Multidimensional Analysis of YouTube Communities in the Indo-Pacific Region

Ugochukwu Onyepunuka (uponyepunuka@ualr.edu), Thomas Marcoux, Mainuddin Shaik, Mayor Inna Gurung, Nitin Agarwal
Collaboratorium for Social Media and Online Behavioral Studies (COSMOS)

cosmos.ualr.edu

Dr. Nitin Agarwal

COSMOS Director - Leading projects with a combined funding of over $20 million from various agencies

- Department of Defense
- DARPA
- Department of State
- National Science Foundation

Ugochukwu Onyepunuka

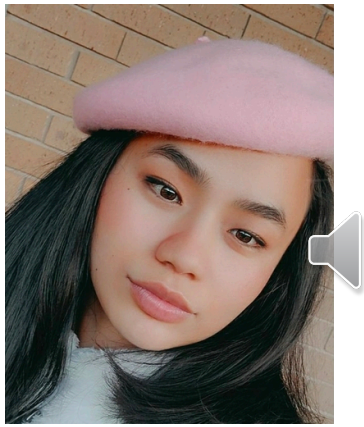
COSMOS Graduate Student

- Data Analysis
- Topic Modelling
- Data Engineering

Dr. Thomas Marcoux

Mainuddin Shaik

Ma, Mayor Inna Gurung



Why?

- Present a pathway on how to analyze video-based platform
- Provide insights into the dynamics of their online discourse
- To obtain situational awareness during any social cyber-influence operation
- To identify suspicious and inorganic content in cyber-influence operations

Agenda

1. Data collection
2. Methodology
3. Results
4. Future directions

Data collection

- 3 narrative themes (China, Uyghur, Covid)
- Group of keywords for each narrative theme provided by subject matter experts at ASU
- Times series data based on channel's engagement stats

VIDEO COLLECTION STATISTICS

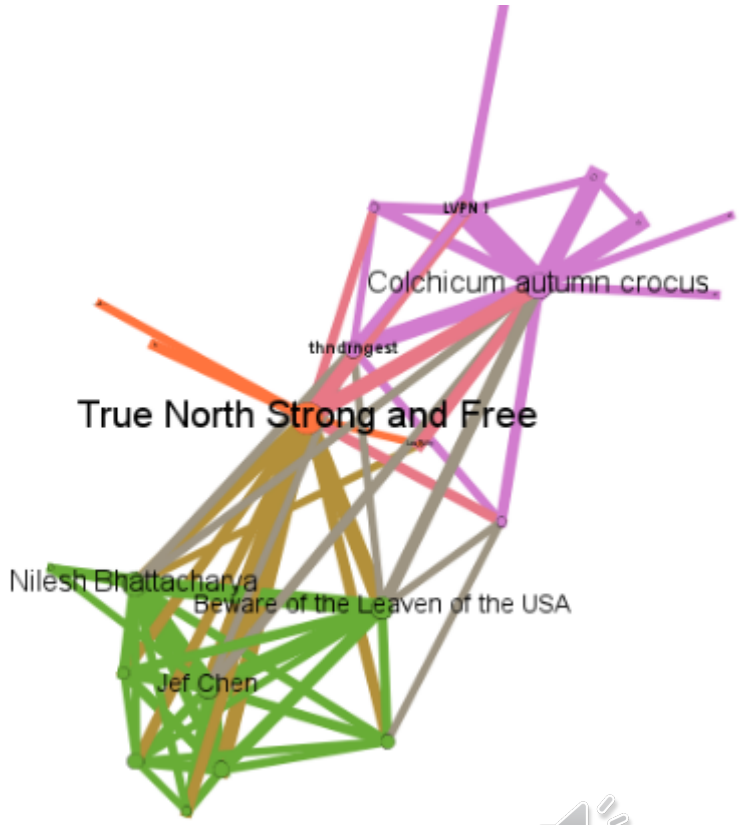
Videos	25,673
Channels	6,806
Comments	5,538,730

NARRATIVE EXTRACTION SAMPLE

Narrative	Date	Keywords
China	2018 - 2021	'Komunis Cina—China pengaruh Indonesia', 'Menguasai Cina–
Uyghur	2018 - 2021	'Uighur—Uyghur Indonesia', 'Penindasan Uighur—Uyghur beb

Methodology

- Narrative Segregation
- Engagement Trend Analysis
- Network Analysis



Narrative Segregation

We collected data from YouTube using a group of keywords relating to each narrative theme.

