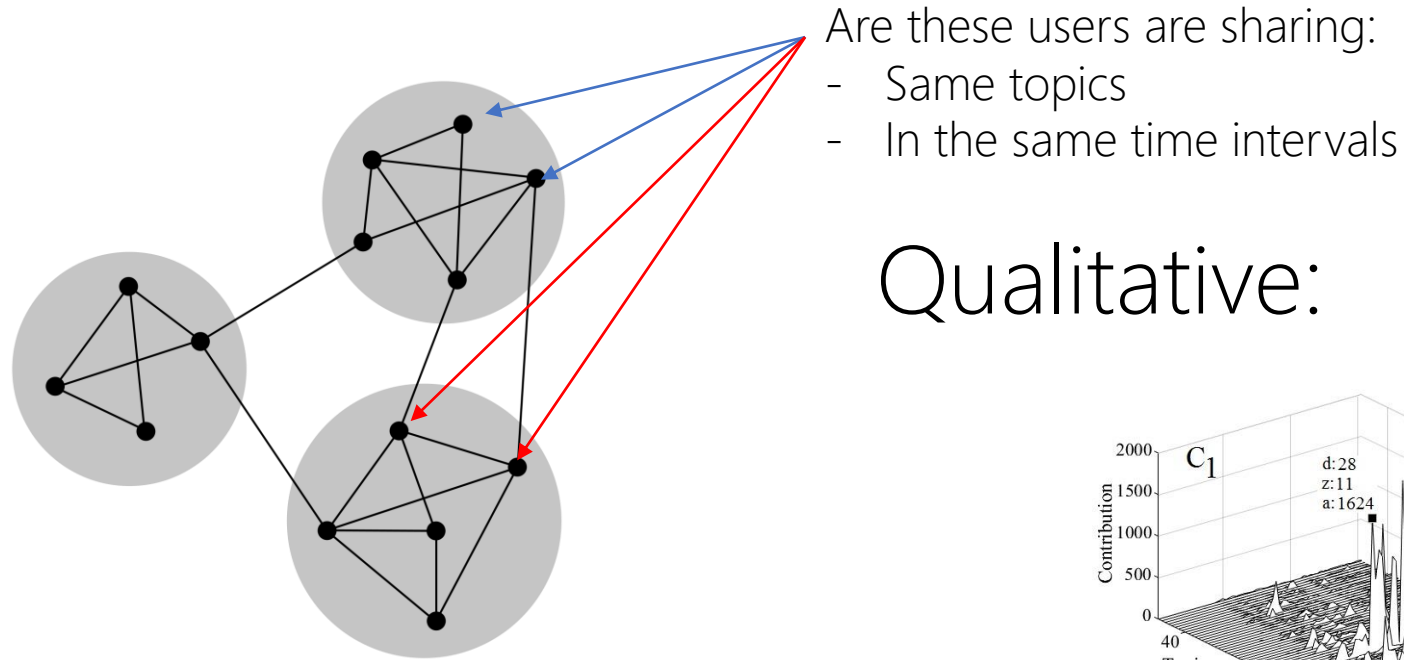
Diachronically Like-minded User Community Detection

Evaluation: how accurate are the communities?

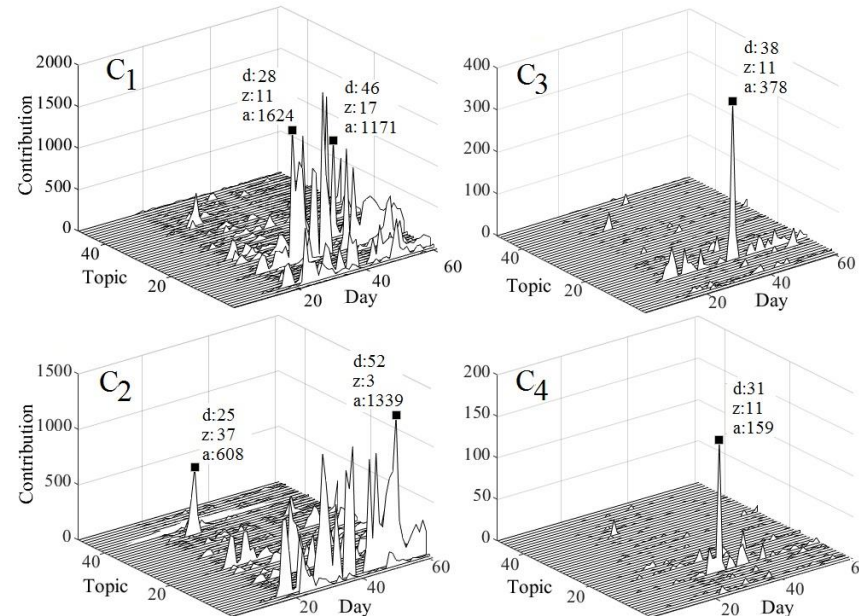


Diachronically Like-minded User Community Detection

Evaluation: how accurate are the communities?

