

Social Cybersecurity and Network Analysis of Pakistani Influence Operations Targeting India

Abstract

Since its inception, Pakistan has interfered in India's internal affairs, a practice that has evolved with the rise of social media. This study examines how Pakistan has adapted to digital platforms to execute propaganda and influence operations aimed at polarizing and destabilizing Indian society. By employing social network analysis and social cybersecurity, this research seeks to understand the mechanisms and impact of these digital influence campaigns, providing insights into their structure, reach, and the flow of information.

Introduction

Nation states are increasingly leveraging social media as a strategic instrument in contemporary warfare, utilizing both state and non-state actors to pursue political goals (Flew & Iosifidis, 2020). The integration of social media into hybrid warfare strategies presents a sophisticated and dynamic challenge in unconventional conflict approaches (Svetoka, 2016). Social media is exploited to propagate propaganda by leveraging existing narratives and employing inauthentic accounts to enhance message reach (Prior, 2020). Hostile actors and opposing nation states infiltrate the civil societies of their targets to intensify dissent and widen societal divisions by exploiting vulnerabilities in socio-political and economic fault lines (Jayamaha, 2018) to further polarize and divide society. Employing social media for hybrid warfare, particularly through the dissemination of disinformation, is a key tactic in these strategies (Rasheed, 2021).

In an era where information warfare has become a potent tool in international relations, the historical conflict between India and Pakistan has evolved into a new digital battleground. Since the inception of the two nations India (a secular democracy with a Hindu majority) and Pakistan (Islamic Republic), conflict between the nations has been ripe, which initially began due to territorial disputes in the state of Jammu and Kashmir. However, territorial disputes are not the only points of contention between the two nations - the states are very different ideologically. Pakistan's persistent conflict with India transcends territorial disputes, rooted deeply in ideological opposition (Fair, 2014). The existence of a secular, diverse India challenges the ideological foundations upon which Pakistan was established, driving its revisionist stance not merely for land, but as part of a broader civilizational contest. Pakistan's founding father, Muhammed Ali Jinnah, had suggested the partition of British India based on "Two Nation Theory" - the idea that Hindus and Muslims could not live together in harmony, and that the only way for Muslims to be free would be by having an Islamic state known as Pakistan (Majid,

Hamid, & Habib, 2014). A neighboring India, contradicts this view - that a nation with Hindus, Sikhs, Muslims, Jains, Parsis and other religious minorities live together. Such ideological divergence has led to repeated attempts by the Pakistani state to interfere with domestic issues within India. Issues pertaining to border skirmishes, proxy warfare (Kanwal, 1990), cross border terrorism (Harshe, 2003) and polarizing political rhetoric (Akhtar, 2009) from the two nations have exacerbated tensions between the two countries over time.

Literature Review

Digital Operations

With the rise of social media, the digital battleground between India and Pakistan has also evolved beyond traditional state-controlled media outlets, marking a new phase characterized by direct confrontations through trolls, spam and bot accounts. This digital activity has become a component of broader strategies to exploit the tensions and extend the narrative conflict into the digital sphere, wherein specific incidents of heightened propaganda can be linked back to key moments of escalating tension between the two nations, where social media was used to amplify the impact of physical geopolitical confrontations (Rasheed, Naseer, & Khawaja, 2021). This transformation has expanded the scope of digital conflict, encompassing both sophisticated cyberattacks and the strategic manipulation of information to influence public perception and exploit national vulnerabilities.

In the aftermath of the 2019 Pulwama attack, there was a marked increase in digital propaganda and misinformation, leveraging social media platforms by both India and Pakistan (Hussain, Shahzad, & Saud, 2021). Throughout these episodes, the strategic manipulation of information and the adept use of digital tools have been pivotal in molding both domestic and international perceptions. The consistent exploitation of social media during conflicts, exemplified by the 2019 standoff between India and Pakistan following the Balakot airstrikes, underscores the profound influence of cyber operations on tangible geopolitical conflicts. These digital skirmishes not only intensify existing tensions but also present substantial challenges to regional stability and security, as both countries persist in a high-stakes contest of narrative control and influence within the cyber domain.

Influence Operations

Influence operations involve orchestrated efforts by state or non-state actors aimed at molding the perceptions, beliefs, or behaviors of specific audiences to meet strategic objectives (Bergh, 2020). These methods, which include the use of propaganda, misinformation, and disinformation, are crafted to exploit vulnerabilities within informational environments and sway public opinion or decision-making processes.

Propaganda uses biased or misleading information, often with emotional appeals, to support specific political causes or viewpoints. It strategically presents information to shift audience opinions, omitting contradictory data (Bernays, 2005). Misinformation, on the other hand, involves unintentionally false or inaccurate information due to errors in understanding or reporting (Suarez-Lledo, V., & Alvarez-Galvez, J, 2021). When used in influence operations, misinformation is exploited by malicious actors to create confusion and erode public trust. Disinformation, however, is intentionally crafted and disseminated with the aim to deceive. It involves creating fictitious narratives or manipulating facts to mislead the public, capitalizing on biases and preconceptions to foster divisiveness or discord (Fallis, 2015).

The role of social networks in influence operations provide a scalable and covert means to transmit targeted information swiftly (Chen, W., Castillo, C., & Lakshmanan, L. V, 2022). Social media platforms allow operators to utilize detailed demographic data to tailor content specifically to resonate with certain populations, enhancing the effectiveness of their campaigns.

Importance of Social Network Analysis

Social Network Analysis (SNA) is a methodology used to analyze patterns and interactions across social networks, which is particularly important within social media analysis (Tabassum, Pereira, Fernandes, & Gama, 2018). SNA can be vital for identifying influence operations and analyzing the flow of information. This will also aid in better understanding the key nodes, links and centrality, to understand the key actors behind disseminating content.

SNA maps and measures relationships and information flows among individuals, groups, organizations, and other connected entities like computers and URLs. Through this analysis, insights are garnered that illustrate how influence spreads across a network and highlight the most pivotal nodes. These central nodes are crucial for the broad dissemination of information and are often the focus of targeted campaigns aimed to maximize influence. Understanding these network dynamics enables analysts to track information spread and develop methods to counteract harmful influence efforts effectively.

SNA also assesses networks' structural features like density, centrality, and connectivity, influencing their resilience to external forces. Networks with centralized structures may be vulnerable to disruptions if crucial nodes fail, while decentralized networks, though more robust, might spread messages less effectively. These insights are key to crafting strategies that either bolster a network's defenses against detrimental information or improve the spread of beneficial content.

In the field of cybersecurity, SNA is utilized to identify and understand the dynamics of network relationships and communications (Vlachos, Stamatiou, Tzamalīs, Nikolettēs, & Chantzi, 2019). It helps cybersecurity professionals detect patterns and anomalies that indicate potential

threats, such as coordinated attacks or breaches. By analyzing the flow of information and the connections between nodes, SNA enables the identification of key points where security measures can be intensified to prevent malicious activities and protect network integrity.

Social Cybersecurity

Social cybersecurity is an applied computational social science with two objectives (National Academies of Sciences et al., 2019):

- “characterize, understand, and forecast cyber-mediated changes in human behavior and in social, cultural, and political outcomes; and
- build a social cyber infrastructure that will allow the essential character of a society to persist in a cyber-mediated information environment that is characterized by changing conditions, actual or imminent social cyberthreats, and cyber-mediated threats.”

Social cybersecurity merges principles from cybersecurity, social sciences, and computational disciplines, examining social dynamics within digital contexts. This field strives to assess and counteract how online influence and misinformation impact human behavior, societal norms, and political landscapes. By leveraging data science, network analysis, and machine learning, it explores the diffusion of information through social networks and its effects on society. Specialists in this area are dedicated to preserving public discourse and upholding the integrity of information in our digitally evolving world.

BEND Maneuver

The BEND framework in social cybersecurity refers to a comprehensive approach for understanding and countering influence campaigns in digital environments (Carley, 2020). This framework conceptualizes influence operations as a series of strategic maneuvers—both narrative and structural—that actors use to alter community interactions and the flow of information within these communities.

