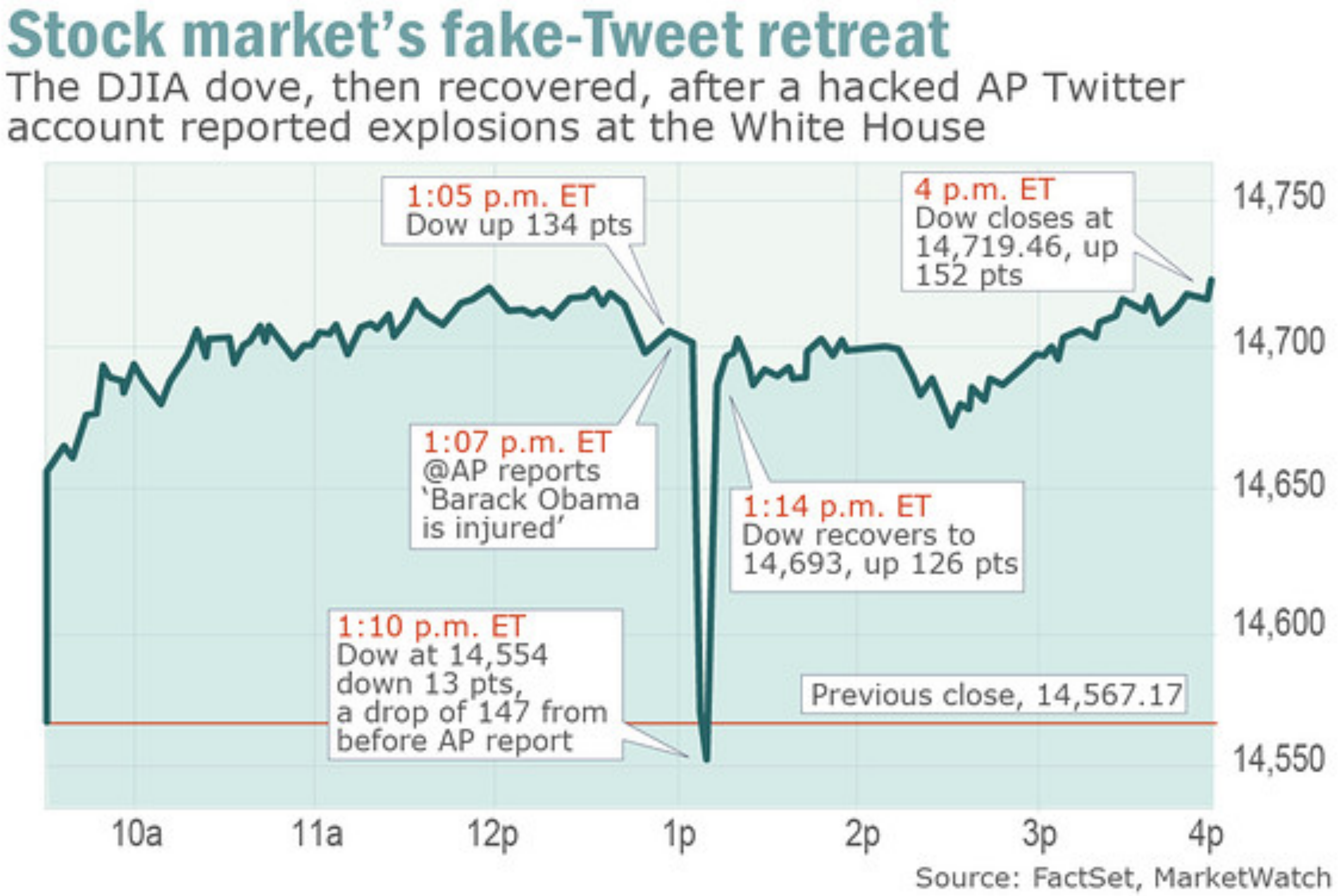
The spread of true and false news online

Soroush Vosoughi, Deb Roy, Sinan Aral

Enis Ceyhun Alp - 23042021


The Associated Press
@AP
Following
Breaking: Two Explosions in the White House and Barack Obama is injured
Reply
Retweet
Favorite
Buffer
More
3,242
RETWEETS
153
FAVORITES
12:07 PM - 23 Apr 13


AP
The Associated Press
@AP
Following
Breaking: Two Explosions in the White House and Barack Obama is injured
Reply
Retweet
Favorite
Buffer
More
3,242
RETWEETS
153
FAVORITES
12:07 PM - 23 Apr 13


Stock market's fake-Tweet retreat
The DJIA dove, then recovered, after a hacked AP Twitter account reported explosions at the White House
1:05 p.m. ET
Dow up 134 pts
4 p.m. ET
Dow closes at 14,719.46, up 152 pts
1:07 p.m. ET
@AP reports 'Barack Obama is injured'
1:14 p.m. ET
Dow recovers to 14,693, up 126 pts
1:10 p.m. ET
Dow at 14,554 down 13 pts, a drop of 147 from before AP report
Previous close, 14,567.17
14,750
14,700
14,650
14,600
14,550
10a
11a
12p
1p
2p
3p
4p
Source: FactSet, MarketWatch

OXFORD INTERNET INSTITUTE | UNIVERSITY OF OXFORD

Research | Study | People | Events | About

HOME > NEWS > PRESS RELEASES >

New report shows that Swedish election second only to US in proportion of 'junk news' shared

Published:
6 September 2018

Research from the Oxford Internet
'junk news' shared on social media
European country studied – and se

1 in 3 news articles shared about Sweden election are fake, finds study

COMMENTS

By Alice Cuddy • Updated: 08/09/2018



BOSTON

Problem definition

- Social media technologies are very prominent in guiding our access to information and news
- Little known about their contribution to falsity online

Problem definition

- Social media technologies are very prominent in guiding our access to information and news
- Little known about their contribution to falsity online

How and why do truth and falsity diffuse differently?

Are there categorical differences that influence diffusion?

What factors of human judgement explain these differences?

Some terminology

Some terminology

- "True" and "false" news
  - Any story or claim with an assertion in it
  - Any asserted claim made on Twitter

Some terminology

- “True” and “false” news
  - Any story or claim with an assertion in it
  - Any asserted claim made on Twitter
- Rumor
  - Social phenomena of a news story/claim spreading through Twitter
  - Involves sharing of claims between people

Some terminology

- “True” and “false” news
  - Any story or claim with an assertion in it
  - Any asserted claim made on Twitter
- Rumor
  - Social phenomena of a news story/claim spreading through Twitter
  - Involves sharing of claims between people
- Rumor cascade
  - Starts with a user making an assertion on Twitter (text, photos, links, etc.)
  - Unbroken retweet (RT) chain with a common & singular origin
  - Rumor diffusion through one or more cascades

Overview

- Investigating the differential diffusion of true, false, and mixed news stories
- Dataset of ~126,000 rumor cascades on Twitter
  - Spread by ~3M people more than ~4.5M times
  - Between 2006-2017. Access to all tweets ever posted.
  - Investigated by six independent fact-checking organizations (95-98% agreement on classifications)

Data collection

Data collection

- Collect all English tweets with link to the organizations’ websites (500k)

Data collection

- Collect all English tweets with link to the organizations’ websites (500k)
- Consider only the replies. For each reply tweet, extract the original tweet and all of its RTs (= RT cascades).

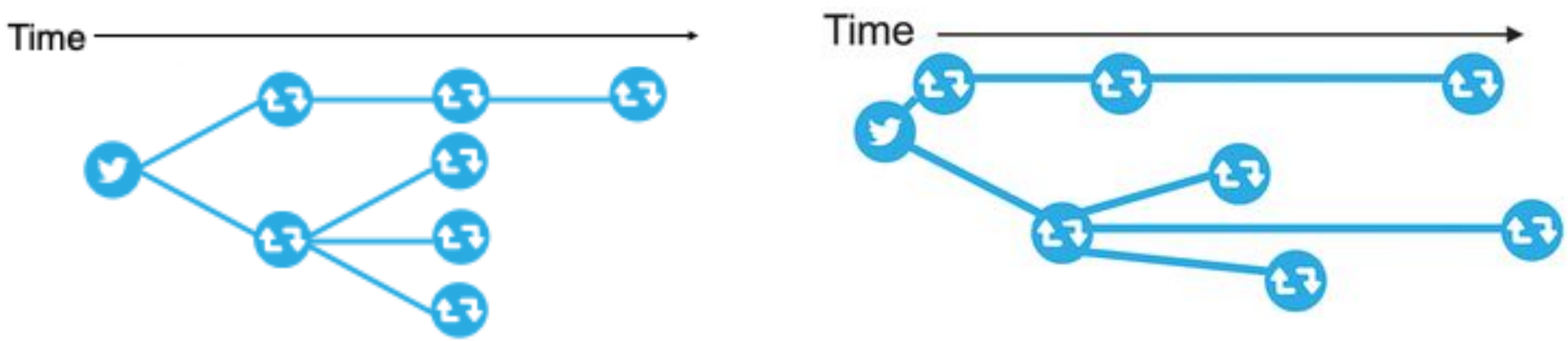
Data collection

- Collect all English tweets with link to the organizations' websites (500k)
- Consider only the replies. For each reply tweet, extract the original tweet and all of its RTs (= RT cascades).
- Each RT cascade = rumor propagating on Twitter
  - Veracity of the RT cascade - through the reply with link

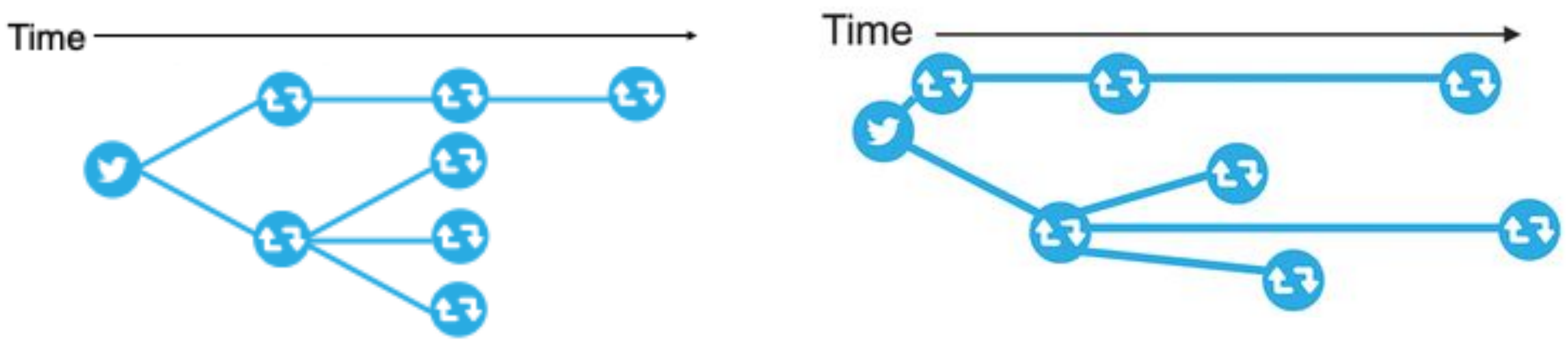
Data collection

- Collect all English tweets with link to the organizations' websites (500k)
- Consider only the replies. For each reply tweet, extract the original tweet and all of its RTs (= RT cascades).
- Each RT cascade = rumor propagating on Twitter
  - Veracity of the RT cascade - through the reply with link
- Manually and automatically ensure that replies address the original tweet
  - Only considered replies that target the original tweet (no replies to replies)
  - Compared the headline of the linked article to that of the original tweet
    - Cosine similarity