The relevant keywords were identified by studying coverage in the Indo-Pacific region with further reviews to improve the inclusiveness

To perform an analysis on a narrative we queried our database using the keywords in the full-text search query

NARRATIVE EXTRACTION SAMPLE

Narrative	Date	Keywords
China	2018 - 2021	'Komunis Cina—China pengaruh Indonesia', 'Menguasai Cina—China—Tiongkok— Tionghoa ekonomi Indonesia', ...
Uyghur	2018 - 2021	'Uighur—Uyghur Indonesia', 'Penindasan Uighur—Uyghur bebaskan', 'Kejam Uighur—Uyghur', ...

Engagement Analysis

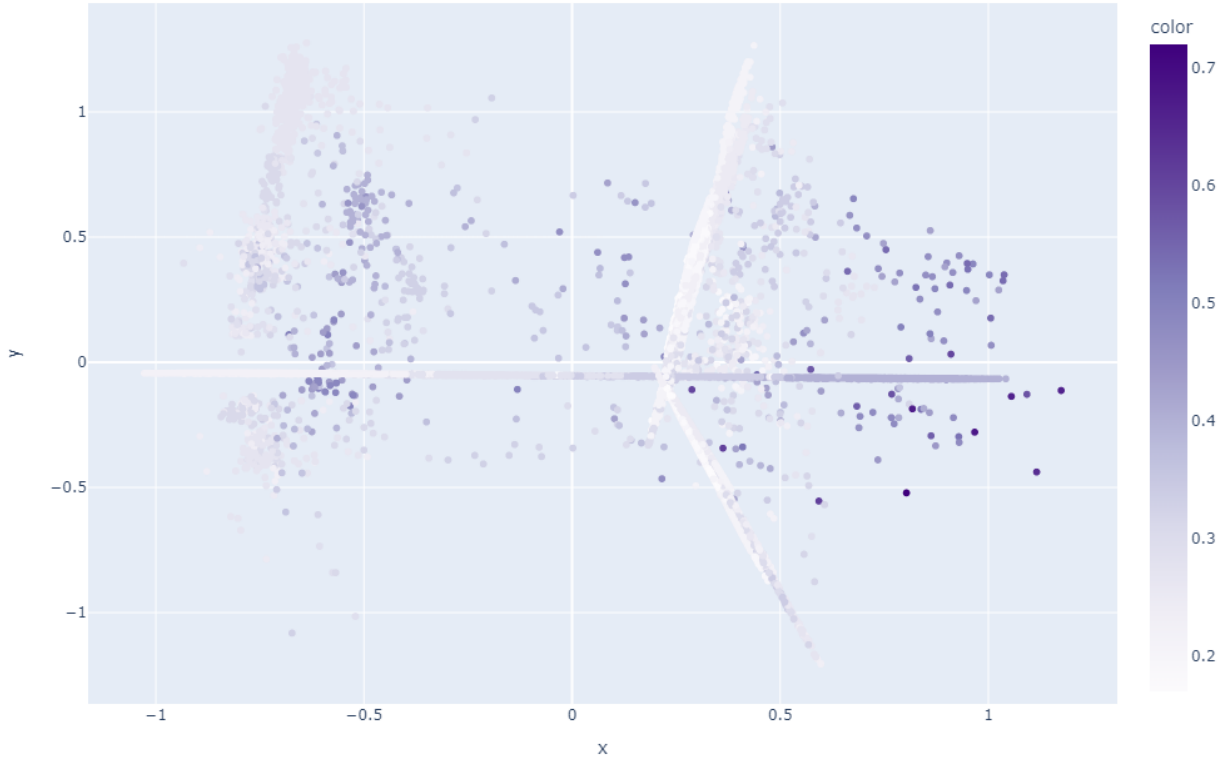
Engagement trend analysis was executed on the channels we collected data for, to discover channels exhibiting suspicious behaviors. There are 5 key steps employed by the script in discovering suspicious behaviors in a channel:

- Rolling window correlation analysis
- Anomaly detection
- Rule-based classification
- Principal component analysis
- Clustering

Metrics:

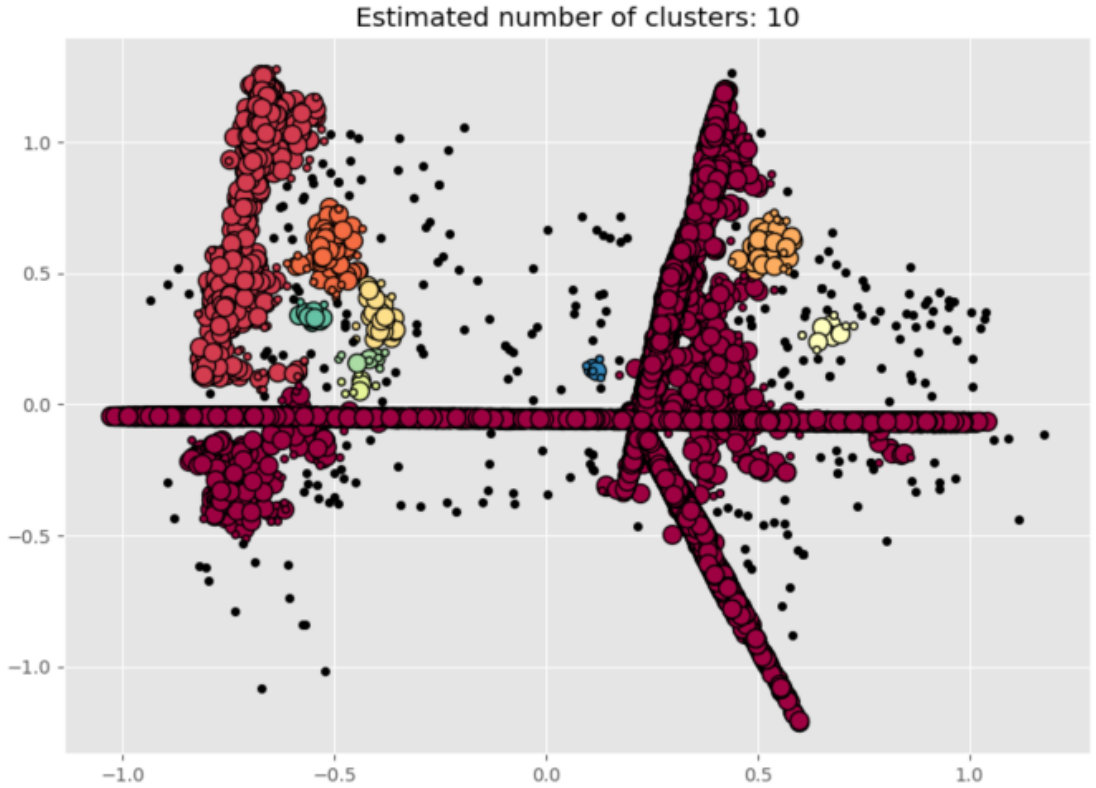
- Views & Subscribers
- Views & Videos
- Views & Comments
- Subscribers & Comments
- Videos & Comments
- Subscribers & Videos

Engagement Analysis



Principal component analysis scatter plot showing suspicion scores

Engagement Analysis



Scatter plot showing clusters identified from DBSCAN.

Engagement Analysis

CHANNELS WITH MOST SUSPICIOUS DATA POINT

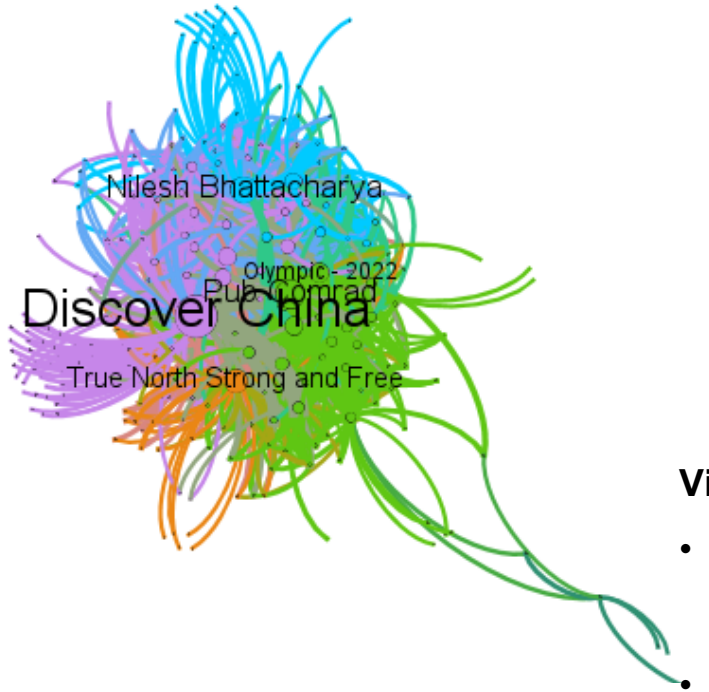
Channel ID	Channel Name	Suspicious Score	Date Range
UCN_qIhm7BAq9Qxa6j9XMVXA	Breaking News TV	0.72	2019-03-24 to 2019-10-08
UCrGZO3wJ20CWiy36Fdu6vdw	Badminton Talk	0.68	2019-04-24 to 2019-11-08
UCkQSMH1vP1Kcx-SPArUYLLQ	viral_makkodak	0.67	2020-02-07 to 2020-09-04
UCM2XWLz9zUcYZRJNl2dsjnw	DCTV Heroes	0.66	2019-05-23 to 2019-12-18
UCmvtaFkiWOSHhnOwt4Fz68g	SAFA News	0.64	2019-11-26 to 2020-06-22