BEND stands for four types of maneuvers aimed at manipulating both the narrative (what is being discussed) and the social network (who is interacting with whom). The "B" and "E" maneuvers (Build, Bridge, Engage, and Enhance) are positive, aiming to create or strengthen connections and clarify or boost favorable narratives. The "N" and "D" maneuvers (Neglect, Neutralize, Dismay, and Distort) are negative, intended to break down community ties or muddy understanding.

These maneuvers can involve a variety of tactics such as deploying bots to amplify specific messages, using trolls to attack opponents or sway public opinion, or creating fake accounts to simulate grassroots support (astroturfing). The overarching goal is to shift the perception and

behavior of the targeted audience in a way that serves the strategic objectives of the campaign initiator, be it a state actor, a political group, or any other entity with vested interests in the outcome of the influence operation.

The BEND framework is particularly useful in analyzing complex information operations where multiple tactics are synchronized to create, manipulate, or disrupt social narratives across various online platforms. By breaking down the components of these campaigns, researchers and cybersecurity professionals can better predict, identify, and mitigate harmful influence operations aimed at misleading the public, skewing political discussions, or causing societal discord.

Online Twittersphere between India and Pakistan

The study by Hussain et al. on the state of digital information warfare between India and Pakistan highlights a pattern of interaction between users from these two nations on Twitter. Particularly in the context of high-tension incidents like the Pulwama attack and subsequent military actions, the online behavior of users from both countries largely involved engaging with content within their own echo chambers. This segregated form of engagement suggests that social media platforms, rather than being spaces for cross-border dialogue and reconciliation, are instead arenas where pre-existing hyper-nationalistic sentiments are reinforced. Indian and Pakistani users predominantly posted and interacted within separate hashtags, which were heavily loaded with nationalistic and propagandistic content, echoing and amplifying the official narratives and sentiments prevalent in their respective countries.

This insular interaction pattern creates what are effectively digital echo chambers, where users are exposed primarily to viewpoints that reinforce their own beliefs and biases. The study found little evidence of substantive engagement between Indian and Pakistani users in terms of dialogue or debate. Instead, each group stayed largely within its respective echo chamber, amplifying its own narratives without significant exposure to or interaction with the opposing viewpoints. This lack of cross-engagement can perpetuate misunderstandings and increase the ideological distance between the two populations, reducing the potential for social media to act as a bridge-building tool between India and Pakistan.

Furthermore, the role of trolls and bots, as highlighted in the study, complicates the nature of online interactions between these two communities. Many of these accounts, often indistinguishable from real users, are operated to push specific political or nationalistic agendas, creating a skewed perception of the general sentiment. This manipulation of social media discourse serves to deepen the divide, fueling a continuous cycle of misinformation and mistrust. The strategic use of digital propaganda not only muddies the waters of genuine discourse but also plays a significant role in shaping the political landscape, influencing public opinion and national sentiment during critical moments of conflict.

Polarization between India and Pakistan

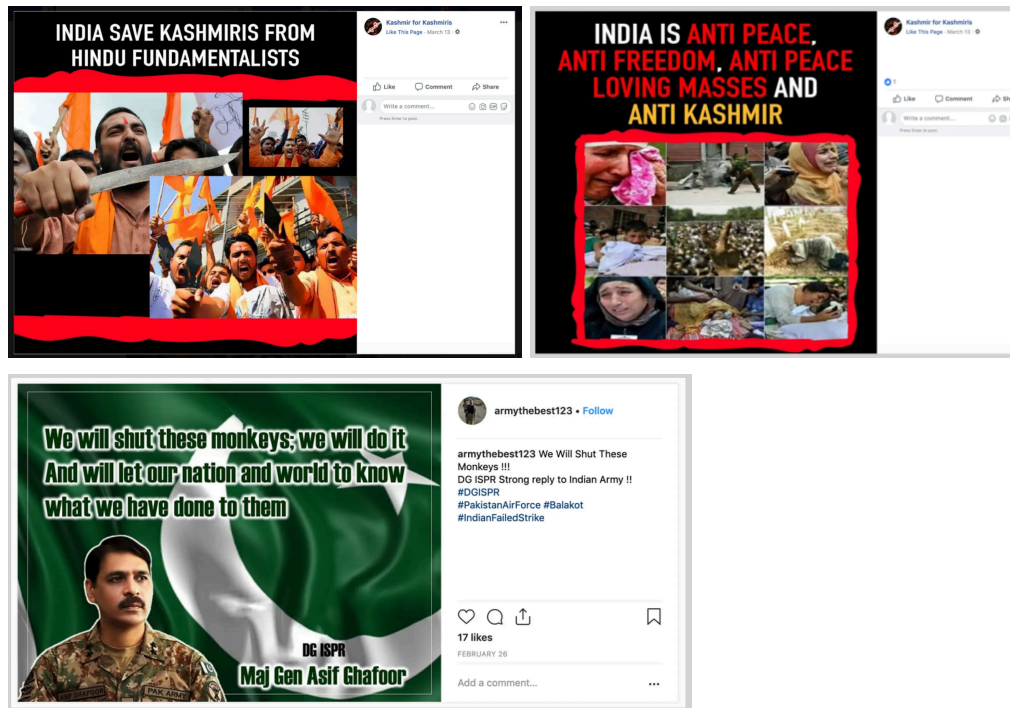
The study "Analyzing the State of Digital Information Warfare Between India and Pakistan on Twitter" highlights significant findings about social media interactions between Indian and Pakistani users. It was observed that there is considerable polarization driven by targeted hashtag usage, which reinforces nationalistic and political leanings rather than promoting cross-cultural dialogue. Each group tends to stick within their digital echo chambers, using specific sets of hashtags that align with their national perspectives, thereby amplifying existing divisions and reducing the chances of constructive interactions.

Further analysis revealed that political affiliations significantly influence online behaviors, particularly among Indian users. The study found that members of India's ruling party were more likely to engage in using polarizing and aggressive hashtags compared to other political parties. This strategic use of social media to promote specific political narratives suggests a deliberate approach to escalate tensions and manipulate public perception, which is a common tactic observed in digital information warfare.

The methodology used in the study, involving label propagation techniques based on hashtag co-occurrences, provides a novel approach to understanding how polarization manifests on social media. These methods offer insights into the dynamics of organized social media strategies and their powerful impact on public discourse. The findings highlight the challenges faced in combating misinformation and polarization, emphasizing the need for strategies to mitigate the influence of divisive content on digital platforms, particularly in politically sensitive contexts.

Past Pakistani Interference

In 2020, Facebook announced the removal of a substantial network of deceptive accounts operated from Pakistan, promoting the Pakistani military. This crackdown involved eliminating 453 accounts, 103 pages, 78 groups, and 107 Instagram accounts.



Based on an analysis (Nimmo, B., & Karan, K, 2019) by Atlantic Council’s Digital Forensics Lab (DFRLab), The images display examples of inflammatory and provocative social media content with messages exacerbating tensions between India and Pakistan over Kashmir, using animalistic language to describe Indians (“Monkeys”) and threatening Pakistani military actions. More specifically, Hindus are depicted as “fundamentalists” who are threatening to kill Kashmiris.

Taxonomy of Pakistani Interference

Based on the literature review, a taxonomy of Pakistani interference was developed to understand Pakistani interference targeting India.