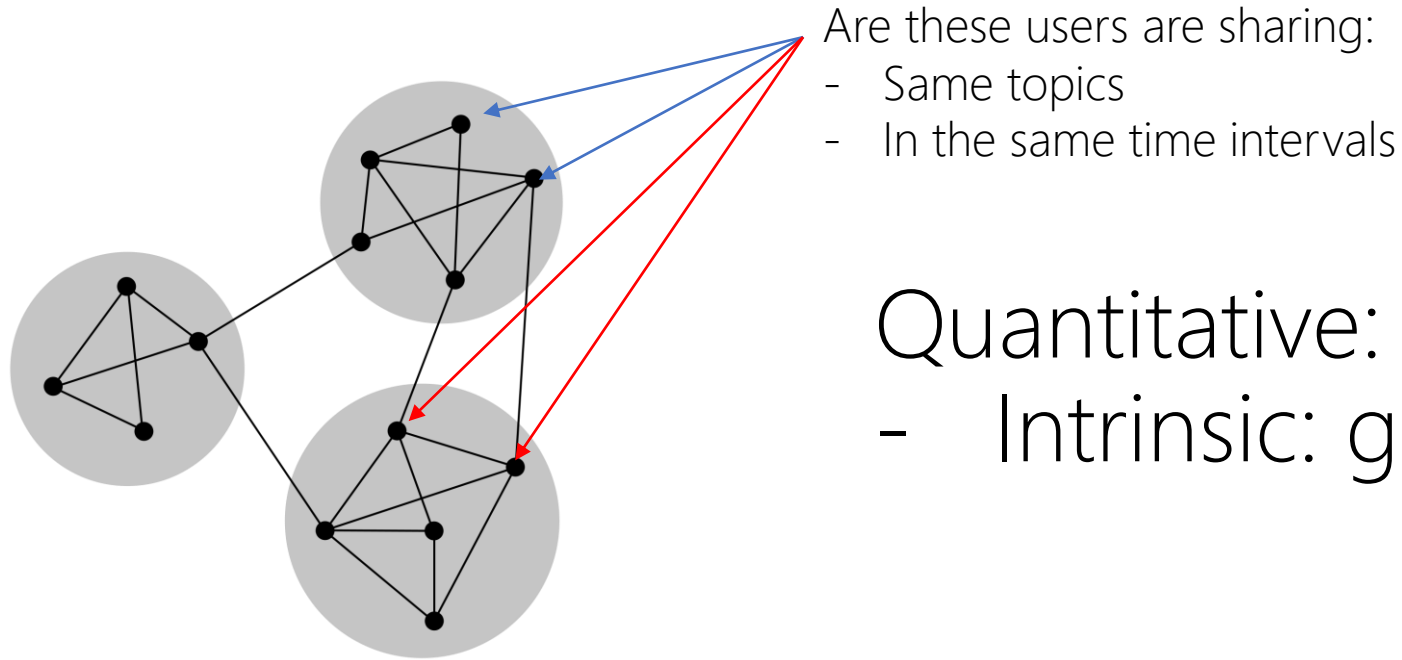


Qualitative:



Diachronically Like-minded User Community Detection

Evaluation: how accurate are the communities?

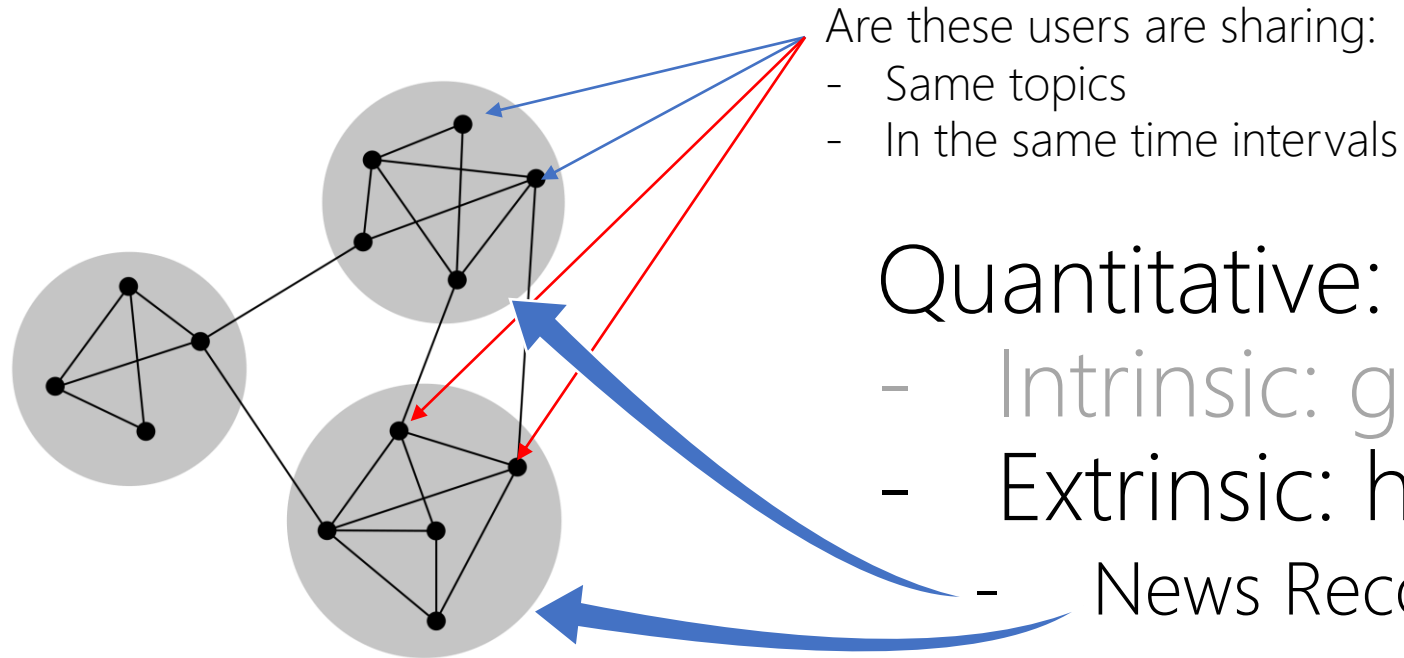


Quantitative:

- Intrinsic: golden communities

Diachronically Like-minded User Community Detection

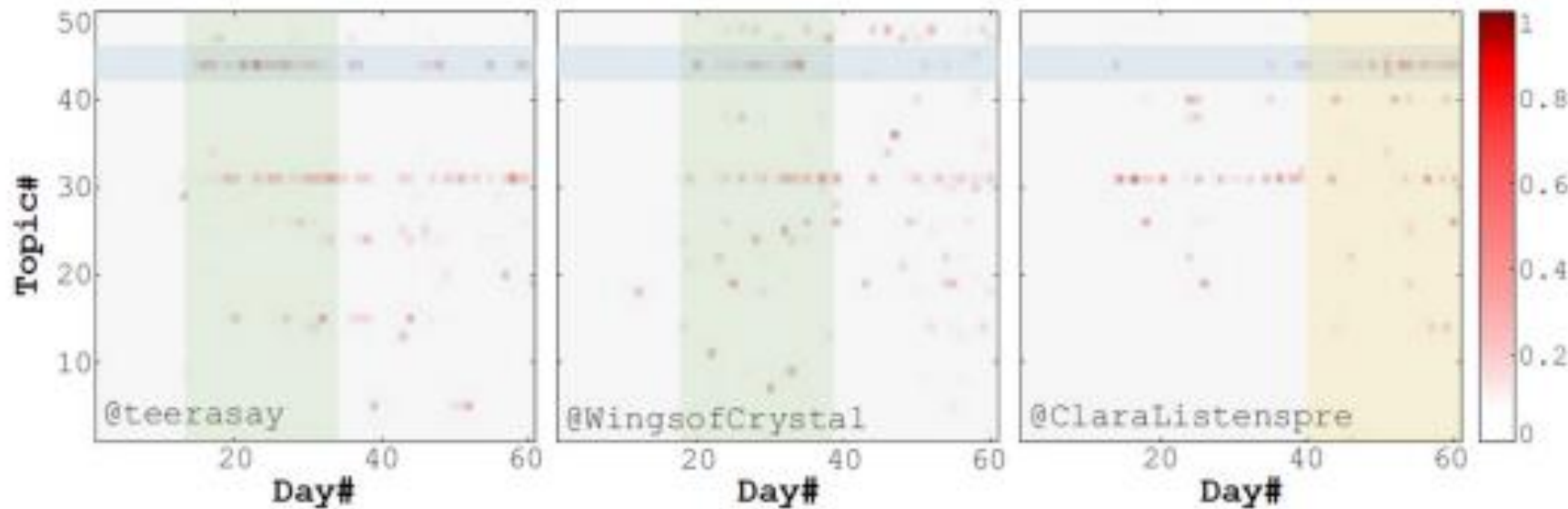
Evaluation: how accurate are the communities?