Rumor cascades

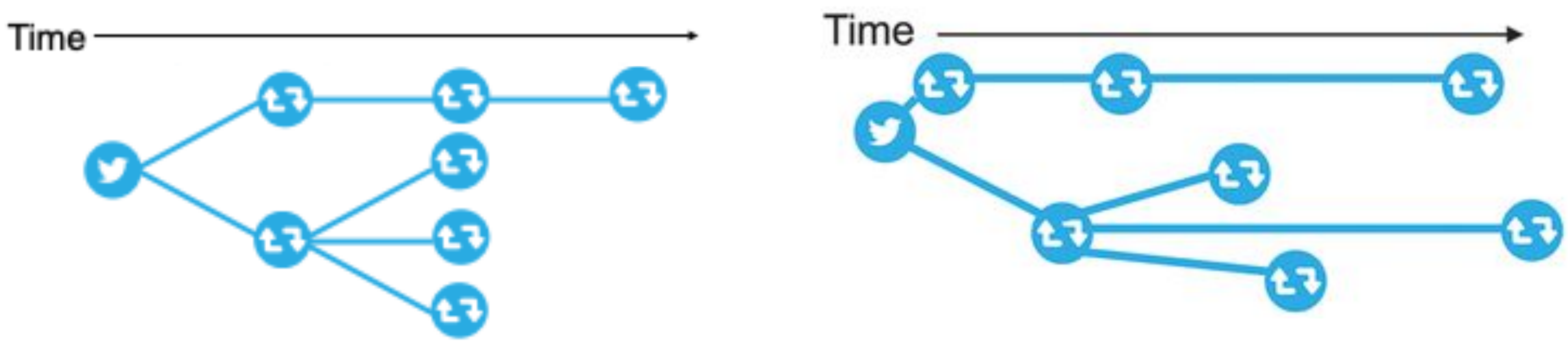


Rumor cascades



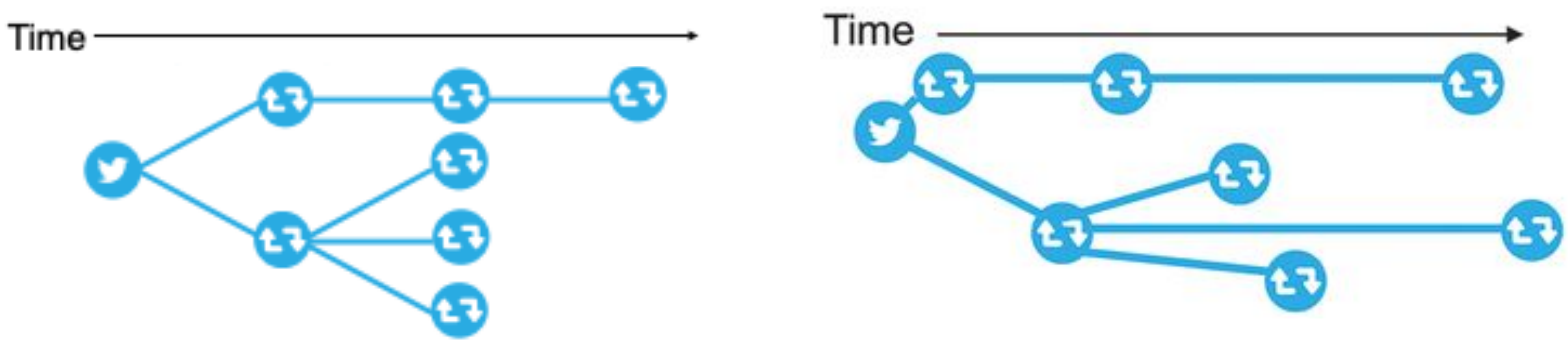
- **Depth:** number of RT hops (by a unique user) from the origin tweet over time

Rumor cascades



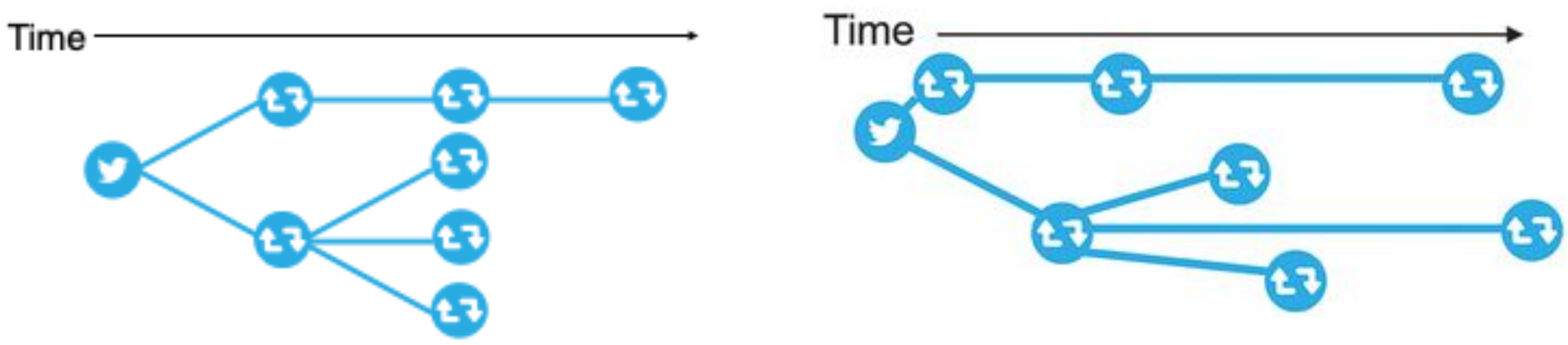
- **Depth:** number of RT hops (by a unique user) from the origin tweet over time
- **Size:** number of users involved in the cascade over time

Rumor cascades



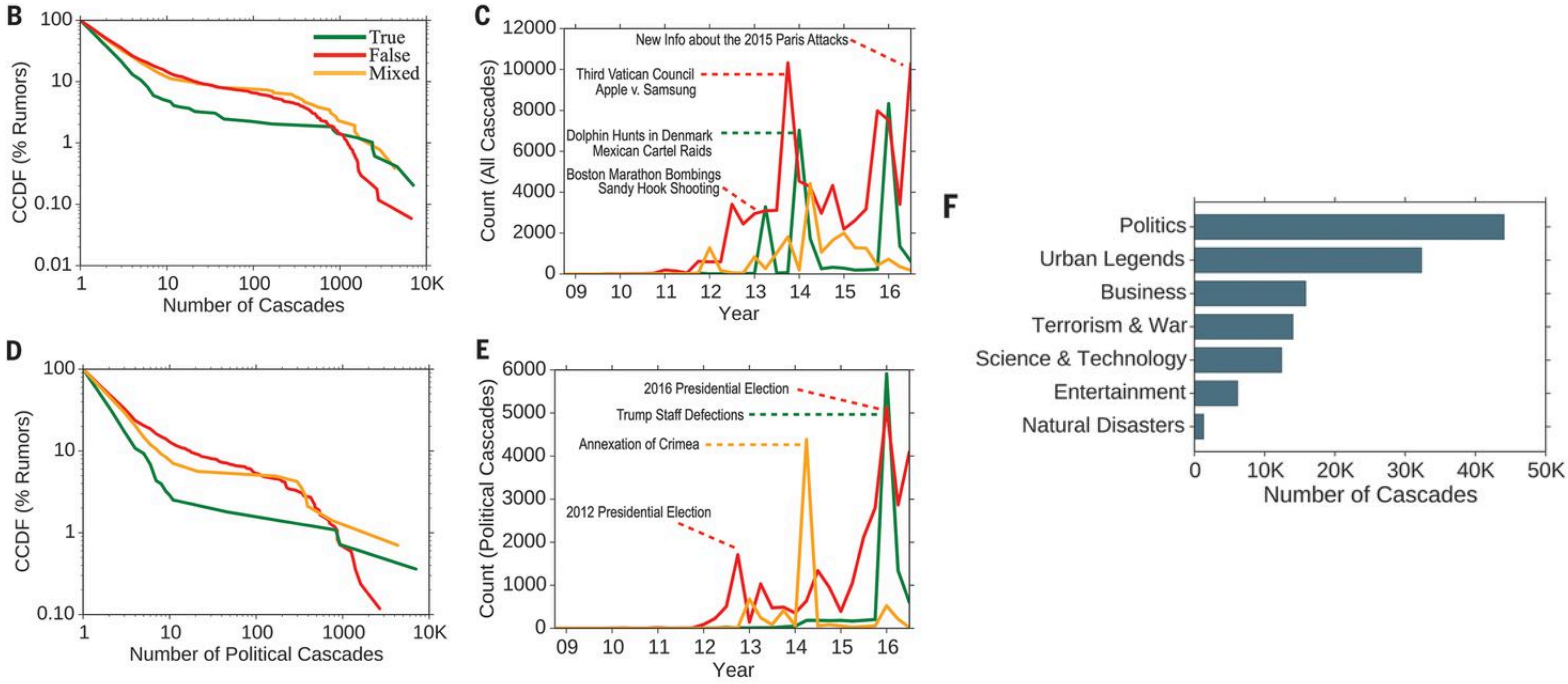
- **Depth:** number of RT hops (by a unique user) from the origin tweet over time
- **Size:** number of users involved in the cascade over time
- **Max. breadth:** max number of users in the cascade at any depth

Rumor cascades

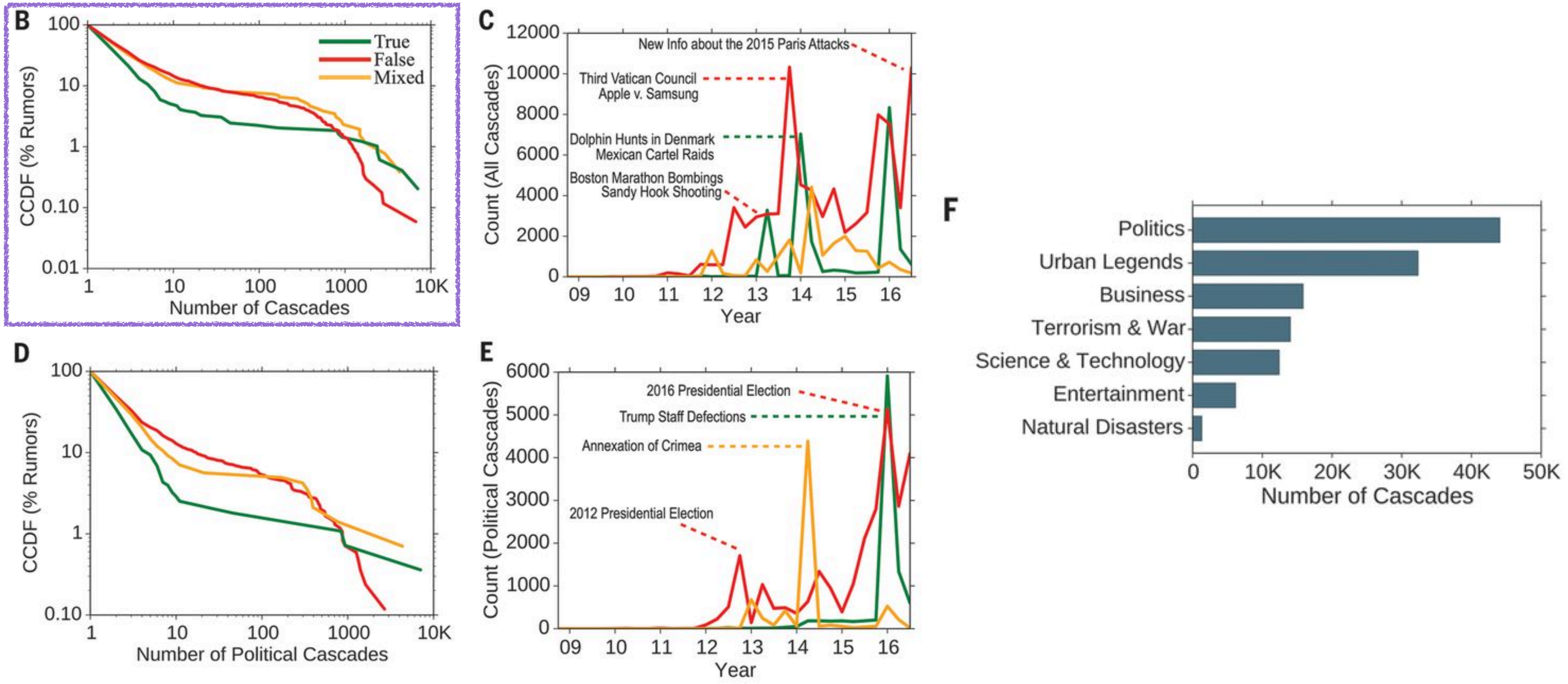


- **Depth:** number of RT hops (by a unique user) from the origin tweet over time
- **Size:** number of users involved in the cascade over time
- **Max. breadth:** max number of users in the cascade at any depth
- **Structural virality:** average distance between all pairs of nodes in the cascade