Table 2: Themes of Pakistani Interference with Twitter examples

Theme	Description	Example
Furthering Communal Divides	Exacerbating religious and sectarian tensions within India	#Islamophobia_In_India
Balkanization of India	Encouraging separatist sentiments and movements across different regions.	#FreeDravidistan

Hindus as a perceived threat	Hindu political participation and activism is received as a threat	#StopIndianAggression
Exaggerating Human Rights violations in Kashmir	Promoting a narrative of harsh governance and human rights abuses in Kashmir	#KashmirSolidarityDay
Diplomatic Isolation	Isolating India diplomatically and tarnishing its image with allies and neighbors.	#IndiaOut
Cyber and Information Warfare	Engaging in cyber attacks and spreading disinformation to disrupt India's state security and society.	#DamnocracyDay
Economic Undermining	Implementing strategies to weaken India's economic position regionally and globally.	#BoycottIndianProducts & variants in Bangla, Urdu and Arabics
Portraying India as the Aggressor and Scapegoating for Pakistan's Challenges	India as the principal instigator of regional conflicts and anti-Pakistan activities, positioning India as a convenient scapegoat for Pakistan's internal challenges.	#IndianSurgicalDrama
Downplaying Indian security apparatus	Denying or downplaying Indian security accomplishments - such as surgical strikes or paramilitary operations - to shield Pakistan from criticism	

Methodology

The methodology leverages tools for comprehensive analysis of social networks. Using Social Network Analysis with ORA and Netmapper, we assess network structures and key influencers. The Bothunter algorithm helps us identify and classify bots on social media. We further enhance our approach with the BEND maneuver in Social Cybersecurity to understand and influence online interactions, supported by data annotation through large language models like GPT-4 for precise profiling.

SNA

In our methodology, we employ SNA using advanced tools like ORA (Altman, N., Carley, K. M., & Reminga, J, 2020) and Netmapper (Carley, K. M., & Reminga, J, 2018), which are integral for analyzing and understanding social networks within cyber-mediated environments. ORA, developed by CASOS at Carnegie Mellon, is a dynamic meta-network assessment tool that provides comprehensive analytics on social network structures over time and space, and it is adept at identifying key influencers and nodes within networks, assessing changes, and evaluating network resilience against disruptions.

Netmapper, compatible and interoperable with ORA, enhances our methodology by facilitating text mining to support network analysis. It can process text in multiple languages and formats, extracting semantic and meta-networks, which are crucial for understanding the underlying structures and flows within data. This tool also supports the extraction of network data from diverse text sources, allowing for detailed sentiment analysis and the identification of cues within texts.

BotHunter Algorithm

Bothunter (Beskow & Carley, 2018) is an algorithm designed to detect and characterize automated activities on social media platforms like Twitter. It analyzes behavioral patterns and network interactions to identify characteristics typical of bots. In our study, we employ a threshold where profiles scoring above 0.7 on Bothunter's scale are classified as bots, allowing for a precise identification of automated accounts while reducing the likelihood of false positives.

Social Cybersecurity

BEND maneuver in social cybersecurity is designed to analyze influence operations. The BEND framework identifies 16 specific communication objectives, divided equally into shaping the social networks (community) and shaping the narrative (content). Each objective is associated with either building or negating aspects of social interactions and discussions. BEND alternates both the community structure (who communicates with whom) and the narrative (what is being discussed). Unlike traditional models that focus on negative impacts like disinformation (often described using tactics like distract, distort, dismay, and disrupt), BEND includes both negative and positive maneuvers, enabling a more nuanced approach to understanding and influencing social networks.

LLMs

We integrate these tools with large language models (LLMs) like GPT-4 for data annotation (Tan et al., 2024), which involves feeding Twitter profile characteristics into the model to classify accounts as Pakistani, news agencies, government accounts, or reporters.

Research Questions

RQ1: How can social cybersecurity frameworks be applied to dissect and understand Pakistani influence operations on Indian social media platforms?

RQ2: What role does coordinated inauthentic activity play in the orchestration of Pakistani influence operations targeting India, and how does it impact the efficacy of these campaigns?

RQ3: Do Pakistani influence operations primarily aim to trend new topics on social media, or do they hijack existing popular topics to inject their narratives?

RQ4: Is there evidence that Pakistani troll operations explicitly support or amplify specific political parties in India to manipulate political discourse?

Hypothesis

The hypothesis for analyzing the objectives and tactics of Pakistani influence operations in relation to India can be articulated as follows:

1. Pakistani influence operations are hypothesized to focus on amplifying religious and ethnic divisions by using distorted narratives to spotlight incidents of religious conflict, particularly highlighting or exaggerating violent clashes to emphasize alleged systematic bias against Muslims by Hindu majorities.
2. These operations are likely to support separatist movements in regions with distinct ethnic identities and historical grievances, such as Kashmir—emphasizing local independence or Pakistani solidarity narratives and alleged human rights abuses by India; Punjab—revitalizing sentiments of Sikh separatism; Northeast India—focusing on ethnic divisions and stories of neglect; and South India—propagating Dravidian nationalist sentiments.
3. These operations aim to undermine India's image as a stable and progressive region by emphasizing incidents of religious violence, poor sanitary conditions, and gender-based violence, thereby preventing India from consolidating its status as a regional power.
4. Strategies may include stoking polarization between major Indian political parties, notably the Bharatiya Janata Party (BJP) and the Indian National Congress, leveraging extremist narratives, and misrepresenting policies or political statements to fuel partisan disagreements.
5. The operations are expected to utilize social media extensively, deploying automated bots and inauthentic accounts to amplify these divisive narratives and exacerbate political and social polarization within India, ultimately weakening India both internally and on the international stage.

These hypotheses suggest that Pakistani influence operations are strategically designed to exploit and widen existing fissures within Indian society, aiming to weaken India both internally and on the international stage.

Hashtags

These hashtags were chosen based on examining day to day trends on Pakistani Twitter, and collected using the latest Twitter API.

Table 1: Hashtags trending on Pakistani Twitter collected

Hashtag	Date Collected	n
DamnocracyDay	January 26, 2022	1347
KashmirSolidarityDay	January 1 - April 1, 2024	956
BabriMasjidDemolition	January 1 - April 1, 2024	742
BoycottIndianProducts	January 1 - April 1, 2024	1505
ByeByeModi	January 1 - April 1, 2024	1958
IndiaOut	January 1 - April 1, 2024	1762

Results

The dynamics of Pakistani influence operations targeting India have undergone notable shifts from 2022 to the present. Previously, key influencers affiliated with the PTI (Pakistan Tehreek-e-Insaf) were observed deploying unique hashtags during significant Indian holidays and historical events, effectively utilizing these occasions to maximize the impact of their campaigns. However, following the PTI's departure from power, there has been a transition in influence tactics. Currently, these operations more frequently employ hashtag hijacking techniques, where existing trends critical of India's human rights record or economic practices are co-opted. This shift suggests a strategic adaptation to more effectively embed disinformation within broader, already trending topics, thereby amplifying their reach and potential influence on public discourse concerning India.