Recommend news articles to users to read at today, tomorrow, next week.
Instead of per user recommendation, we recommend to the communities!

Diachronically Like-minded User Community Detection

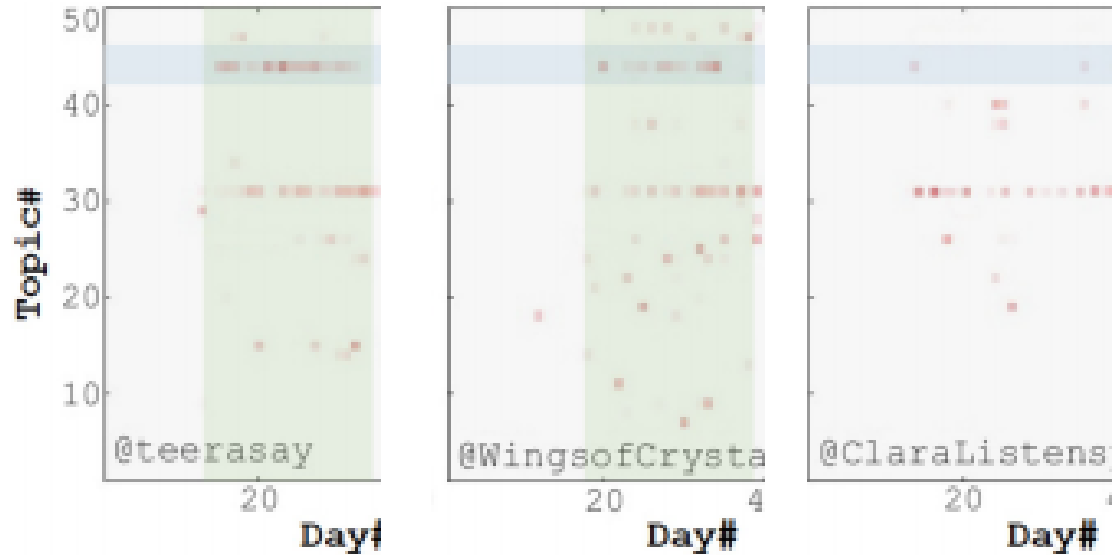
Evaluation: how accurate are the communities?



Recommend news articles to users to read at today, tomorrow, next week.
Instead of per user recommendation, we recommend to the communities!

Diachronically Like-minded User Community Detection

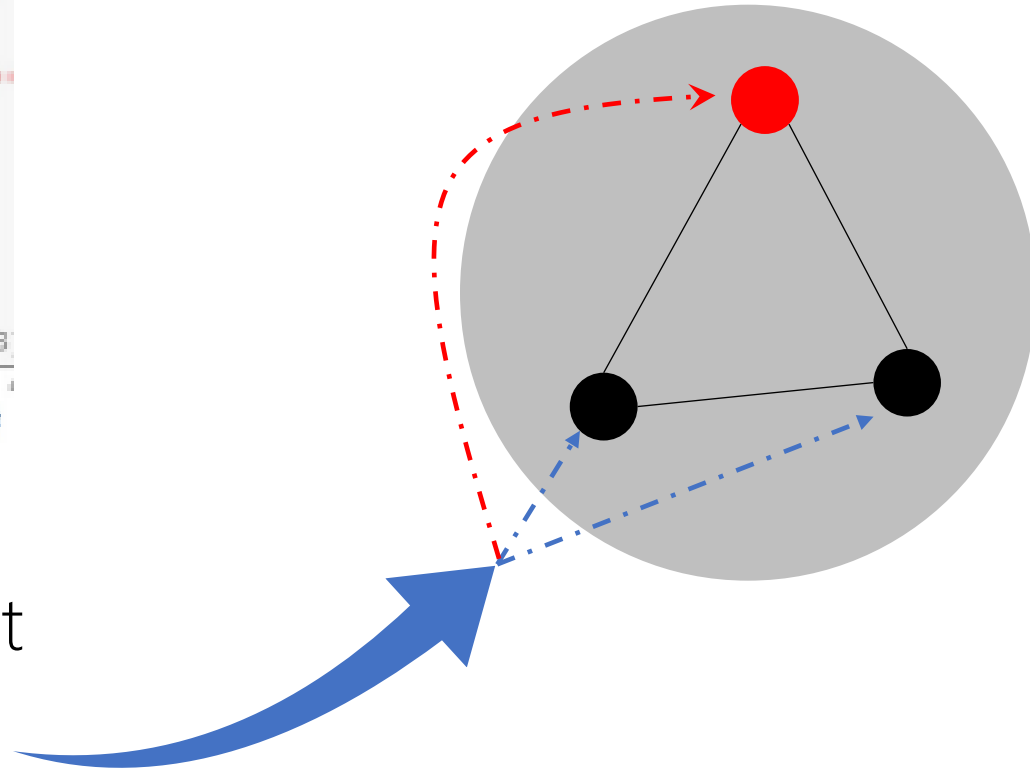
Evaluation: how accurate are the communities?