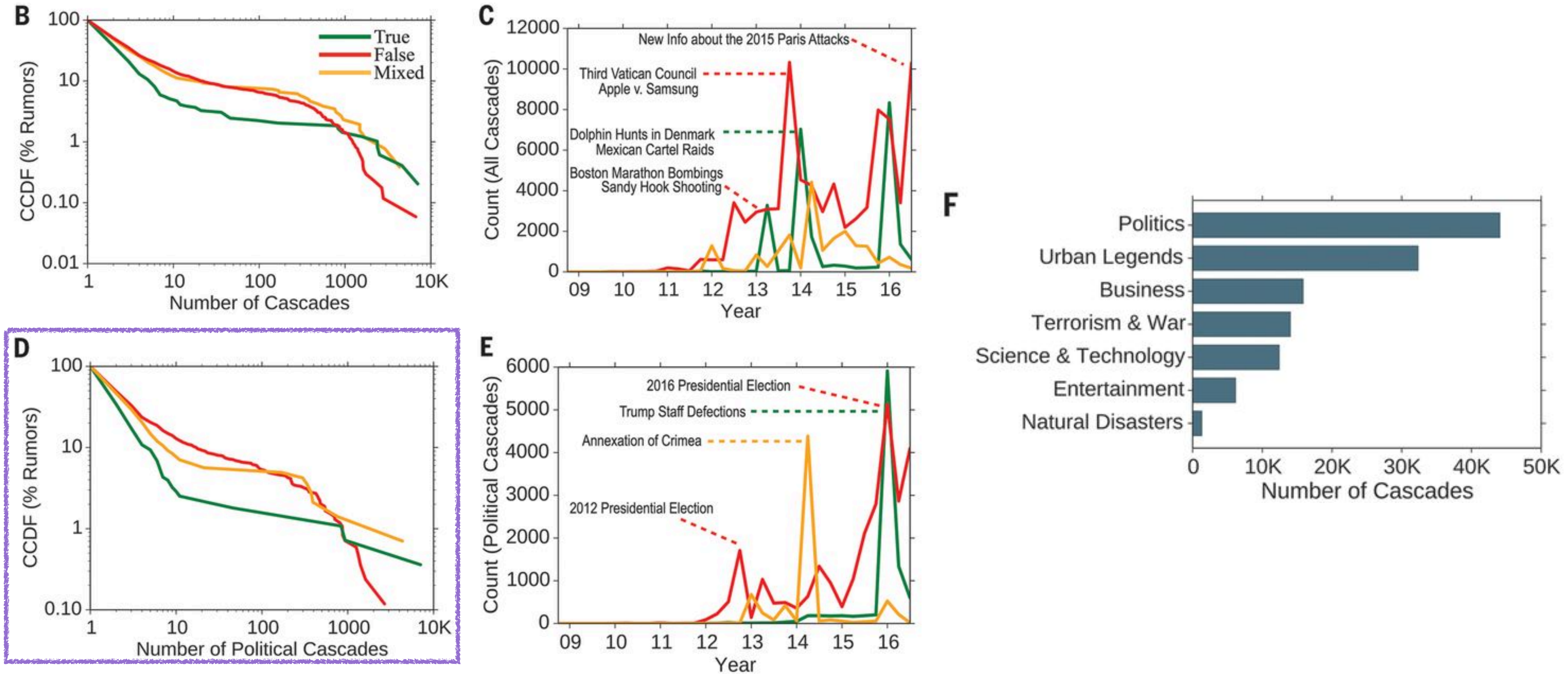
Rumor cascades



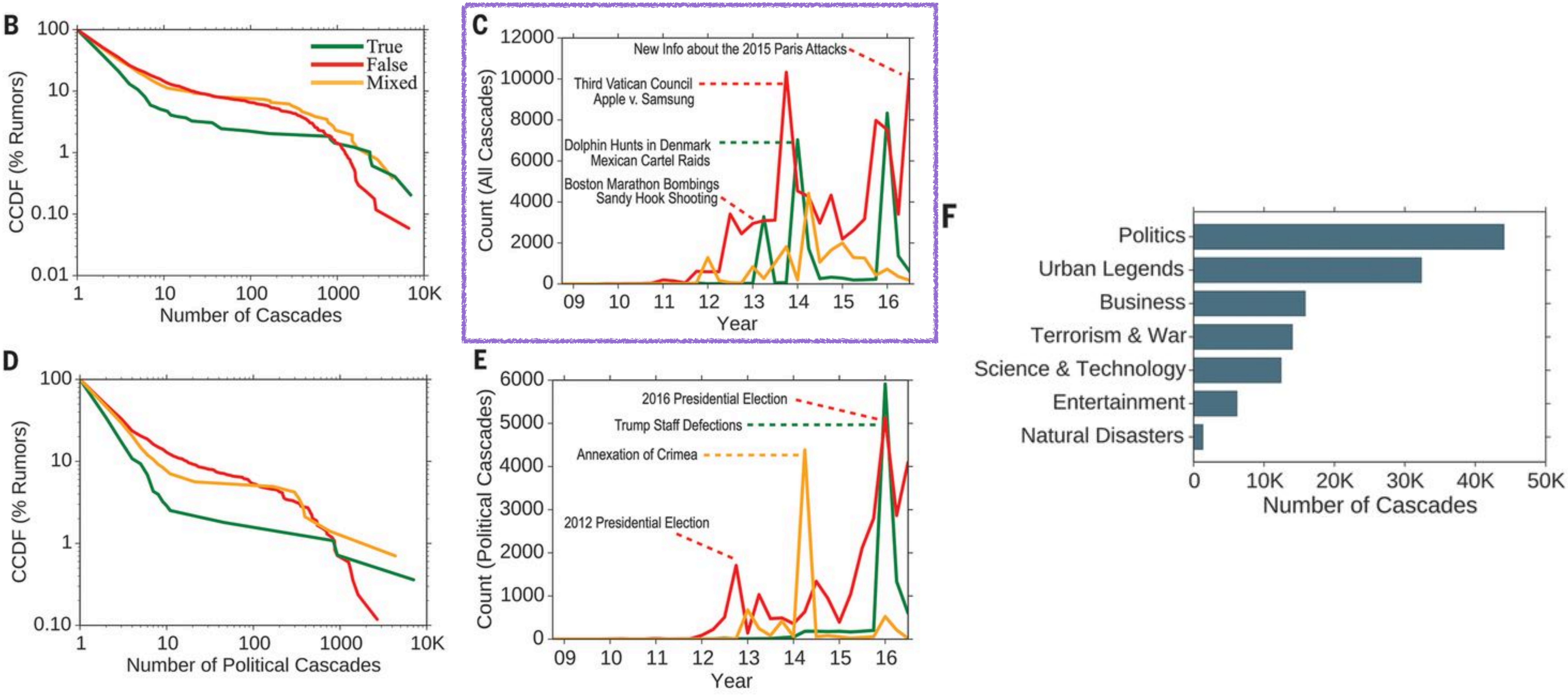
Rumor cascades



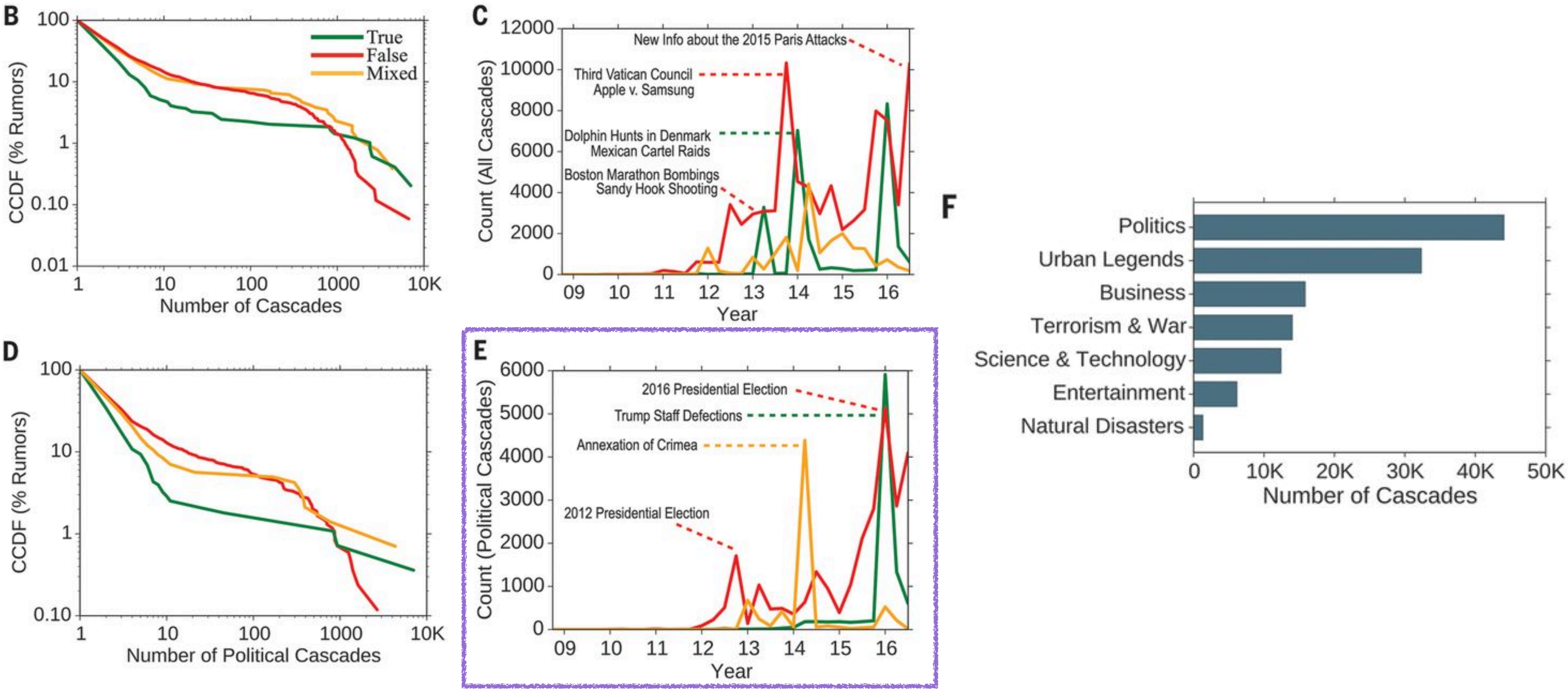
Rumor cascades



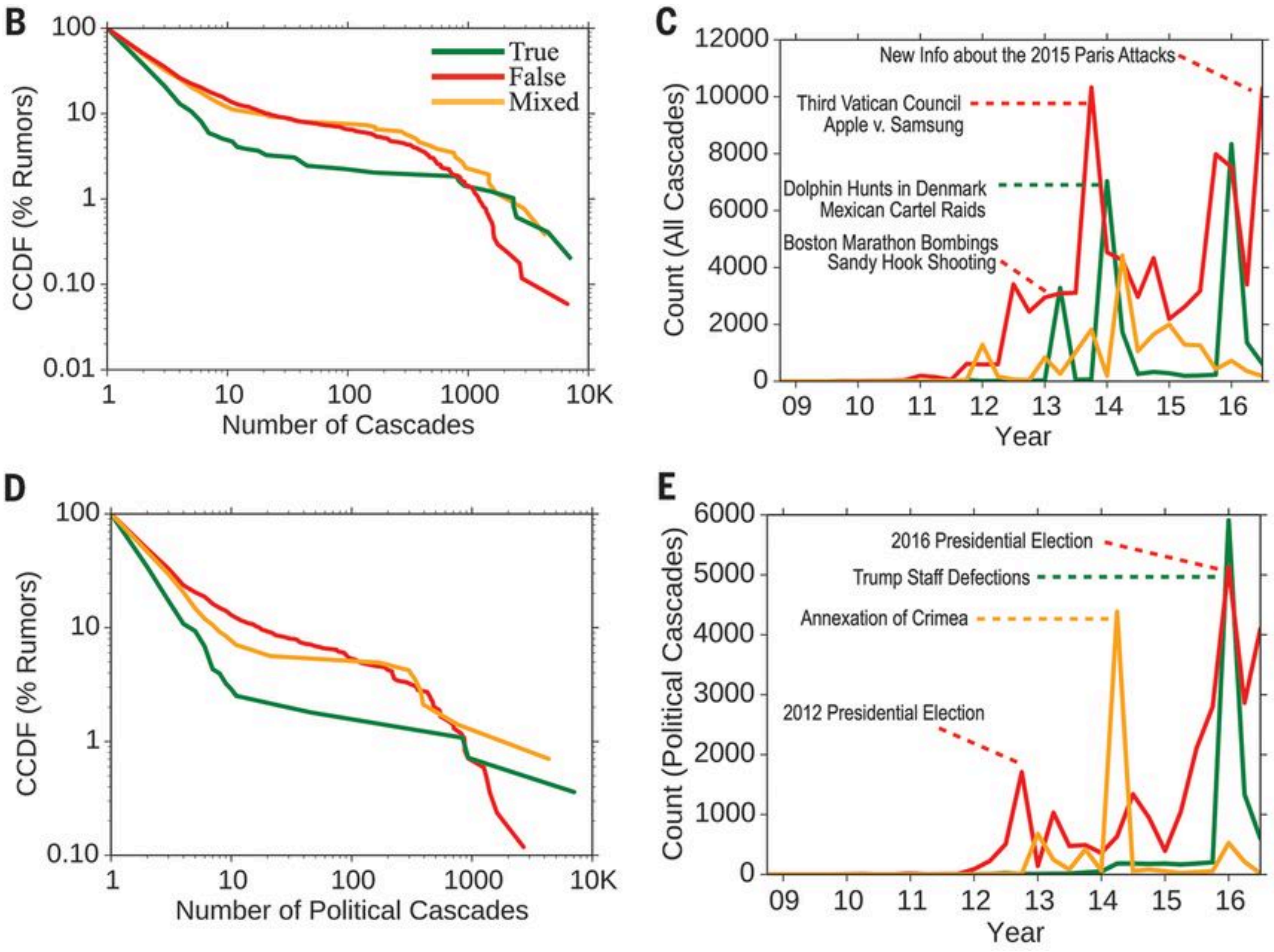
Rumor cascades

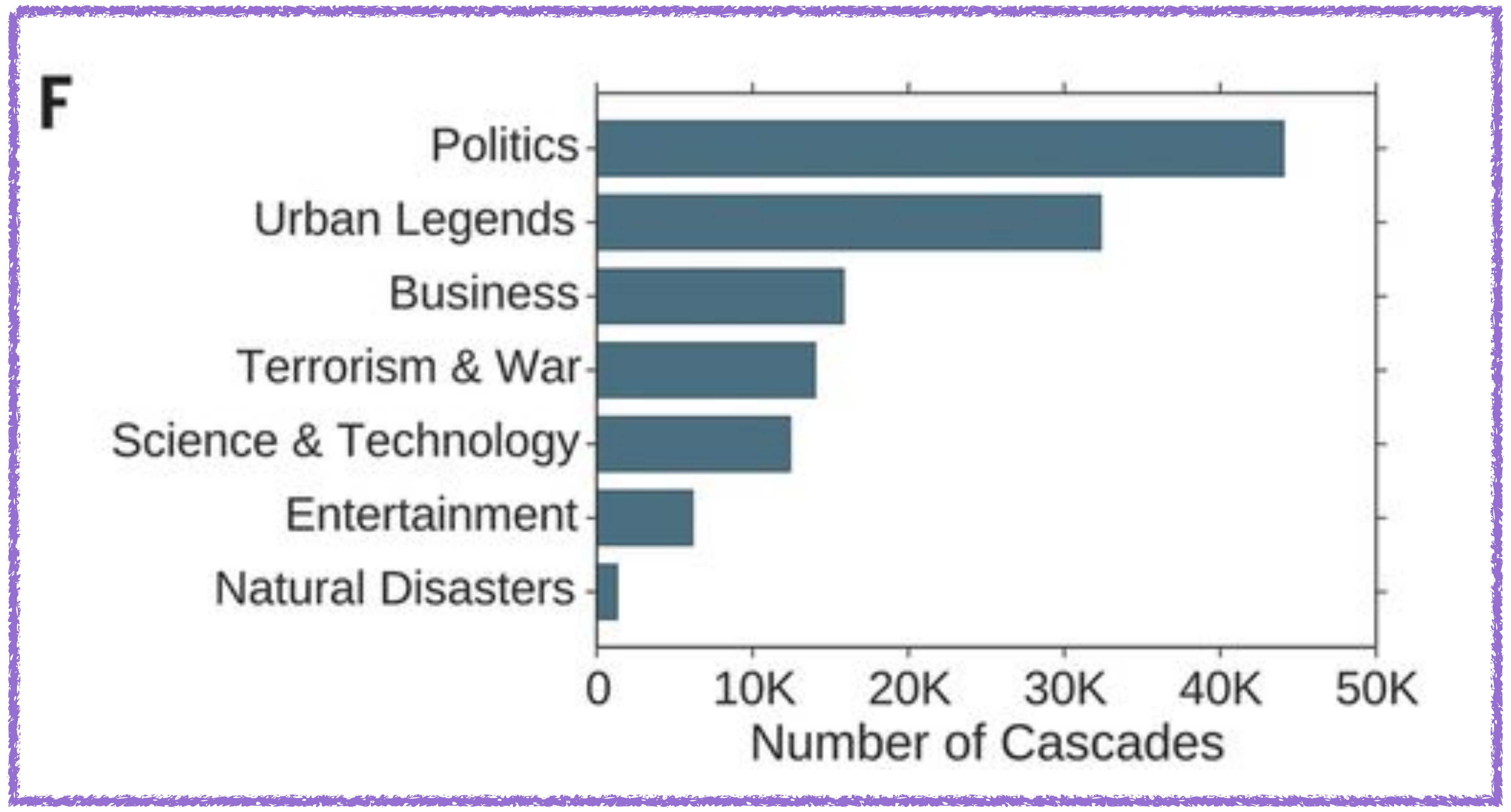