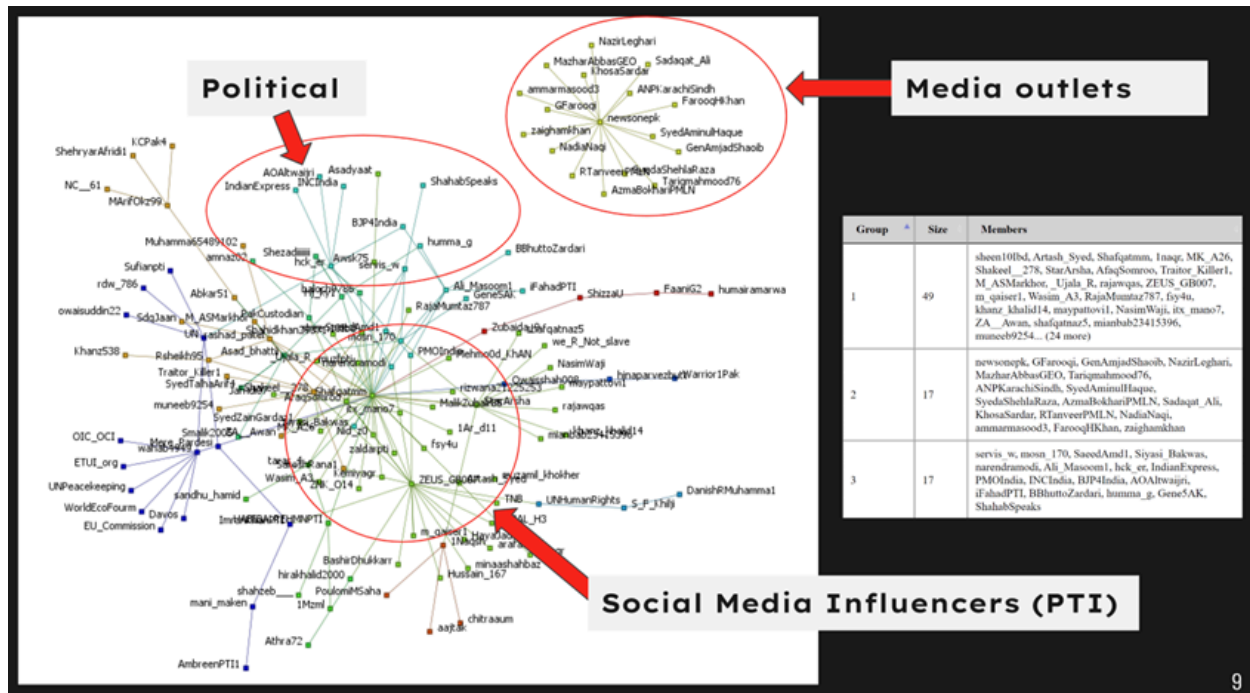
#DamnocracyDay

On January 26, 2022 on India's Republic Day, self-identified Pakistani Twitter users and self-described social media activists used the hashtag #DamnocracyDay to disparage Indian democracy and how it treats minorities.

Figure 1: Self Identified Pakistani accounts



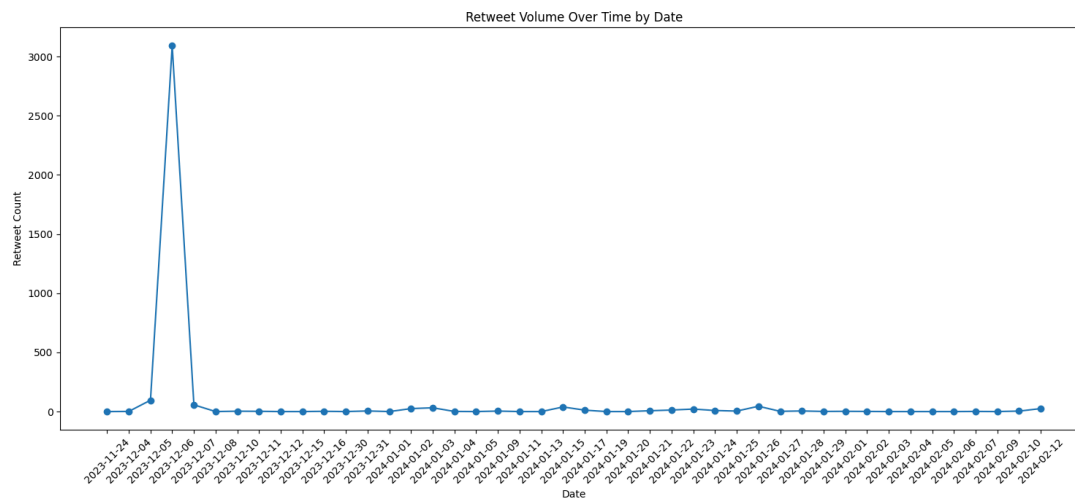
Figure 2: Social Network for #DamnocracyDay and grouped by Louvain



The most influential actors, according to the #DemocracyDay network's analysis, are Pakistani influencer and nationalist accounts. Three main clusters of the network each have a predominance of Pakistani media outlets, influencers, and political opinion. The community detection pinpoints three important groups: media outlets, PTI social media influencers, and internet users who make political party and media references in India. Numerous communities and clusters are present within the network, which suggests that a large number of participants are working together to promote the hashtag and advance their own stories by consistently replying to one another. The network's prominence of PTI-affiliated influencers and Pakistani nationalism accounts is another clue that political players were participating in the influence effort.

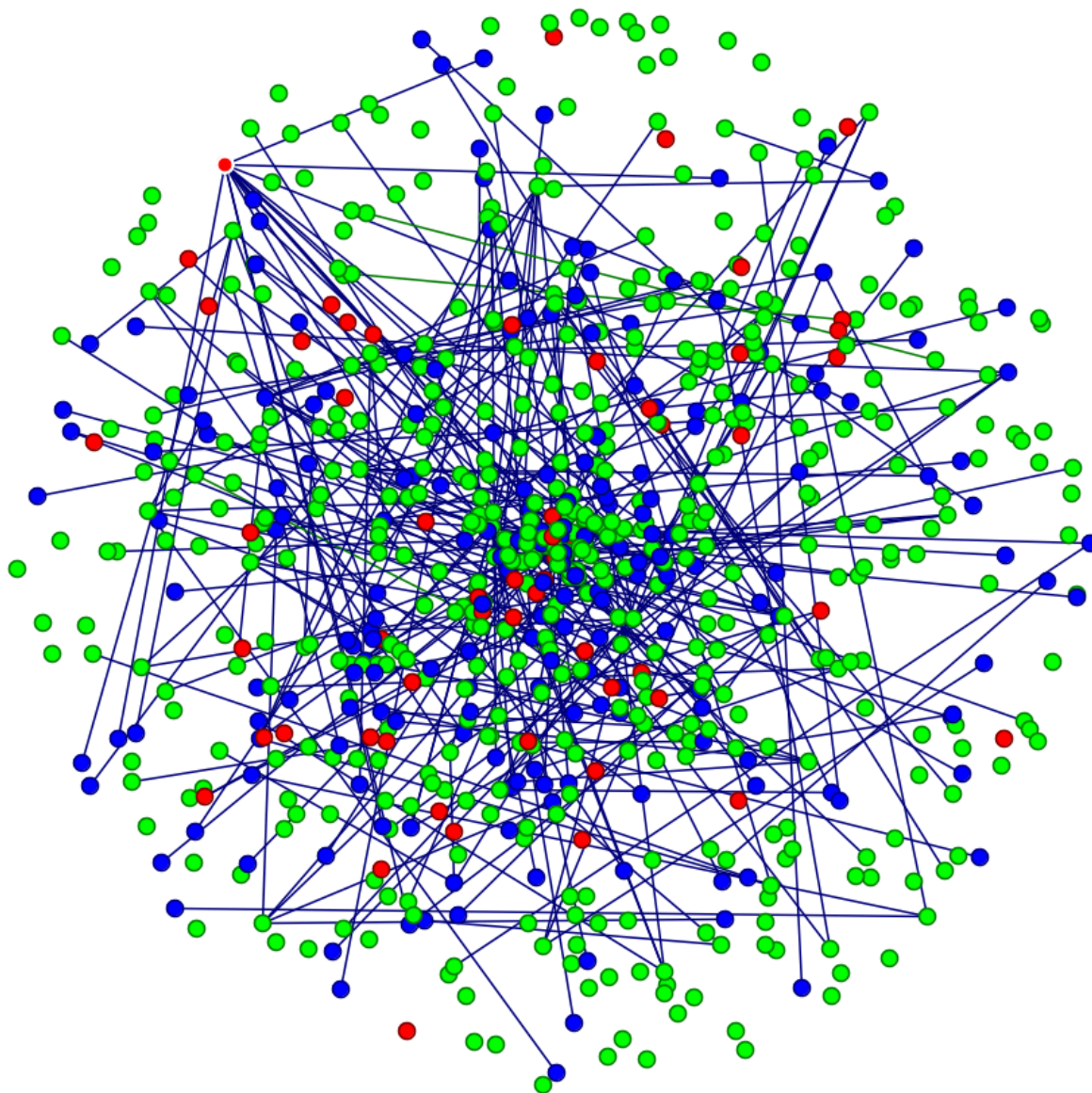
#BabriMasjidDemolition

Figure 3: Timeseries analysis of #BabriMasjidDemolition



The graph illustrates a spike in retweet activity on 6th December, the day which Babri Masjid was demolished in 1992. This pattern could suggest coordinated or inauthentic social media activity, potentially aimed at amplifying the visibility of the topic.