all three users in same community

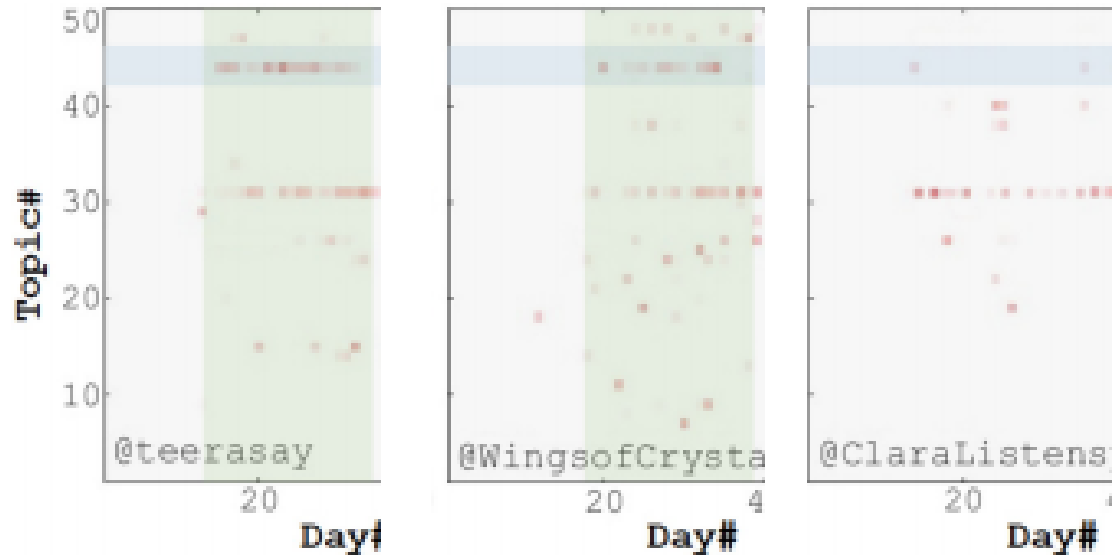
Recommend news articles about

- Z_{44} : "War in Afghanistan"
- at day = 40

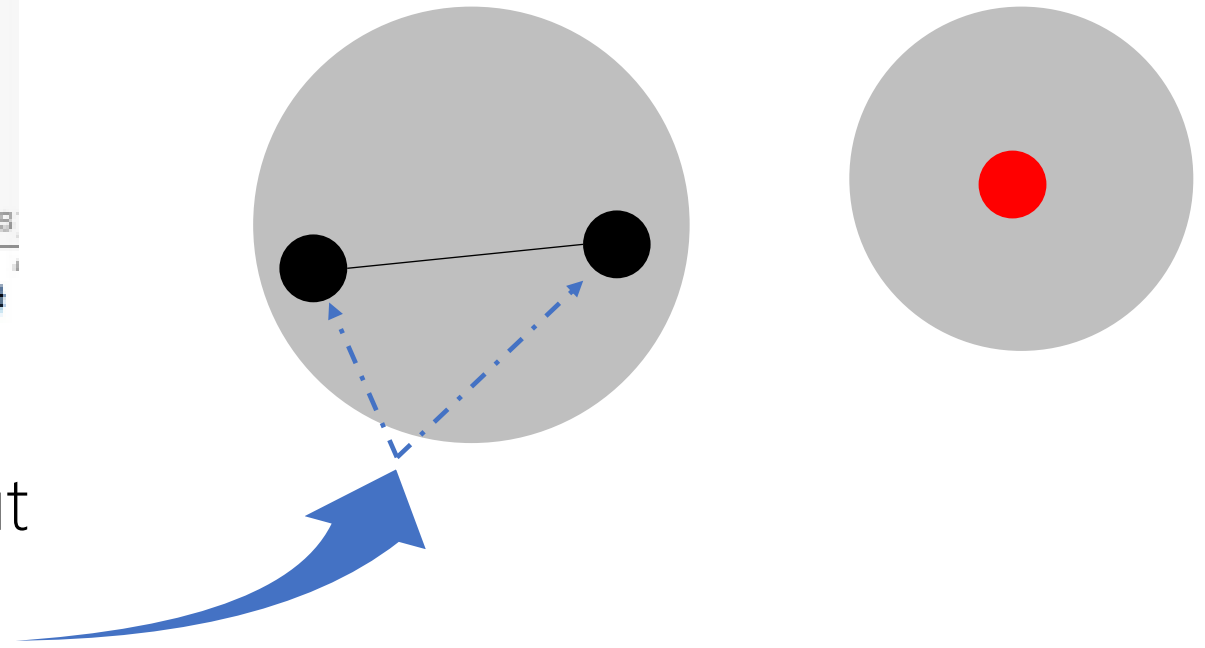


Diachronically Like-minded User Community Detection

Evaluation: how accurate are the communities?