Rumor cascades

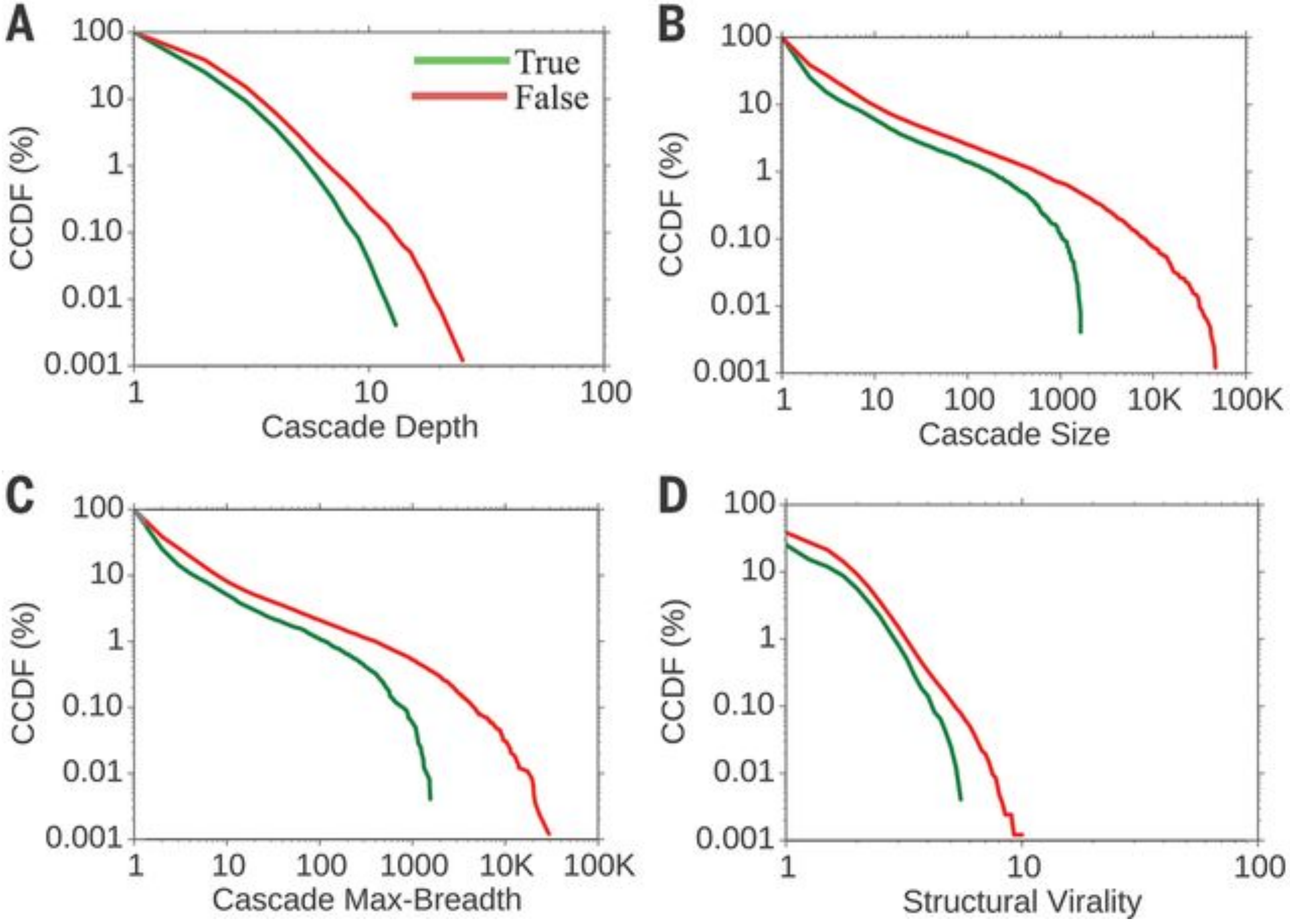


Rumor cascades

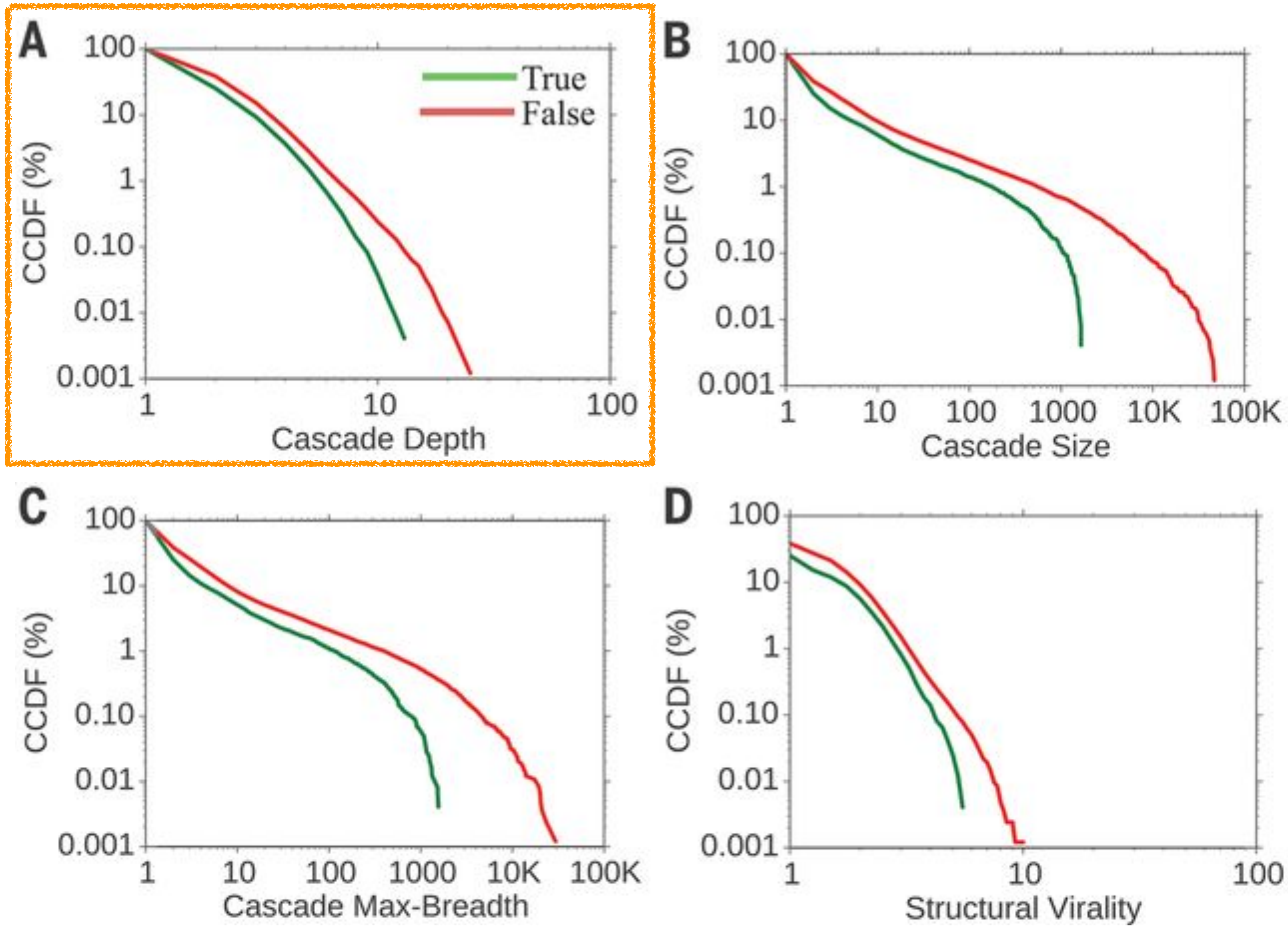




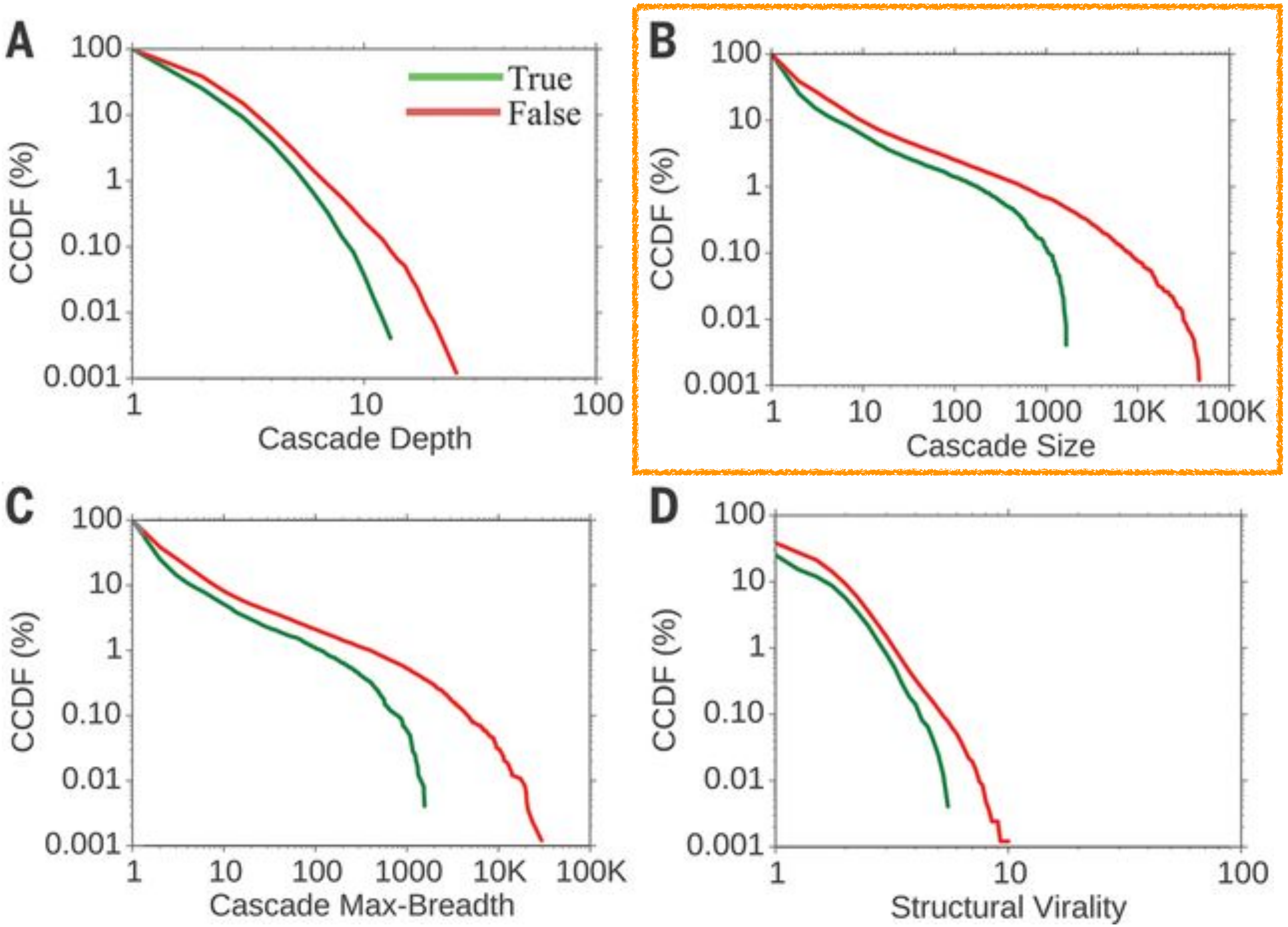
Diffusion dynamics of rumors



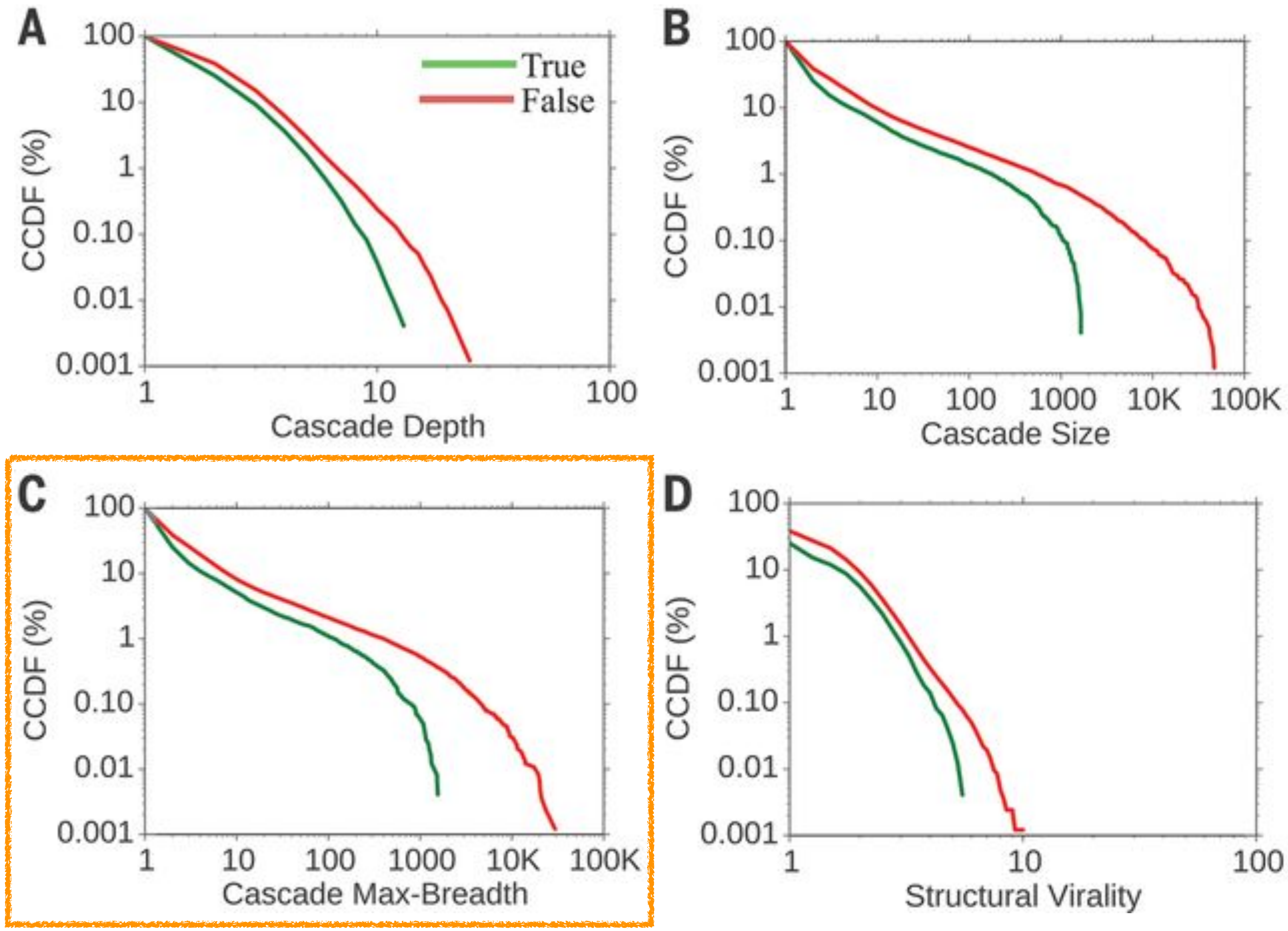
Diffusion dynamics of rumors