Network Analysis

The YouTube co-commenter network represents the connections between commenters on YouTube videos, where the edge weight indicates the number of videos they commented on together

This also showed the top node in each community, and the type of content the community engaged with. The modularity measure was used to partition the network into clusters or communities. As the degree centrality measure was used to find out the top nodes, it also shows how many ties/edges a node has

Network Analysis



CHINESE NARRATIVE COMMUNITIES

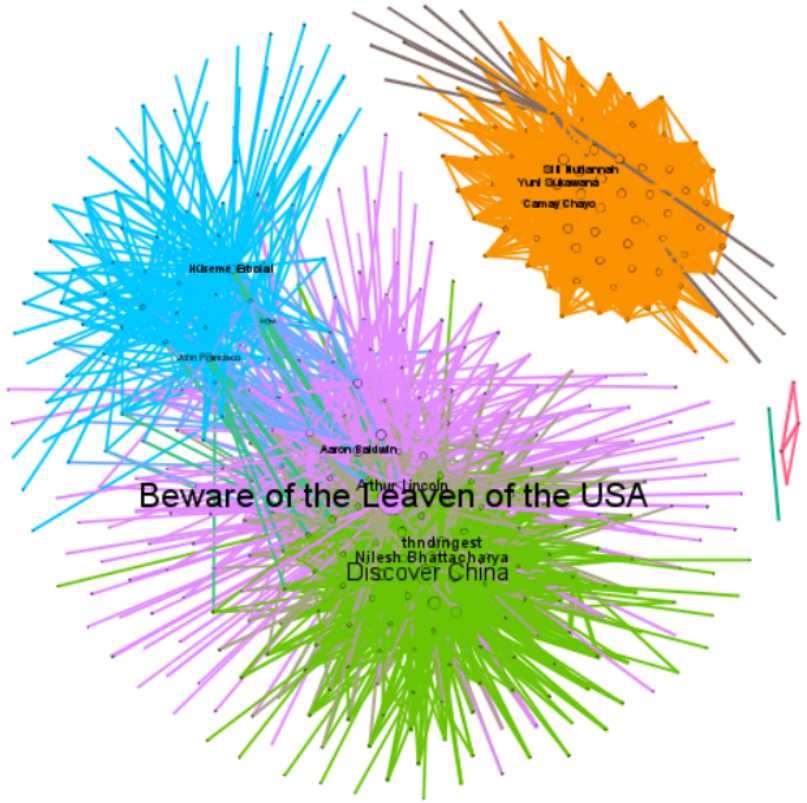
S/N	Top Commenter Name	Community	Degree
1	Nilesh Bhattacharya	0 (Blue)	108
2	thndrngest	0 (Blue)	72
3	Beware of the Leaven of the USA	0 (Blue)	66
4	Discover China 探索中	1 (Purple)	173
5	True North Strong and Free	1 (Purple)	95
6	Colchicum autumn crocus	1 (Purple)	66
7	Pub Comrad	4 (Green)	113
8	Olympic - 2022	4 (Green)	76
9	Last Chang	4 (Green)	72

Videos:

- “Xi Jinping: China’s president and his quest for world power”
- “Global brands face backlash in China for rejecting Xinjiang cotton”.
- “Taiwan: ’China preparing for final military assault”,

Comments:

- “CPC is indeed the Chinese people’s party
- “Of the 14 countries bordering China, it has conflict with 13 of them including Russia”,