the first two users in same community
the last user in another community



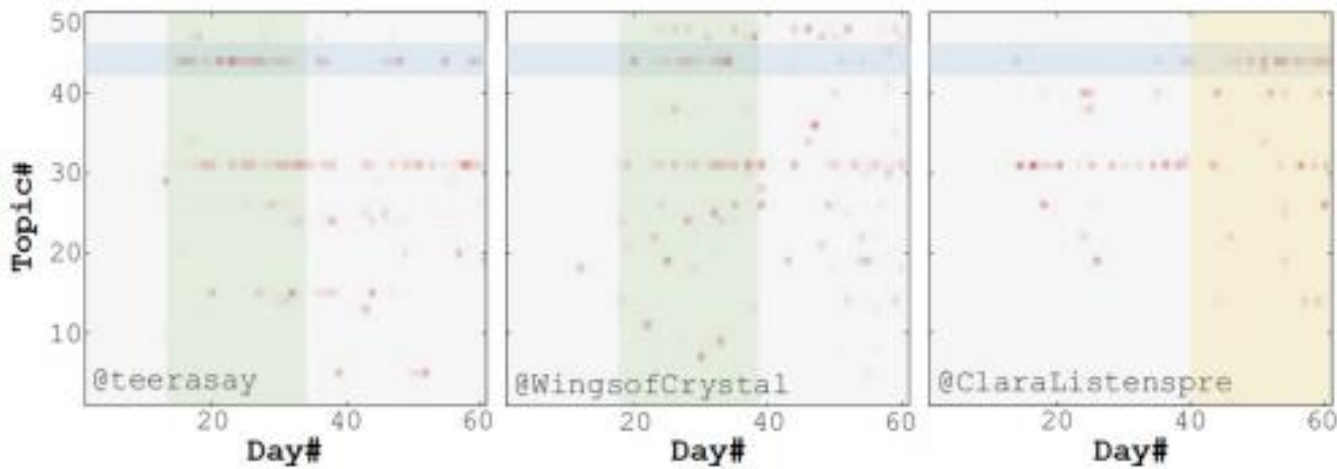
Recommend news articles about

- Z_{44} : "War in Afghanistan"
- at day = 40

Diachronically Like-minded User Community Detection

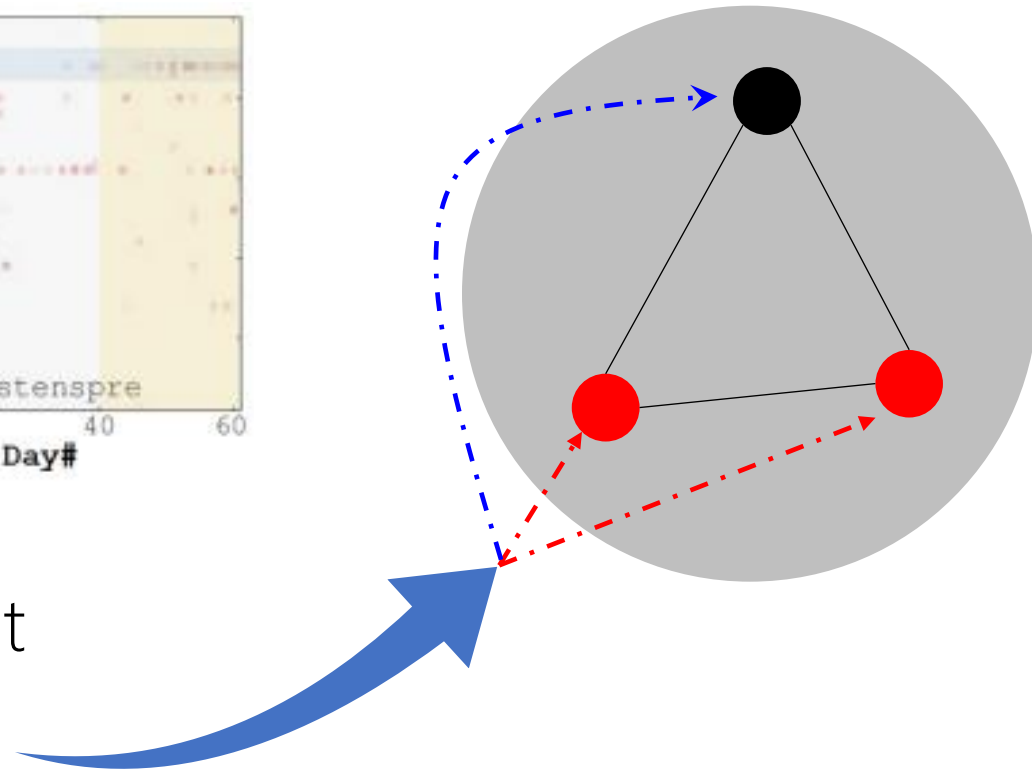
Evaluation: how accurate are the communities?

all three users in same community



Recommend news articles about

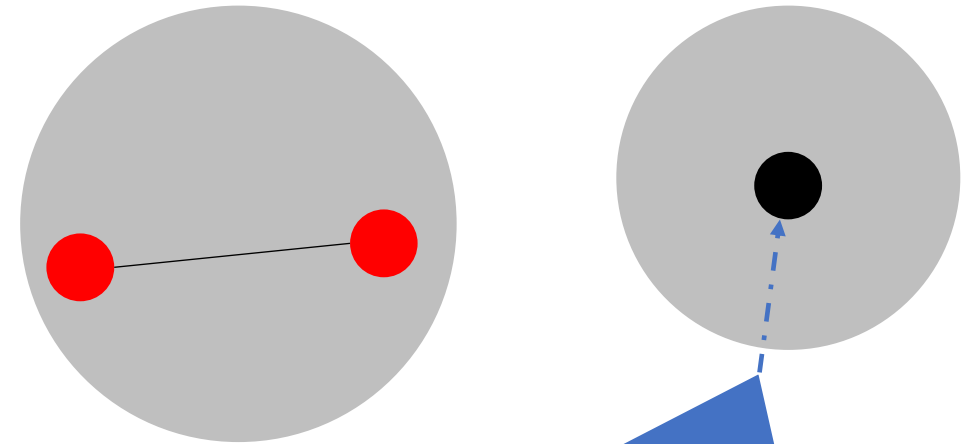
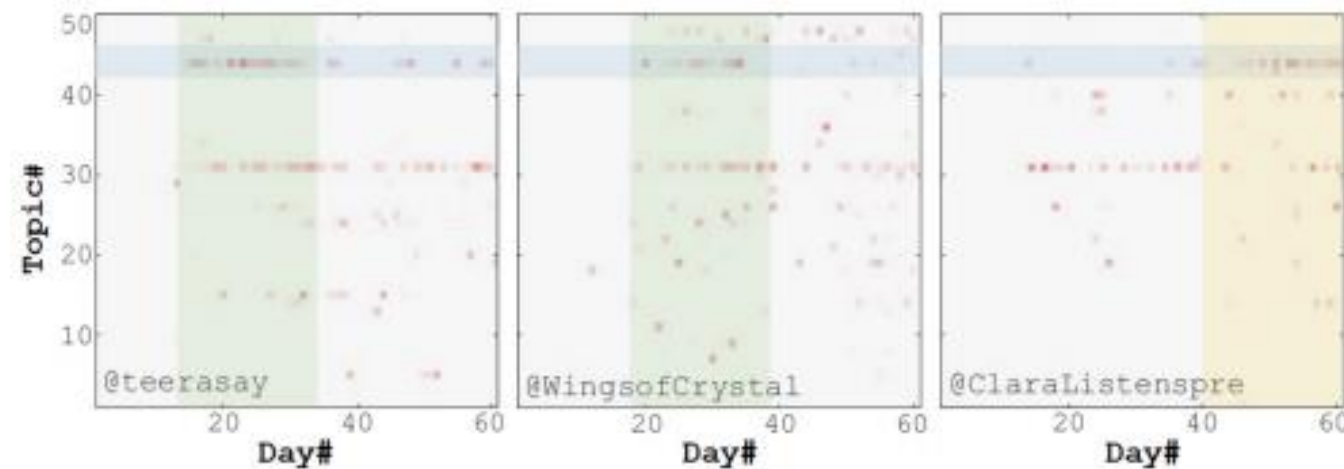
- Z_{44} : "War in Afghanistan"
- at day = 60



Diachronically Like-minded User Community Detection

Evaluation: how accurate are the communities?

the first two users in same community
the last user in another community



Recommend news articles about

- Z_{44} : "War in Afghanistan"
- at day = 60

Diachronically Like-minded User Community Detection

Evaluation Strategy

Assumption:

users are interested in the topics of the news article about which they have posted



Christopher Manning @chrmanning · Mar 24

"The US should establish a standalone, open-source intelligence agency, because existing agencies will never give open-source the attention it needs to succeed"—@AmyZegart. You'd think the US has enough intelligence agencies but this could just be right 🧐



HAI_USIntelligence_FINAL.pdf
drive.google.com



1



1



14



HAI
Stanford University
Human-Centered
Artificial Intelligence

MARCH 2021

The Moment of Reckoning: AI and the Future of U.S. Intelligence

Amy Zegart

THE U.S. INTELLIGENCE COMMUNITY FACES A MOMENT OF RECKONING and AI lies at the heart of it. Since 9/11, America's intelligence agencies have become hardwired to fight terrorism. Today's threat landscape, however, is changing dramatically, with a resurgence of great power competition and the rise of cyber threats enabling states and non-state actors to spy, steal, disrupt, destroy, and deceive across vast distances — all without firing a shot.