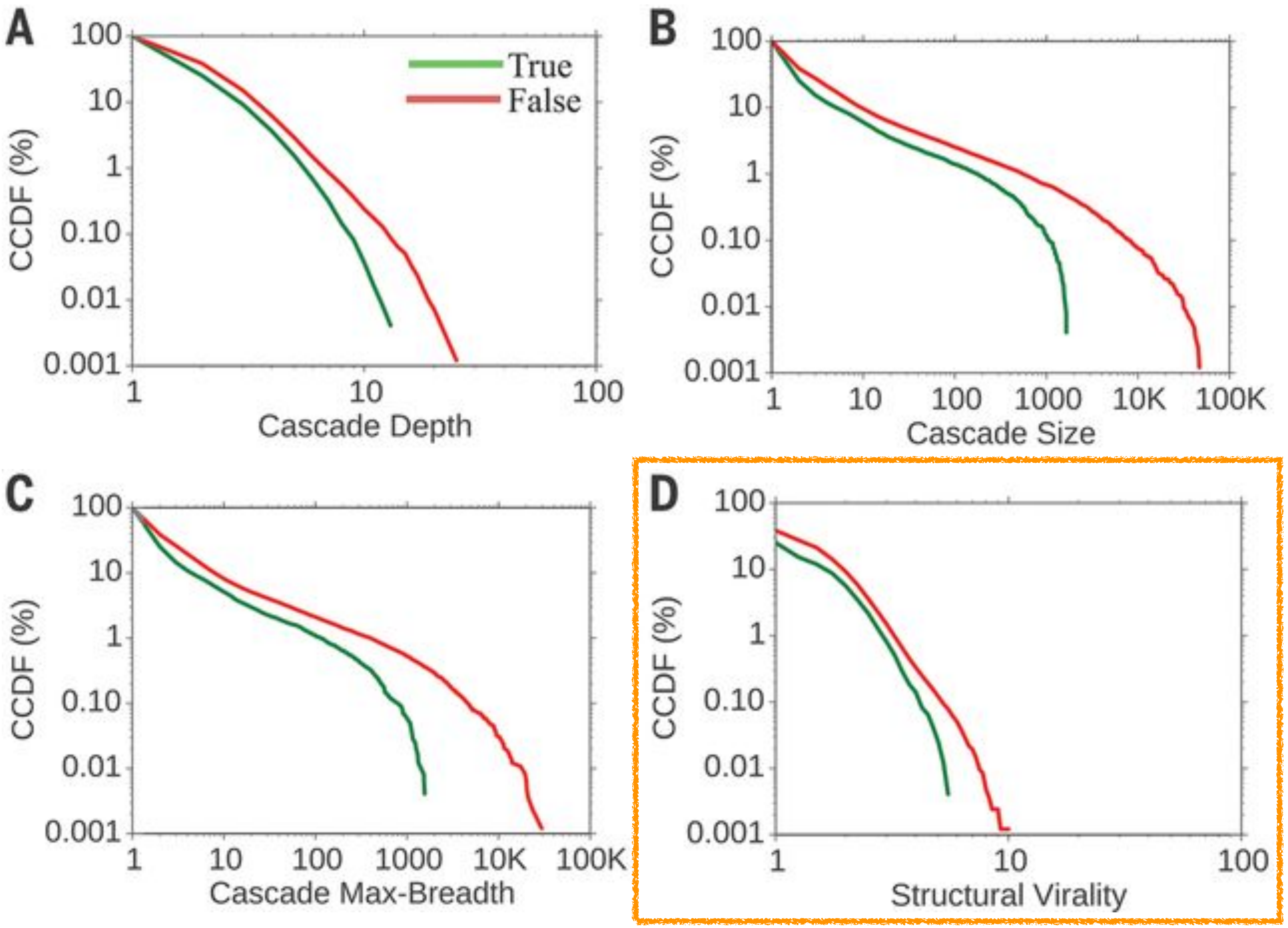
Diffusion dynamics of rumors



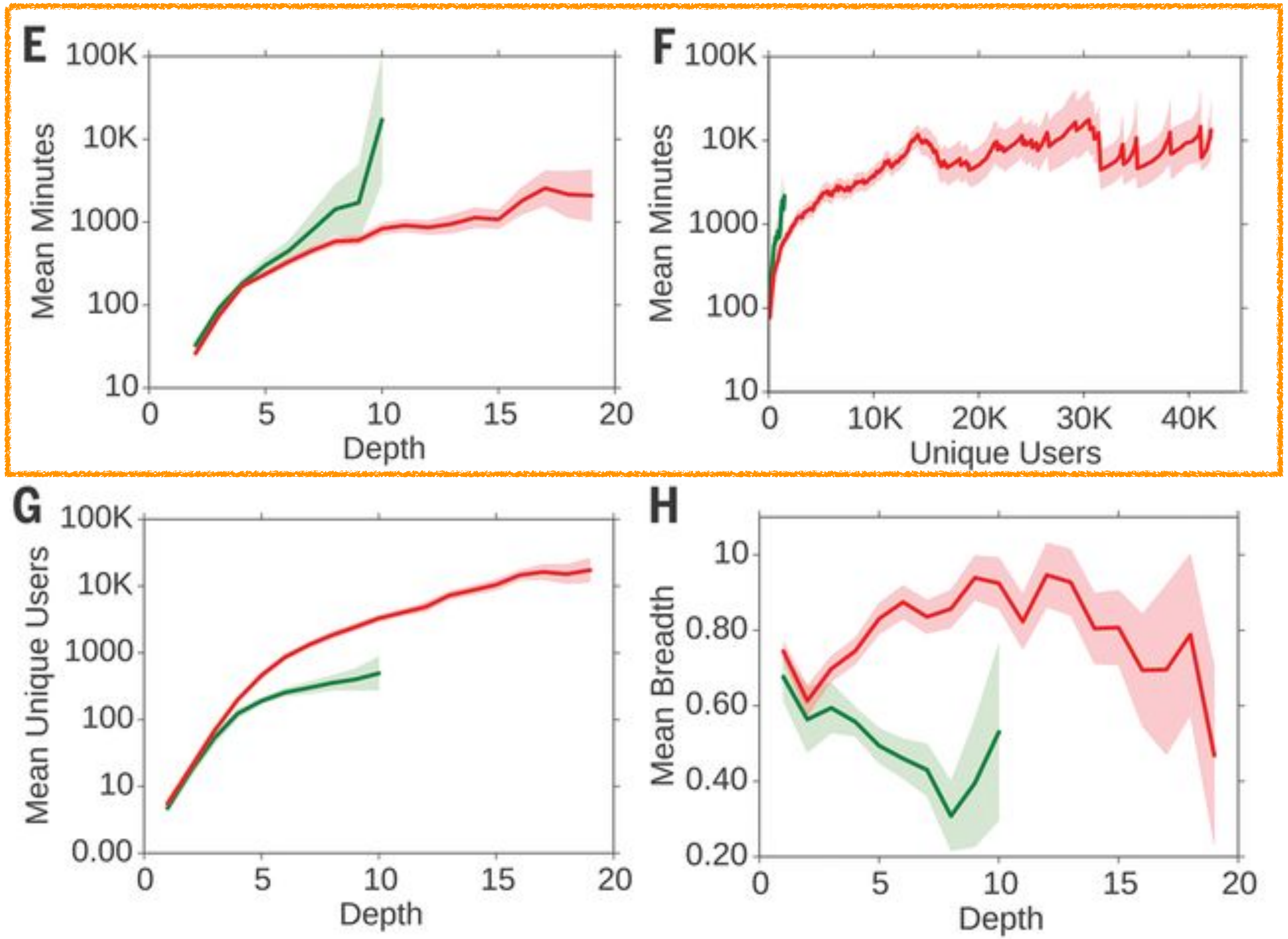
Diffusion dynamics of rumors



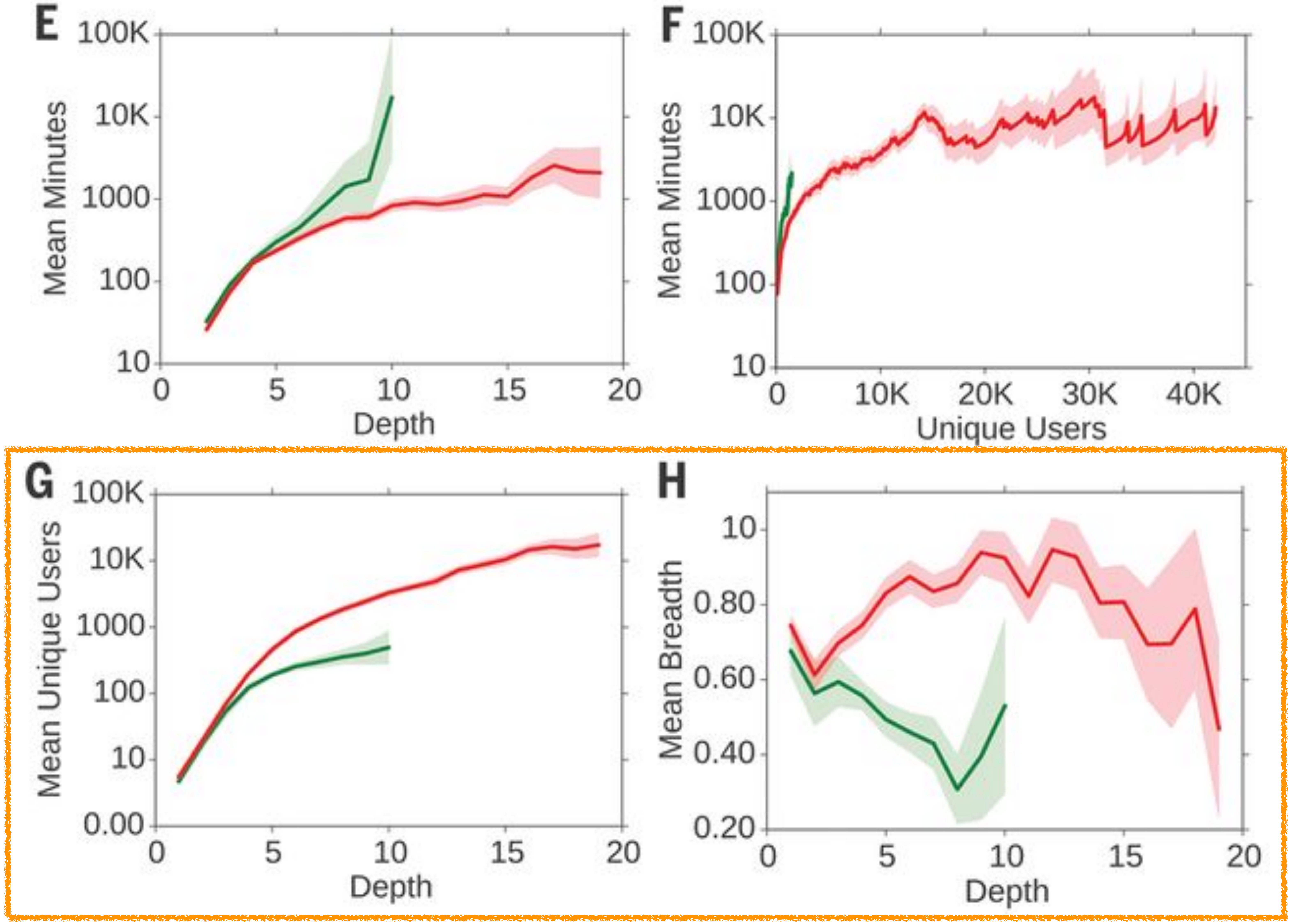
Diffusion dynamics of rumors



Diffusion dynamics of rumors