Figure 4: Network of #BabriMasjidDemolition colored by Pakistan Classification



The red nodes indicate Pakistan accounts, green as not Pakistani and blue indicates no value. Given the presence of red nodes, this indicates Pakistani presence in the network.

Figure 5: Pakistani actors in #BabriMasjidDemolition



Figure 6: Examples of “copypasta” campaigns

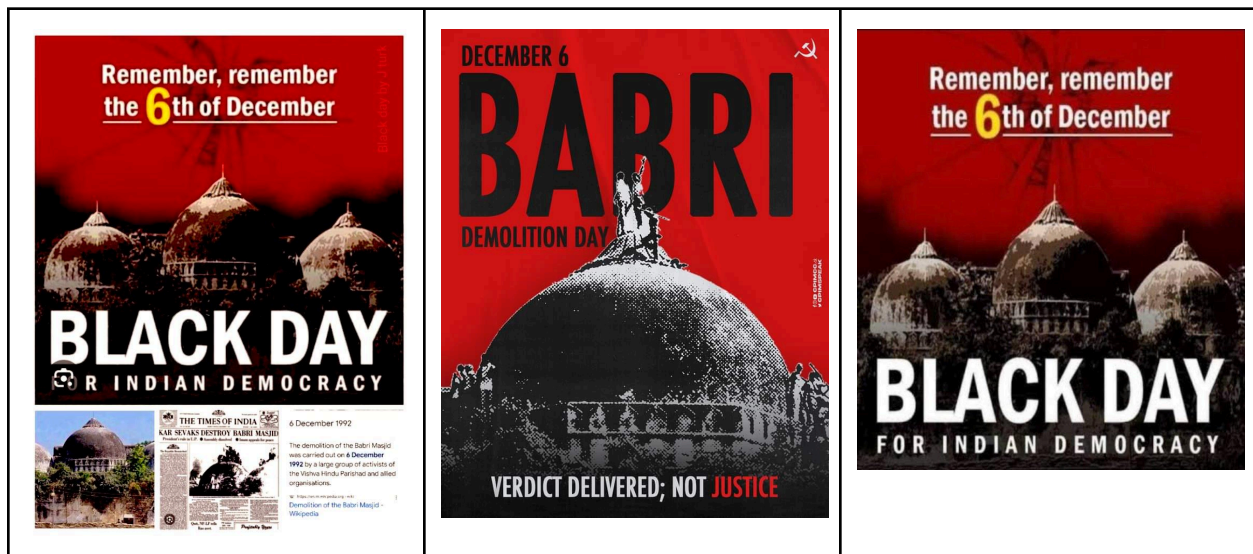
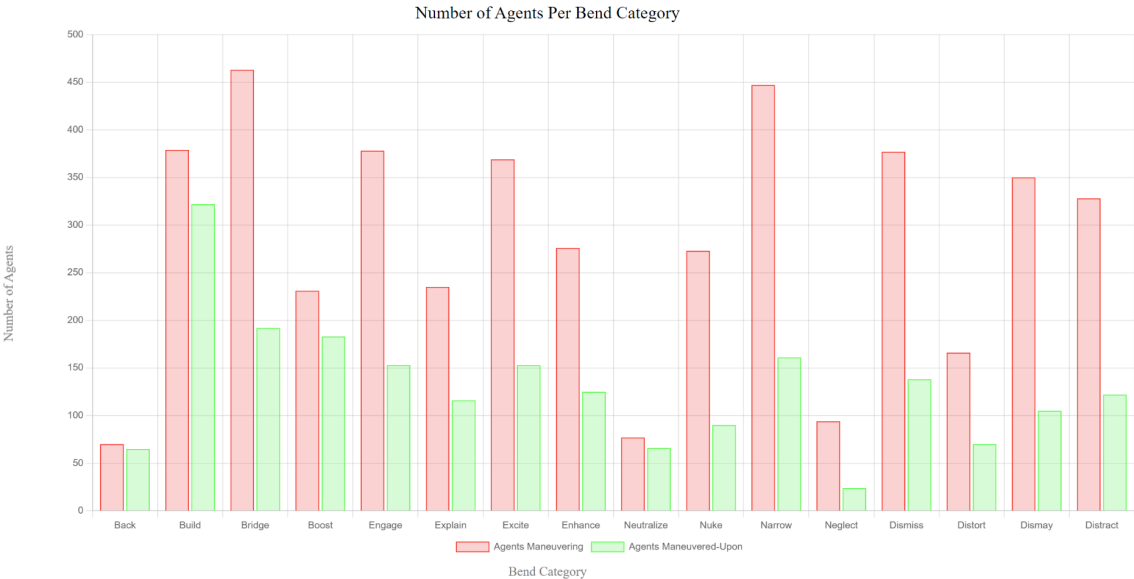


Figure 7: BEND Agent Maneuvers

BEND Agent Maneuvers

The chart below shows the number of agents in each BEND category that maneuvered or were maneuvered-upon. An agent is computed to be in a BEND category by averaging a collection of metrics to produce a likelihood and strength (or confidence) of category membership. An agent with a likelihood value greater than the Mean is counted as a member.



Number of agents	Number of agents maneuvering in at least one BEND category	Number of agents maneuvered-upon in at least one BEND category	Sum of agent maneuverers across all BEND categories	Sum of agents maneuvered-upon across all BEND categories
772	554 (71.76% of total agents)	511 (66.19% of total agents)	4,513	2,085

The high activity levels in the "Narrow," "Bridge," and "Distract" categories from the BEND Agent Maneuvers chart linked to the #BabriMasjidDemolition indicates a strategic use of social media to influence public discourse by some agents speaking within narrow groups, other attempting to bridge with different groups and those seeking to distract from the subject. “Narrow” suggests efforts to focus the conversation among more tightly knit groups, potentially to reinforce specific narratives within communities more likely to amplify them. “Bridge” shows a high level of activity in this category implies attempts to connect different groups or communities, possibly to spread the narrative across diverse social segments and increase the reach of the messaging. “Distract” shows the prominence of this category and points to tactics aimed at diverting attention from potentially damaging or controversial aspects of the demolition issue by introducing unrelated or less impactful topics.

Figure 8: BEND Community Assessment - Bot Activitys

COMMU	Degree	Density	E/I Index	Echo	Echo	Reciprocit	Specialty	Total	Total	Total	Total Url
-------	--------	---------	-----------	------	------	------------	-----------	-------	-------	-------	-----------

NITY	Centrality			Chamber Index 1	Chamber Index 2	y	Concepts	Agent	Hashtag	Tweet	
ISBOT=0	0.020627	0.952018	-0.49106	0.983743	1	1	NellieM assacre, PolicyPa ralysis, CoalSca m, JayantiN atarajan, Muslim	493	550	632	426
No Value	0	0	1	0	0	1		198	0	0	0
1	0.042335	1	0.844796	1	0	1	Reclaim Bharat, backday, sajaiku mar, in_sha_a llah, Naseem	81	129	110	66

Entities classified as bots demonstrate integration and influence, with high degree centrality indicating their central roles within the network. The network's high density points to a closely interconnected community where nodes frequently interact. The External/Internal Index values suggest varying degrees of external versus internal interactions among these entities, with some showing a greater propensity for external engagements. The presence of strong echo chamber indices highlights the prevalence of echo chambers, where similar or reinforcing content circulates predominantly, limiting exposure to diverse perspectives. High reciprocity values confirm that these interactions are typically mutual, supporting the notion of an engaged and active community.