At the same time, new technologies are eroding the "decision advantage" that the U.S. has traditionally enjoyed. Until recently, intelligence was a superpower contest. Not anymore. The rapid global expansion of cell phones, internet connectivity, and commercial satellites has created a world awash in open-source data that can be collected, analyzed, and used by anyone. Today, geopolitical success, whether it's preventing war or advancing American economic interests, requires harnessing all this information to understand trends, events, threats, and opportunities faster and better than adversaries who are unencumbered by America's constitutional and ethical obligations to protect civil liberties and privacy. AI promises to be at the forefront of these new capabilities — with the potential to transform the collection, analysis, and dissemination of intelligence vital to American national security.

KEY TAKEAWAYS

- The Intelligence Community (IC) faces a moment of reckoning. If the IC cannot adopt AI and other emerging technologies successfully, it risks failure.
- AI is critical to maintaining America's decision advantage — helping intelligence agencies harness the explosion of open-source data with increased precision, speed, and analytic power to detect hidden and emerging patterns.
- AI's promise lies in augmenting human intelligence collectors and analysts, not replacing them.
- The most important near-term steps are developing a comprehensive intelligence and technology strategy.

Diachronically Like-minded User Community Detection

Evaluation Strategy

Golden Dataset Curation:

news articles to which a user has explicitly linked in her tweets

mentions = {(user, news article, timestamp)}

Diachronically Like-minded User Community Detection

Evaluation Strategy

News Recommendation:

$$\text{mentions} = \{(\text{user}, ?, \text{timestamp})\}$$

We recommend news article n

- About topic z
- At timestamp t
- To a community that shows overall burst at time t about z

Diachronically Like-minded User Community Detection

Evaluation Strategy

News Recommendation:

mentions = {(user, ?, timestamp)}

We hope the community

- Read news article n
- Tweet news article n

Diachronically Like-minded User Community Detection

Evaluation Strategy

News Recommendation:

mentions = {(user, ?, timestamp)}

We evaluate

- Our hope: (user, *n*, timestamp)
- The reality: (user, *news article*, timestamp)

Diachronically Like-minded User Community Detection

Evaluation Strategy

Dataset:

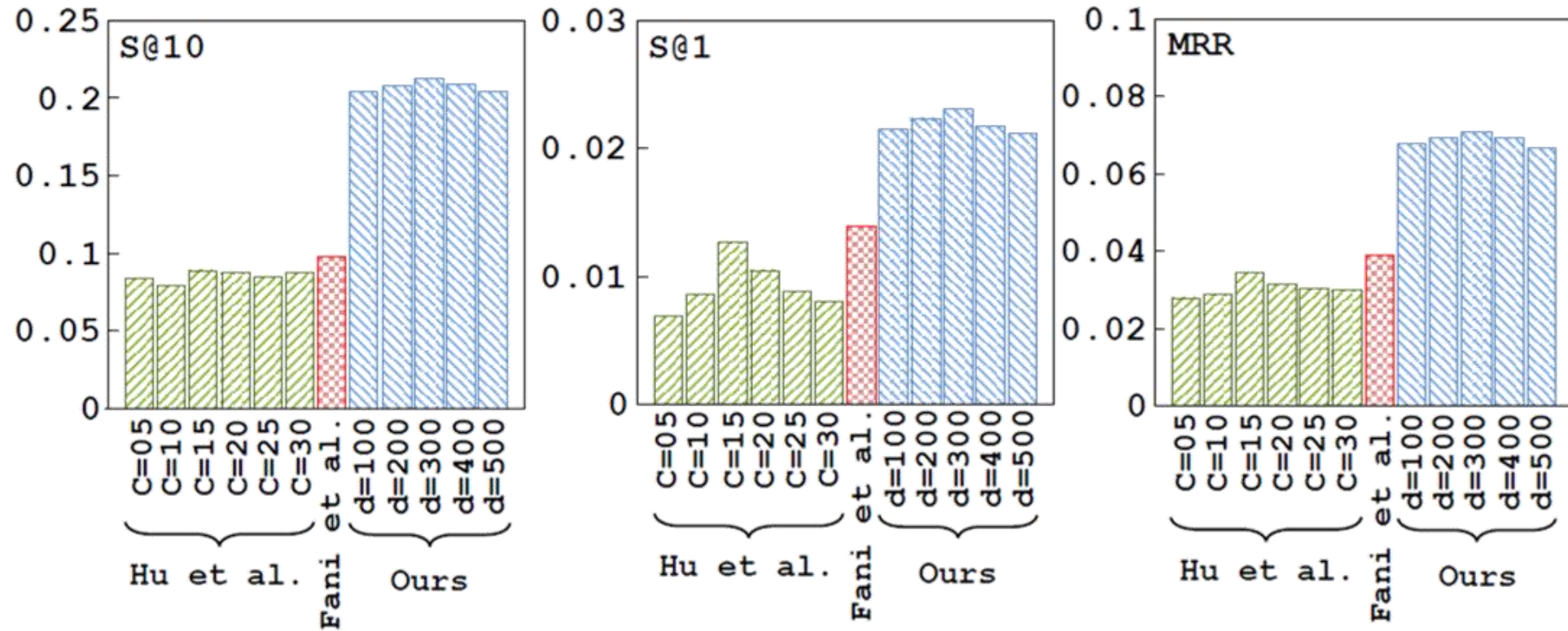
- Abel et al.: Twitter,
 - 3M tweets
 - Posted by 135K users
 - Between Nov. 1 and Dec. 31, 2010.

Golden Entries:

- 25,756 triples extracted from 3,468 distinct news articles posted by 1,922 users

Diachronically Like-minded User Community Detection

Evaluation Strategy



Diachronically Like-minded User Community Detection

Evaluation Strategy (Second)

User Prediction:

Seen a stray tweet, who is the author?