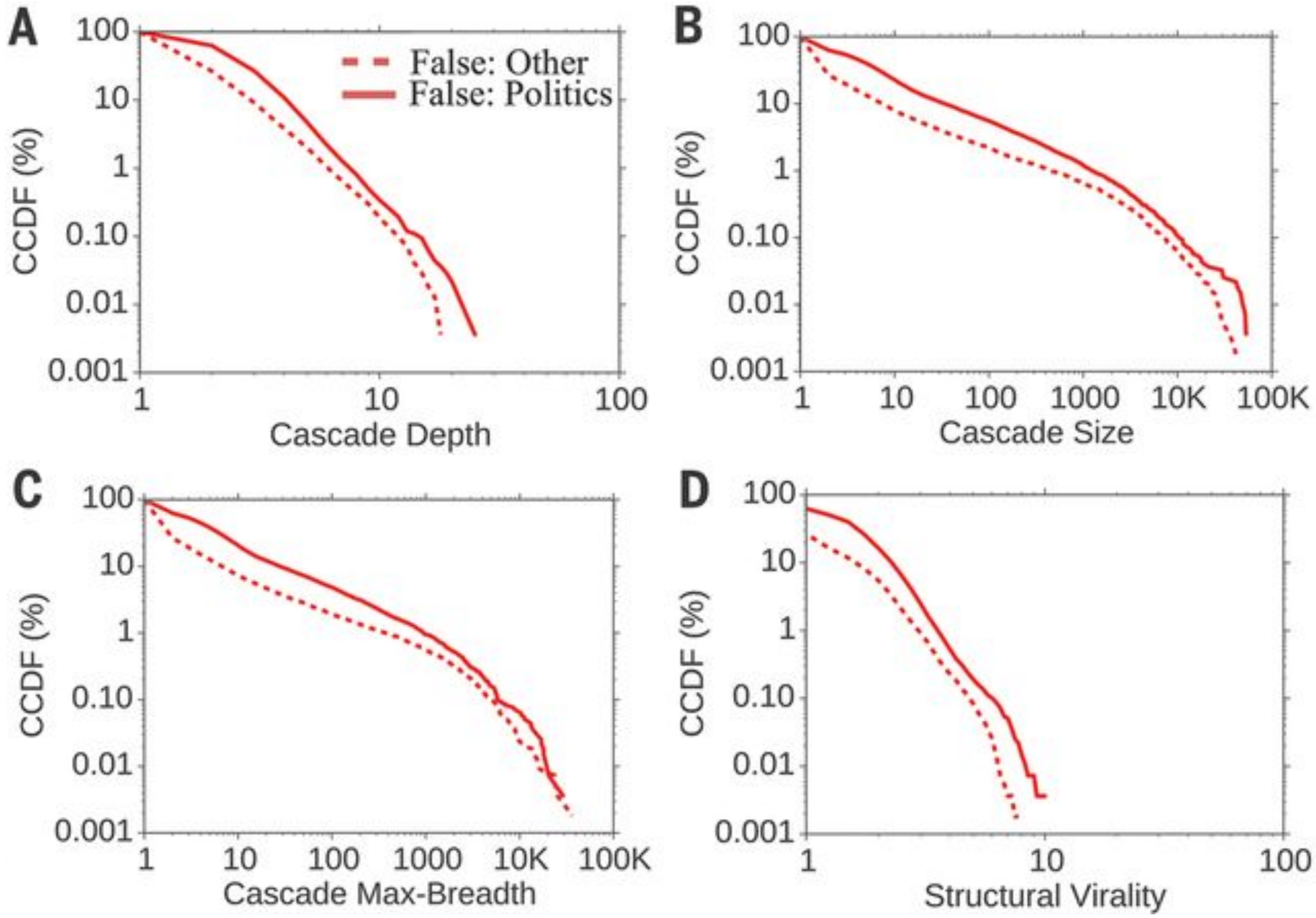


Diffusion dynamics of rumors

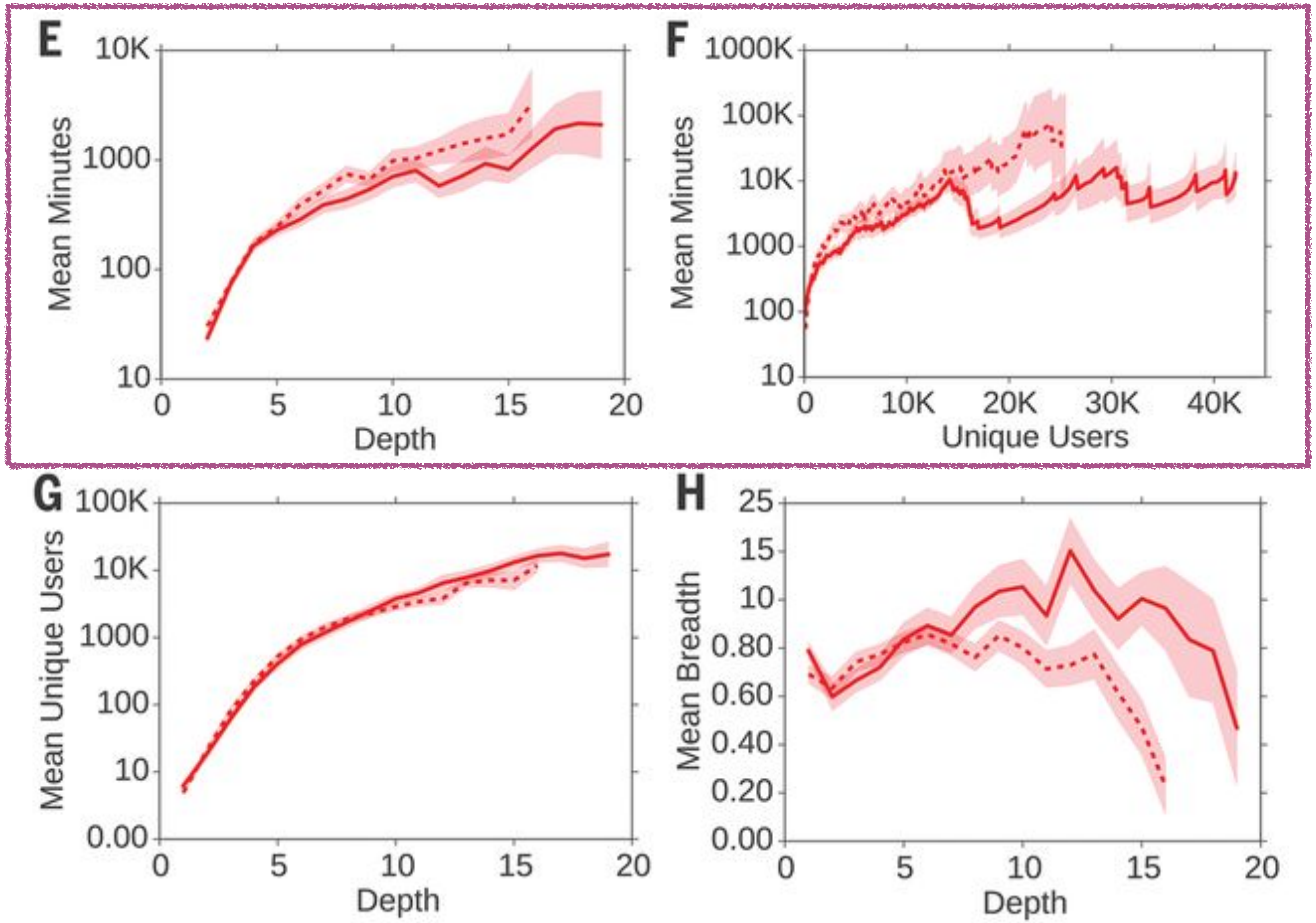


Falsehood diffused significantly <u>farther</u>, <u>faster</u>, <u>deeper</u> and <u>more broadly</u> than the truth in <u>all</u> categories of information.