Figure 9: BEND Community Assessment - Pakistan

COMM UNITY	Degree Centralit y	Densit y	E/I Index	Echo Chamber Index 1	Echo Chamber Index 2	Recip rocity	Specialty Concepts	Total Agent	Total Hasht ag	Total Tweet	Total Url
pakistan i_classif ication= 0	0.021	1	-0.623	1	1	1	NellieMassacre, PolicyParalysis, CoalScam, JayantiNatarajan, babri	504	544	657	435
pakistan i_classif ication= 1	0.048	1	0.893	1	0	1	in_sha_allah, اب_ابا_بھی_تبدیل_ہوگیا, HindutvaTerror, VHP, Yogi	57	108	84	56
pakistan i_classif ication= No Value	0.012	0.000202	0.989	0.059	0	1	twitter.com/F_Malik11/status/1732429030585180356/photo/1 , BabriMasjidDemoliti on, babrizindahai	211	2	1	1

Entities not classified as Pakistani exhibit a strong presence in discussions related to politically and socially sensitive topics, as indicated by specialty concepts like "NellieMassacre," "PolicyParalysis," "CoalScam," and notably "babri," which refers to the Babri Masjid demolition. These topics underscore a significant degree of polarization, aligning with the high density and the perfect echo chamber indexes suggesting a closed-loop communication among participants, reinforcing a singular perspective. Entities classified as Pakistani show an even more targeted approach, focusing on phrases that are significantly more polarizing, such as "HindutvaTerror" and mentions of Indian political figures like "Yogi" (referring to Yogi Adityanath). The presence of Urdu hashtags specific to Pakistani issues further draws into question how Pakistani accounts appear to be hijacking Indian political trends to gain notoriety.

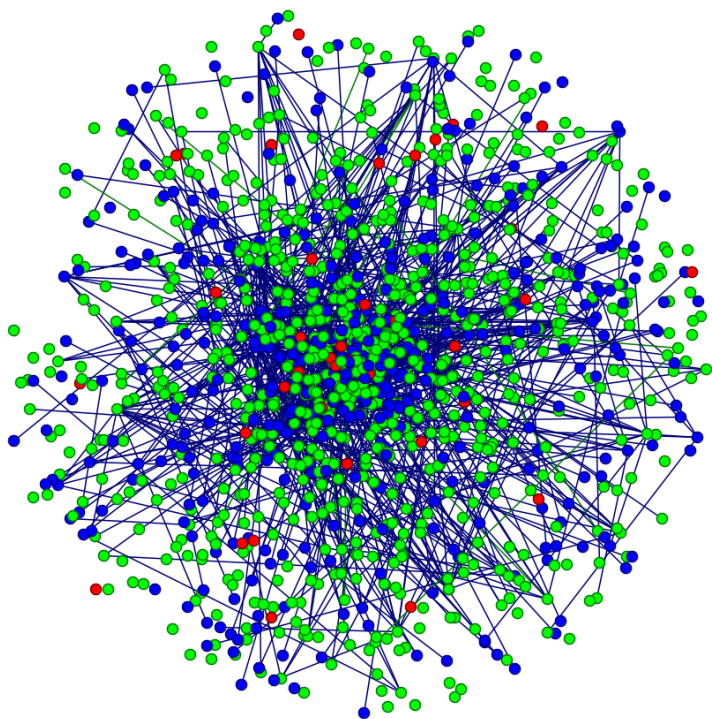
Figure 10: Communities Interaction-Targeted the Most

Rank	BEND Community	Value	Unscaled
1	pakistani_classification=0	0.885	683.194
2	pakistani_classification=1	0.108	83.521
3	pakistani_classification=No Value	0.007	5.283

This shows the communities that are the most targeted by agents based on exclusive interactions and interactions with the community's super-spreaders, where the targets appear to be accounts which are not classified as Pakistani.

Indian Political Hashtags #ByeByeModi

Figure 11: Network of #ByeByeModi colored by Pakistan Classification

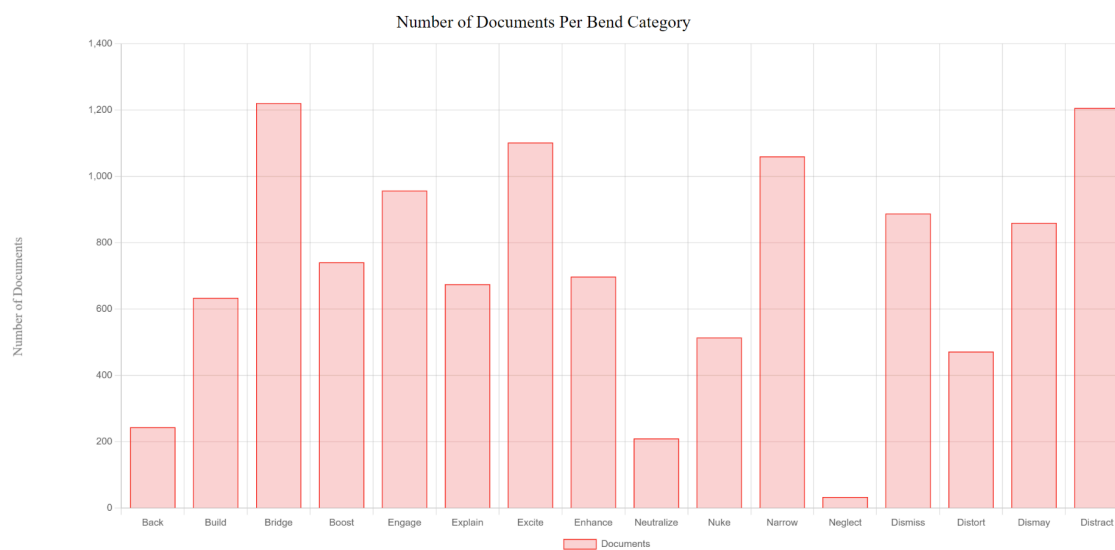


The red nodes indicate Pakistan accounts, green as not Pakistani and blue indicates no value. Given the presence of red nodes, this indicates Pakistani presence in the network.

Figure 12: ByeByeModi BEND Document Analysis

BEND Documents

The chart below shows the number of documents in each BEND category. A document is computed to be in a BEND category by averaging a collection of metrics to produce a likelihood and strength (or confidence) of category membership. A document with a likelihood value greater than the Mean is counted as a member.

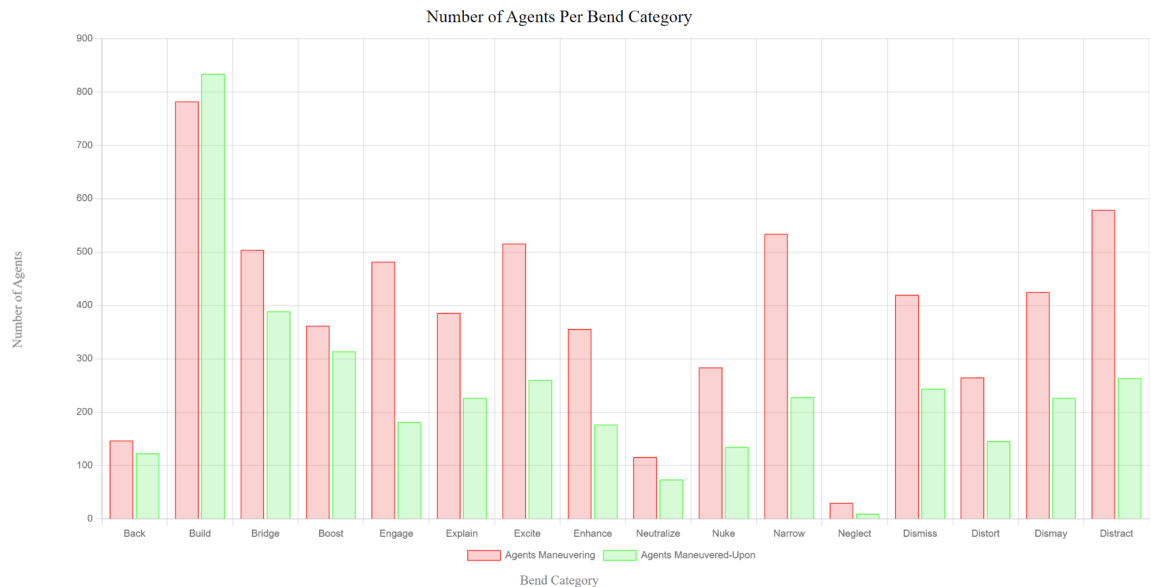


Number of documents	Number of documents in at least one BEND category	Sum of documents per BEND category
1,999	1,712 (85.64% of total documents)	11,516

Figure 13: ByeByeModi BEND Agent Analysis

BEND Agent Maneuvers

The chart below shows the number of agents in each BEND category that maneuvered or were maneuvered-upon. An agent is computed to be in a BEND category by averaging a collection of metrics to produce a likelihood and strength (or confidence) of category membership. An agent with a likelihood value greater than the Mean is counted as a member.