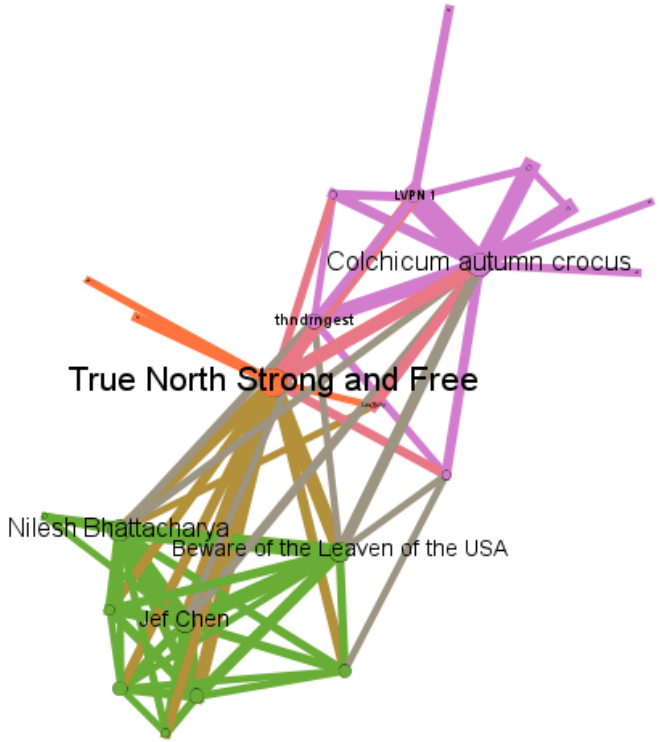


UYGHUR NARRATIVE COMMUNITIES

S/N	Top Commenter Name	Community	Degree
1	Discover China	0 (Green)	145
2	Nilesh Bhattacharya	0 (Green)	102
3	thndrngest	0 (Green)	95
4	Beware of the Leaven of the USA	1 (Purple)	210
5	Arthur Lincoln	1 (Purple)	85
6	Aaron Baldwin	1 (Purple)	63
7	Hüseme Erbolat	2 (Blue)	67
8	John Francisco	2 (Blue)	55
9	lin hai	2 (Blue)	35
10	Yuni Sukawana	3 (Orange)	71
11	Siti Nurjannah	3 (Orange)	71
12	Camay Chayo	3 (Orange)	68

Videos:

- “Uyghurs Who Fled China Now Face Repression in Pakistan
- Xinjiang Became Muslim ft. Let’s Talk Religion”,
- “China’s Vanishing Muslims: Undercover In The Most Dystopian Place In The World”.



COVID-19 NARRATIVE COMMUNITIES

S/N	Commenter Name	Community	Degree
1	True North Strong and Free	0 (Orange)	16
2	Lau Billy	0 (Orange)	3
3	Frederic Chen	0 (Orange)	1
4	Nilesh Bhattacharya	1 (Green)	12
5	Jef Chen	1 (Green)	11
6	Beware of the Leaven of the USA	1 (Green)	11
7	Colchicum autumn crocus	2 (Purple)	13
8	thndrngest	2 (Purple)	8
9	LVPN 1	2 (Purple)	6

Future directions

- Considering bot accounts and their impact
- Automating the process of connecting community detection and clustering to engagement trends
- Community level suspicion score

Thank you!

COSMOS Tools Developed:

- **COVID-19 -** https://cosmos.ualr.edu/covid-19
- **Blogtrackers** - https://btracker.host.ualr.edu
- **YouTubeTracker** - https://vtracker.host.ualr.edu
- **Focal Structure Analysis** - http://fsa.host.ualr.edu

Blogtrackers YouTubeTracker

FOCAL STRUCTURE ANALYSIS

cosmos.ualr.edu

Nitin Agarwal – nxagarwal@ualr.edu

Follow **cosmographers** on

ACKNOWLEDGEMENTS

This research is funded in part by the U.S. National Science Foundation (OIA-1946391, OIA-1920920, IIS-1636933, ACI-1429160, and IIS-1110868), U.S. Office of Naval Research (N00014-10-1-0091, N00014-14-1-0489, N00014-15-P-1187, N00014-16-1-2016, N00014-16-1-2412, N00014-17-1-2675, N00014-17-1-2605, N68335-19-C-0359, N00014-19-1-2336, N68335-20-C-0540, N00014-21-1-2121, N00014-21-1-2765, N00014-22-1-2318), U.S. Air Force Research (FA9550-22-1-0332), U.S. Army Research Office (W911NF-20-1-0262, W911NF-16-1-0189), U.S. Defense Advanced Research Projects Agency (W31P4Q-17-C-0059), Arkansas Research Alliance, the Jerry L. Maulden/Entergy Endowment at the University of Arkansas at Little Rock, and the Australian Department of Defense Strategic Policy Grants Program (SPGP) (award number: 2020-106-094). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations. The researchers gratefully acknowledge the support.

COSMOS Collaboratorium for Social Media and Online Behavioral Studies

National Science Foundation · Office of Naval Research · Air Force Research Laboratory · Defense Advanced Research Projects Agency · Army Research Office · Department of Homeland Security · Arkansas Research Alliance