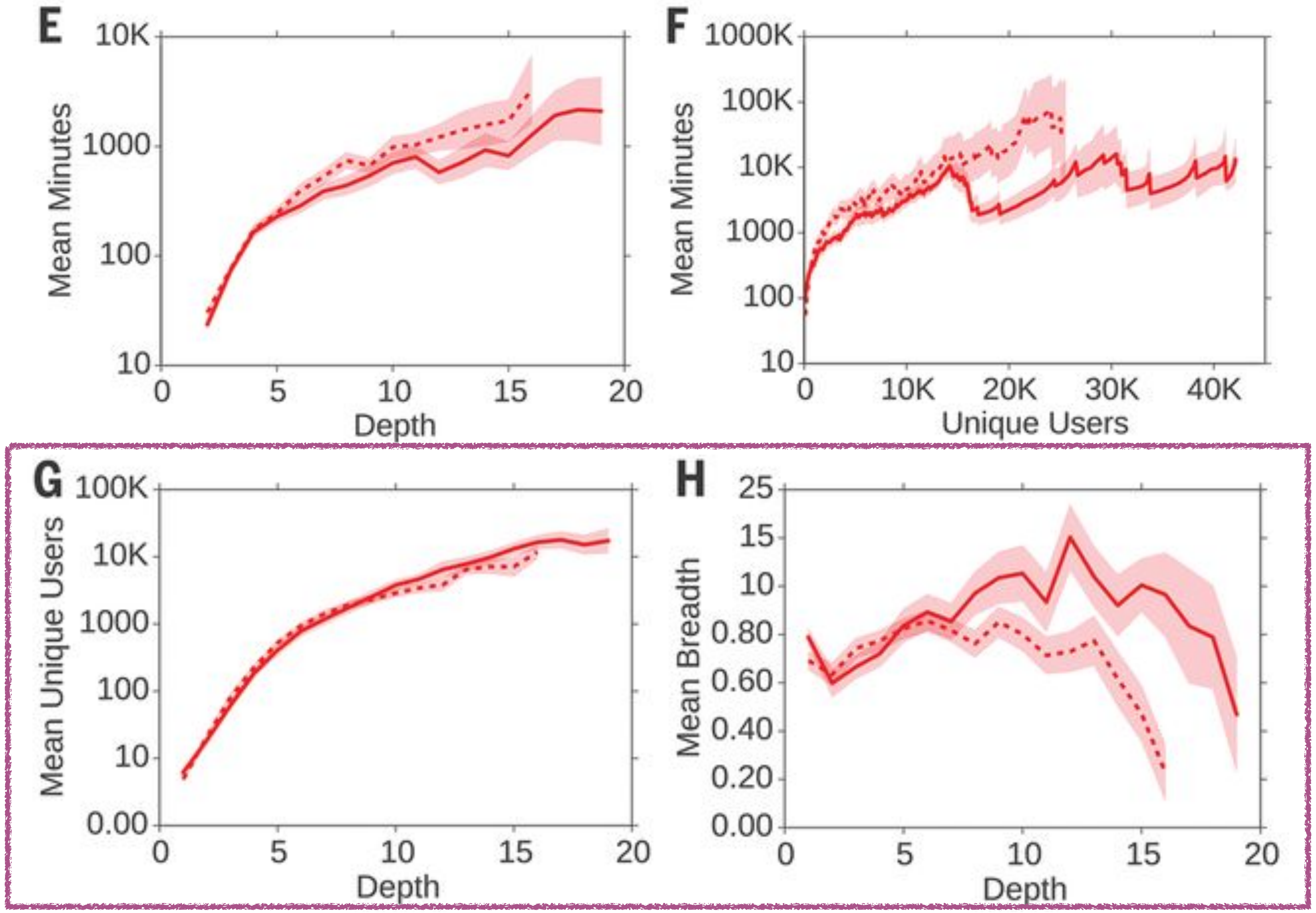
Diffusion dynamics of false political news



Diffusion dynamics of false political news



Diffusion dynamics of false political news



False political news diffused farther, faster, deeper and more broadly than any other type of false news.