Number of agents	Number of agents maneuvering in at least one BEND category	Number of agents maneuvered-upon in at least one BEND category	Sum of agent maneuverers across all BEND categories	Sum of agents maneuvered-upon across all BEND categories
1,261	1,132 (89.77% of total agents)	1,132 (89.77% of total agents)	6,188	3,835

For both the document analysis for #BJPFreezesIndianDemocracy and #ByeByeModi, the most prevalent documents, "Distract," suggests tactics aimed at shifting public focus from significant issues to less crucial ones to mitigate criticism. The "Explain" category signifies extensive efforts to shape public perception by providing detailed rationales or justifications for certain policies or actions, likely to reinforce support or understanding. Meanwhile, "Dismiss" shows a frequent strategy of invalidating opposing views or critiques, reducing their impact by portraying them as irrelevant or incorrect, thus controlling the narrative and suppressing dissent.

While the BEND agent analysis finds that while there are significant maneuvers in categories like 'Bridge' and 'Explain,' only 284 agents (27.26% of total agents) are actually being maneuvered upon across any BEND category. This suggests that while there are active engagements, the number of agents directly impacted or targeted by maneuvers is relatively low compared to the total number involved for BJPFreezesIndianDemocracy. The high number of maneuvers, particularly in 'Explain' and 'Dismiss', indicates a focus on shaping narratives or countering opposing views. However, the lower proportion of agents maneuvered upon highlights a potential disparity between the breadth of strategic messaging efforts and the actual reach or influence on other agents within the network. However, this does not hold true for #ByeByeModi - a potential reason could be that there is higher Pakistani presence.

Figure 14: ByeByeModi BEND Community Assessment - Bot Activity

COMMUNITY	Degree Centrality	Density	E/I Index	Echo Chamber Index 1	Echo Chamber Index 2	Reciprocity	Specialty Concepts	Total Agent	Total Hashtag	Total Tweet	Total Url
IS_BO T=0	0.024	0.741	0.943	0.905	0	1	SuperstarRajinikanth, CongressWithIndia, RahulGandhiForIndia, SalaarEnglishOnNetflix, TVKVijay	50	157	257	135
IS_BO T=1	0.034	1	0.983	1	0	1	AskBabar, بنون خان گندگی, ReleaseSalmanAzhar, PSL9, Modi_Hatao	13	61	62	34
IS_BO T=No Value	0.009	0.368	-0.733	0.717	0	1	goodbyeBJP, Hinduism, MaharashtraPolitics, BharatJodoNayayYatra, BharatJodoNyayaYatra	1,198	847	1,437	701

The analysis suggests that bots, potentially affiliated with Pakistani interests, are actively involved in campaigns focusing on specific political and nationalistic themes. Topics such as "AskBabar," which include Urdu hashtags, and calls to "ReleaseSalmanAzhar," referring to an Islamic preacher linked to extremism, underscore targeted efforts to engage or mobilize support on issues pertinent to Pakistan. These bots exhibit considerable activity, particularly in communities where the interaction is heavily one-sided, indicated by a high E/I Index.

Figure 15: ByeByeModi BEND Community Assessment - Pakistan

COMMUNITY	Degree Centrality	Density	E/I Index	Echo Chamber Index 1	Echo Chamber Index 2	Reciprocity	Specialty Concepts	Total Agent	Total Hashtag	Total Tweet	Total Url
pakistan_i_classif	0.008	0.850	-0.806	0.947	1	1	goodbyebjp, Modi, Hinduism,	806	882	1,708	830

ication=0							MaharastraPolitics, BharatJodoNayayYatra				
pakistan_i_classification=1	0.024	1	0.950	1	0	1	ByeByeIndia, AskBabar, ووٹ ڈالو خان نکالو, ReleaseSalmanAzhari, PSL9	38	94	47	27
pakistan_i_classification=No Value	0.001	5.751e-06	0.999	0.018	0	1	BJPFailsIndia, byebyemodi	417	2	1	0

The ByeByeModi BEND Community Assessment for Pakistan shows significant overlap between the topics discussed by entities classified as bots and those associated with Pakistan, particularly with specialty concepts such as "ReleaseSalmanAzhar" and "AskBabar". This alignment suggests a potential linkage between bots and Pakistani interests, highlighting the use of automated accounts to amplify narratives that might influence or interfere with Indian sociopolitical discourse, specifically targeting security issues and trending topics within India.

Economic Undermining #IndiaOut, #BoycottIndianProducts

The #IndiaOut campaign initially emerged in the Maldives and subsequently gained traction in Bangladesh in 2024, following Bangladesh's election and Sheikh Hasina's electoral win. Hasina was found by the opposition, the right wing Bangladesh Nationalist Party (BNP), to be "Pro-India" hence #BoycottIndianProducts emerged as political attacks between parties in Bangladesh, and the same is true in Maldives. This campaign reflects a broader sentiment of dissatisfaction and perceived "interference" by India in the domestic politics of neighboring countries. The BNP, which has utilized the campaign to voice grievances against Indo-Bangladesh relations, culminating in calls for boycotting Indian products under the banner of #BoycottIndianProducts.