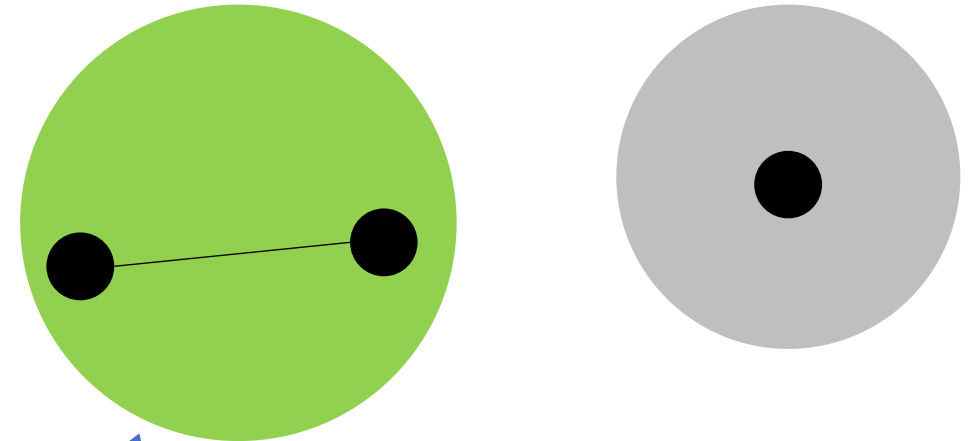
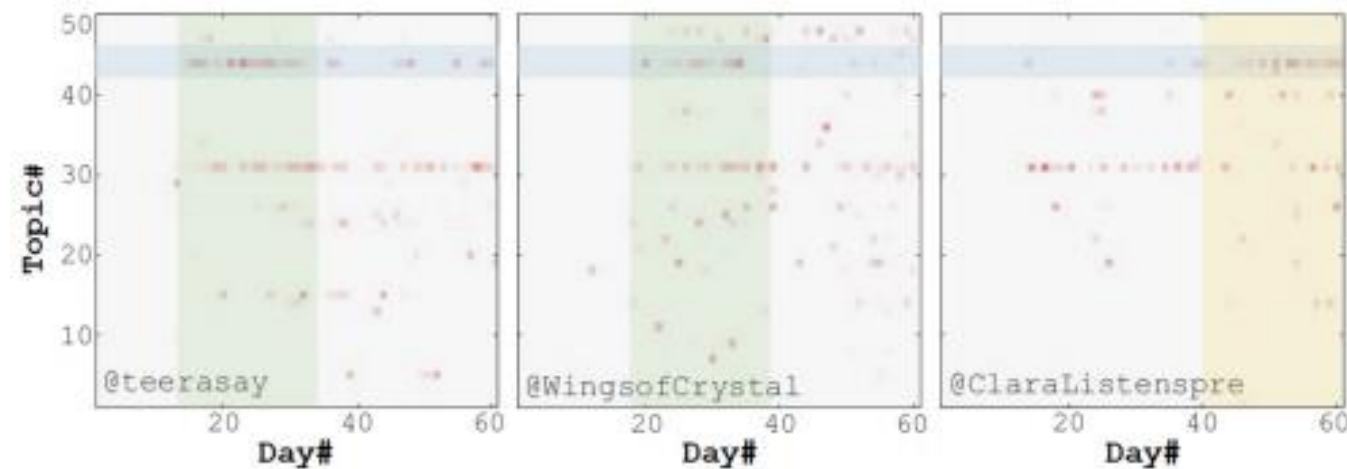
Probably from community \mathcal{C} because this community talked a lot about the topics of the mentioned news article n at timestamp t

mentions = $\{(? , \text{news article}, \text{timestamp})\}$

Diachronically Like-minded User Community Detection

Evaluation: how accurate are the communities?

the first two users in same community
the last user in another community



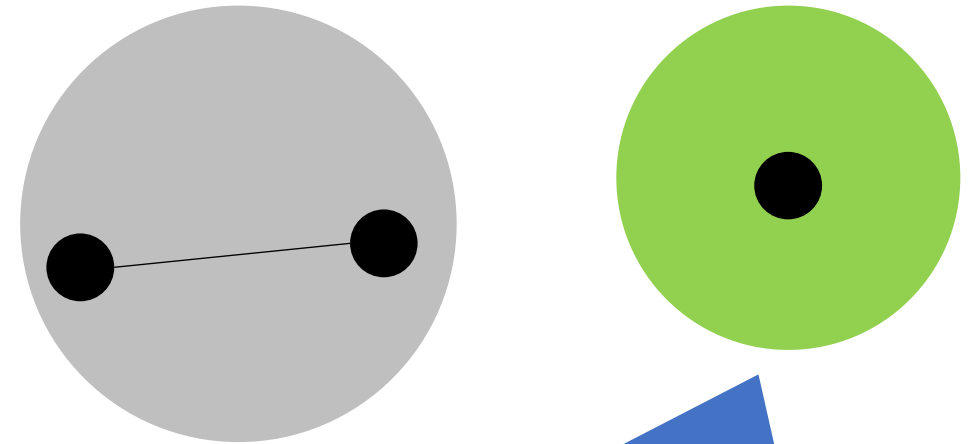
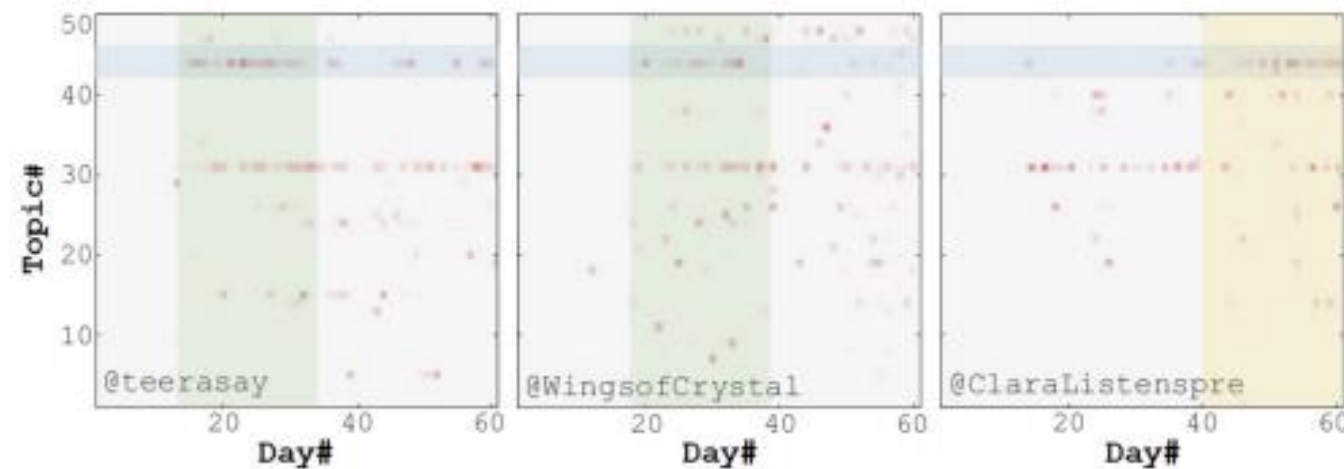
Stray tweet about

- Z_{44} : "War in Afghanistan"
- at day = 40

Diachronically Like-minded User Community Detection

Evaluation: how accurate are the communities?