Network structure

- What if the structural elements of the network or individual characteristics of the users explain why falsehood travels faster than the truth?

Network structure

- What if the structural elements of the network or individual characteristics of the users explain why falsehood travels faster than the truth?

A

	median		mean		mean (log)		stdv (log)		ks-test
	false	true	false	true	false	true	false	true	
followers	410	466	2234	5240	2.62	2.68	0.69	0.88	D=0.104, p~0.0
followees	383	509	1002	1707	2.59	2.72	0.85	0.96	D=0.136, p~0.0
verified	0	0	0.002	0.006	nd	nd	nd	nd	D=0.005, p<0.001
engagement	9.52	9.54	19.70	24.65	0.91	0.90	0.65	0.76	D=0.054, p~0.0
account age	982	1214	1072	1269	2.90	2.97	0.39	0.42	D=0.125, p~0.0

Network structure

- What if the structural elements of the network or individual characteristics of the users explain why falsehood travels faster than the truth?

A

	median		mean		mean (log)		stdv (log)		ks-test
	false	true	false	true	false	true	false	true	
followers	410	466	2234	5240	2.62	2.68	0.69	0.88	D=0.104, p~0.0
followees	383	509	1002	1707	2.59	2.72	0.85	0.96	D=0.136, p~0.0
verified	0	0	0.002	0.006	nd	nd	nd	nd	D=0.005, p<0.001
engagement	9.52	9.54	19.70	24.65	0.91	0.90	0.65	0.76	D=0.054, p~0.0
account age	982	1214	1072	1269	2.90	2.97	0.39	0.42	D=0.125, p~0.0

Falsehood diffused farther and faster than the truth despite the differences in the network structure, not because of them.

Novelty of information

Novelty of information

- Novel information:
  - Attracts human attention
  - Contributes to decision-making
  - Encourages information sharing
  - More valuable both from an information theoretic perspective and from a social perspective

Novelty of information

- Novel information:
  - Attracts human attention
  - Contributes to decision-making
  - Encourages information sharing
  - More valuable both from an information theoretic perspective and from a social perspective
- **Q1:** Was falsity more novel than the truth?

Novelty of information

- Novel information:
  - Attracts human attention
  - Contributes to decision-making
  - Encourages information sharing
  - More valuable both from an information theoretic perspective and from a social perspective
- **Q1:** Was falsity more novel than the truth?
- **Q2:** Were Twitter users more likely to RT information that was more novel?

Novelty of information

Novelty of information

- Randomly selected ~5000 users

Novelty of information

- Randomly selected ~5000 users
- Extracted a random sample of ~25,000 tweets they were exposed to in the 60 days prior to their RT of a rumor

Novelty of information

- Randomly selected ~5000 users
- Extracted a random sample of ~25,000 tweets they were exposed to in the 60 days prior to their RT of a rumor
- Latent Dirichlet Allocation Topic model with 200 topics and trained on 10M English tweets
  - Calculate information distance between the rumor tweets and exposed tweets

Novelty of information

- Randomly selected ~5000 users
- Extracted a random sample of ~25,000 tweets they were exposed to in the 60 days prior to their RT of a rumor
- Latent Dirichlet Allocation Topic model with 200 topics and trained on 10M English tweets
  - Calculate information distance between the rumor tweets and exposed tweets
- Compare topic distributions of the rumor tweets and exposed tweets

Novelty of information

- Randomly selected ~5000 users
- Extracted a random sample of ~25,000 tweets they were exposed to in the 60 days prior to their RT of a rumor
- Latent Dirichlet Allocation Topic model with 200 topics and trained on 10M English tweets
  - Calculate information distance between the rumor tweets and exposed tweets
- Compare topic distributions of the rumor tweets and exposed tweets

False rumors were significantly more novel than the truth.

Perceived novelty

Perceived novelty

- False rumors are <u>measurably</u> more novel than true rumors

Perceived novelty

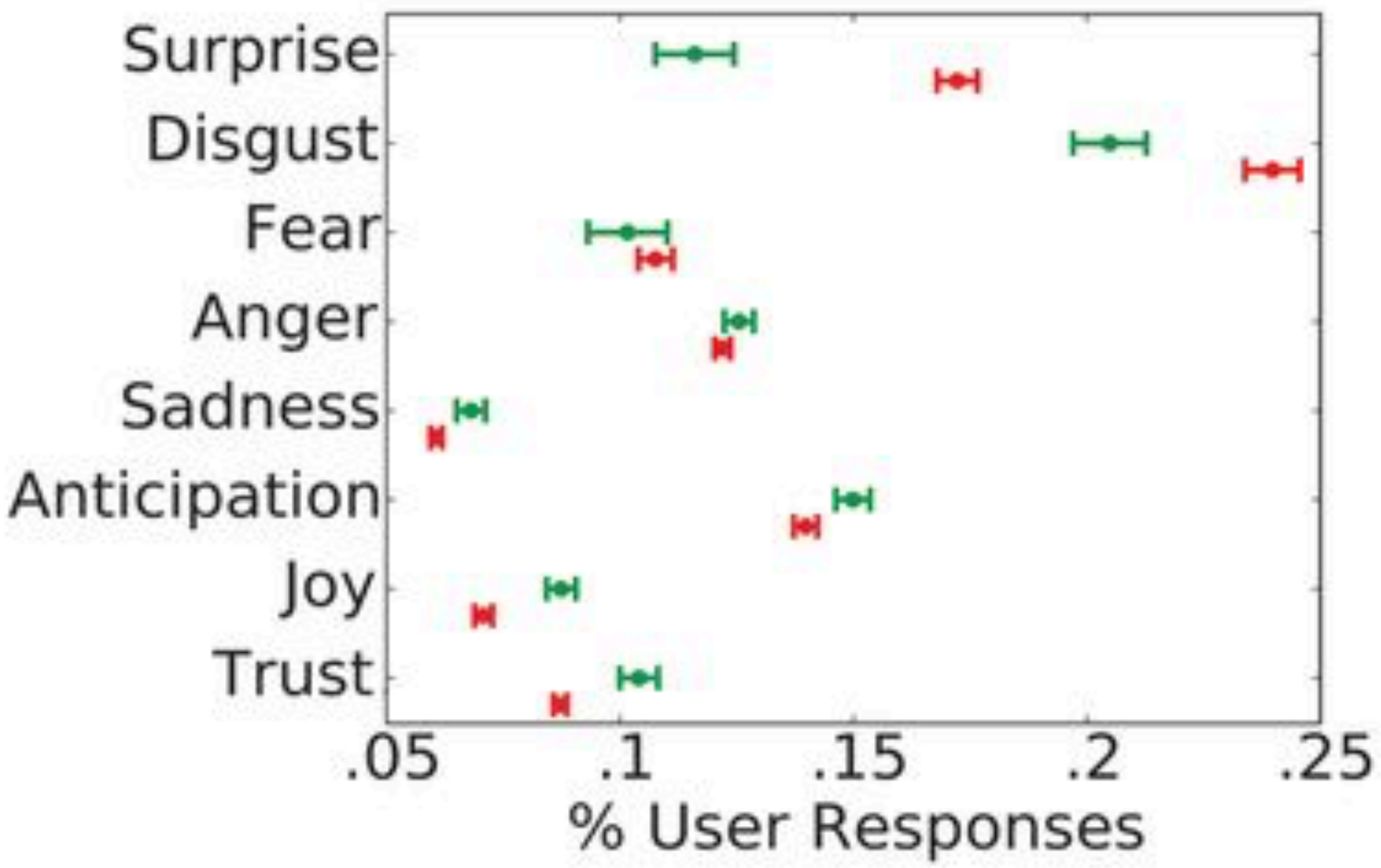
- False rumors are measurably more novel than true rumors
- But is this how users perceive them as well?

Perceived novelty

- False rumors are measurably more novel than true rumors
- But is this how users perceive them as well?
- Compared the emotional content of replies to true and false rumors
  - ~140,000 English words & their associations with eight emotions
  - ~32,000 Twitter hashtags & their weighted associations with the same emotions

Perceived novelty

- False rumors are measurably more novel than true rumors
- But is this how users perceive them as well?
- Compared the emotional content of replies to true and false rumors
  - ~140,000 English words & their associations with eight emotions
  - ~32,000 Twitter hashtags & their weighted associations with the same emotions



	mean		variance		ks-test
	false	true	false	true	
surprise	0.172	0.116	0.0167	0.0072	D=0.205, p~0.0
disgust	0.240	0.205	0.0260	0.0227	D=0.102, p~0.0
fear	0.108	0.102	0.0120	0.0095	D=0.021, p~0.164
anger	0.122	0.126	0.0074	0.0111	D=0.023, p~0.078
sadness	0.061	0.068	0.0038	0.0065	D=0.037, p~0.0
anticipation	0.140	0.150	0.0093	0.0154	D=0.038, p~0.0
joy	0.071	0.087	0.0054	0.0104	D=0.061, p~0.0
trust	0.087	0.104	0.0058	0.0119	D=0.060, p~0.0

Validation of findings (1)

Validation of findings (1)

- Multiple cascades for every true and false rumor

Validation of findings (1)

- Multiple cascades for every true and false rumor
- Variance and error terms associated with cascades corresponding to the same rumor will be correlated

Validation of findings (1)

- Multiple cascades for every true and false rumor
- Variance and error terms associated with cascades corresponding to the same rumor will be correlated
- Solution:
  - Specified cluster-robust standard errors and calculated variance at rumor-cluster level
  - Clustering reduced precision but the significance of the results did not change

Validation of findings (2)

Validation of findings (2)

- Potential selection bias
  - Restricting sample to tweets that are fact-checked by the six organizations
  - Fact checking may select certain types of rumors or draw additional attention to them

Validation of findings (2)

- Potential selection bias
  - Restricting sample to tweets that are fact-checked by the six organizations
  - Fact checking may select certain types of rumors or draw additional attention to them
- Solution:
  - Independently verified another sample of rumor cascades that was not verified by any fact-checking organizations
  - Manually fact-checked by three undergrads (agreed on the veracity of 90% of the 13,240 cascades)
  - Diffusion dynamics of true & false rumors are nearly identical to the main data set

Validation of findings (3)

Validation of findings (3)

- Potential presence of bots that can bias conclusions about human judgement

Validation of findings (3)

- Potential presence of bots that can bias conclusions about human judgement
- Solution:
  - Bot-detection algorithm to identify and remove all bots before the analysis
  - Adding back bot traffic into the analysis did not change the main conclusions
  - Inclusion of bots accelerated the spread of both true and false news and affected their spread equally

Conclusion

Conclusion

- False news spread more pervasively than the truth online
  - Faster, farther, deeper and more broadly than the truth in all categories of information

Conclusion

- False news spread more pervasively than the truth online
  - Faster, farther, deeper and more broadly than the truth in all categories of information
- Effects are more pronounced for political news

Conclusion

- False news spread more pervasively than the truth online
  - Faster, farther, deeper and more broadly than the truth in all categories of information
- Effects are more pronounced for political news
- Network structure and individual characteristics of spreaders do not favor promoting false news

Conclusion

- False news spread more pervasively than the truth online
  - Faster, farther, deeper and more broadly than the truth in all categories of information
- Effects are more pronounced for political news
- Network structure and individual characteristics of spreaders do not favor promoting false news
- False news are more novel than true news - greater likelihood of being shared with others

Conclusion

- False news spread more pervasively than the truth online
  - Faster, farther, deeper and more broadly than the truth in all categories of information
- Effects are more pronounced for political news
- Network structure and individual characteristics of spreaders do not favor promoting false news
- False news are more novel than true news - greater likelihood of being shared with others
- Bots accelerated the spread of true and false news at the same rate