the first two users in same community
the last user in another community



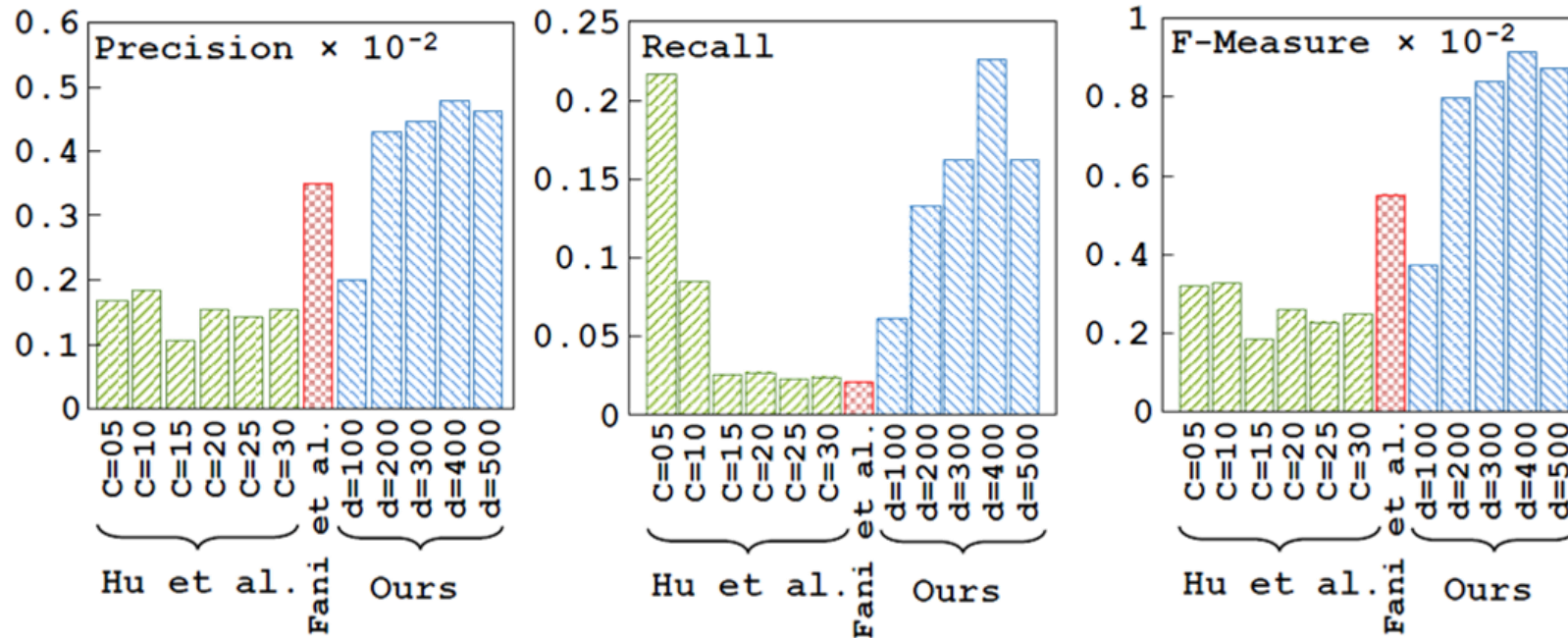
Stray tweet about

- Z_{44} : "War in Afghanistan"
- at day = 60

Diachronically Like-minded User Community Detection

Evaluation Strategy (Second)

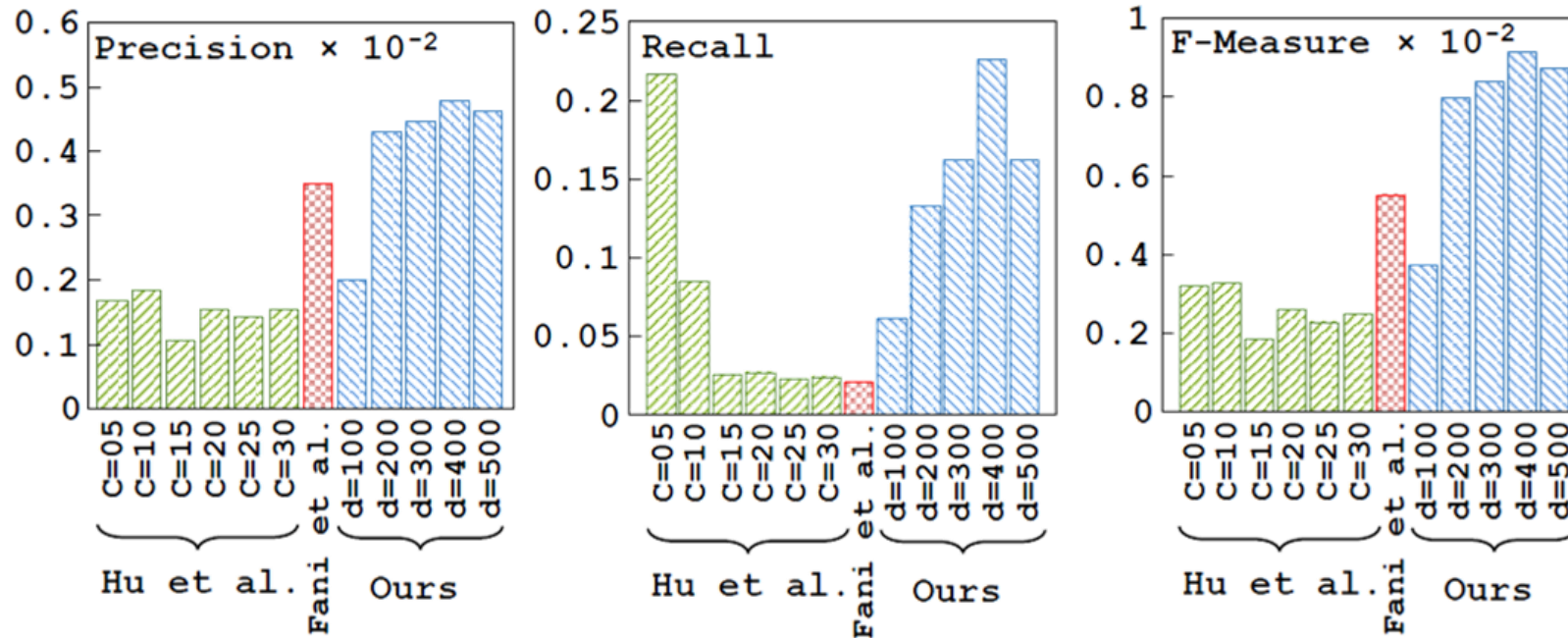
User Prediction:



Diachronically Like-minded User Community Detection

Evaluation Strategy (Second)

User Prediction:



User Community Prediction

User Community Detection in Future!

User Community Prediction