Figure 16: Network of #IndiaOut colored by Pakistan Classification

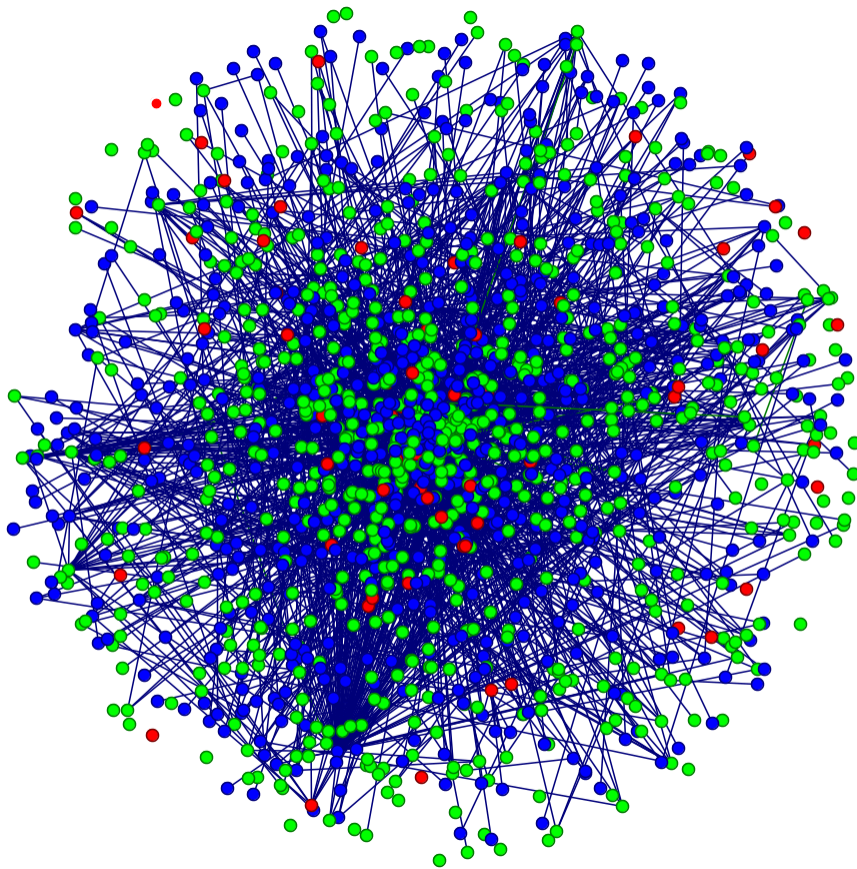
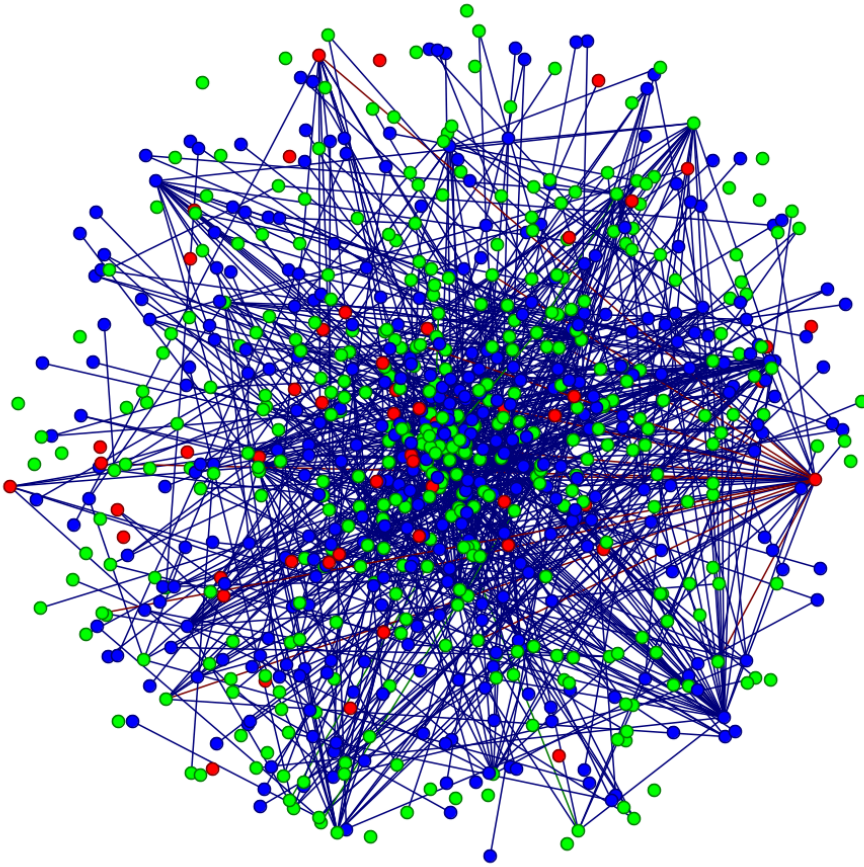


Figure 17: Network of #BoycottIndianProducts colored by Pakistan Classification

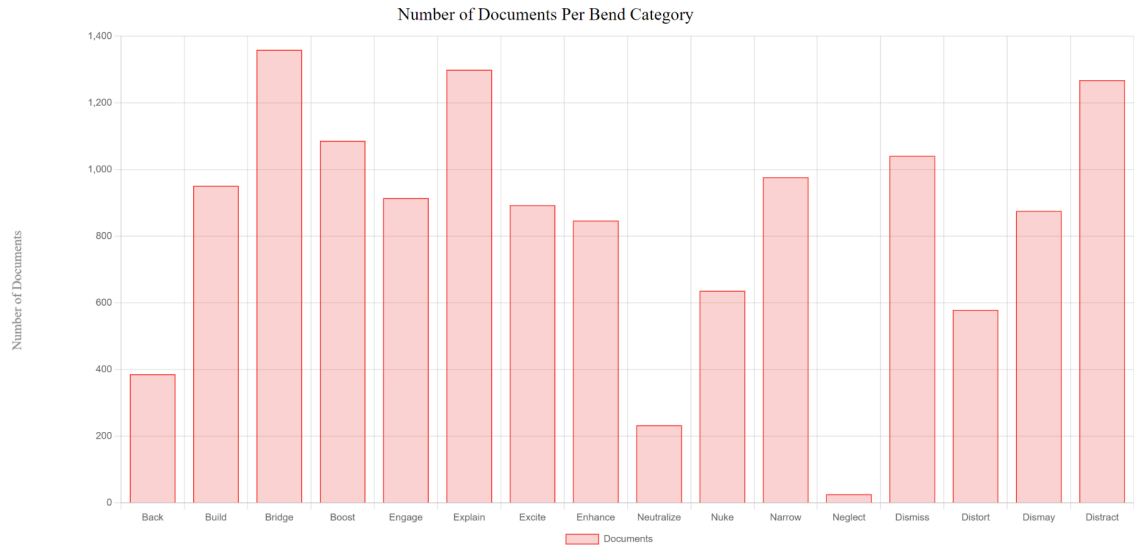


The red nodes indicate Pakistan accounts, green as not Pakistani and blue indicates no value. Given the presence of red nodes, this indicates Pakistani presence in the network.

Figure 18: IndiaOut BEND Document Analysis

BEND Documents

The chart below shows the number of documents in each BEND category. A document is computed to be in a BEND category by averaging a collection of metrics to produce a likelihood and strength (or confidence) of category membership. A document with a likelihood value greater than the Mean is counted as a member.

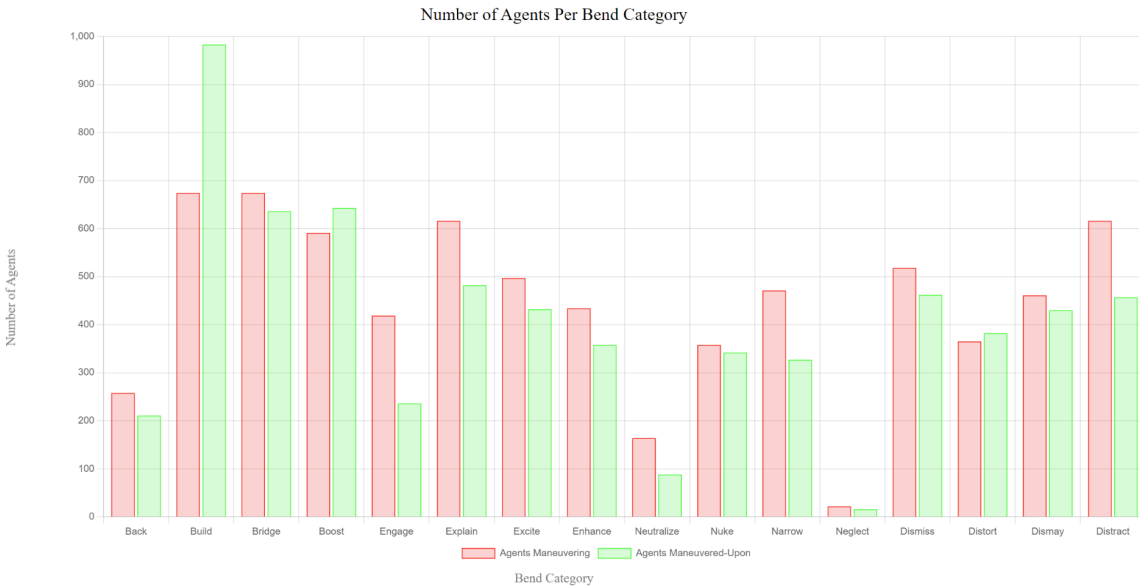


Number of documents	Number of documents in at least one BEND category	Sum of documents per BEND category
1,992	1,747 (87.70% of total documents)	13,371

Figure 19: IndiaOut BEND Agent Analysis

BEND Agent Maneuvers

The chart below shows the number of agents in each BEND category that maneuvered or were maneuvered-upon. An agent is computed to be in a BEND category by averaging a collection of metrics to produce a likelihood and strength (or confidence) of category membership. An agent with a likelihood value greater than the Mean is counted as a member.

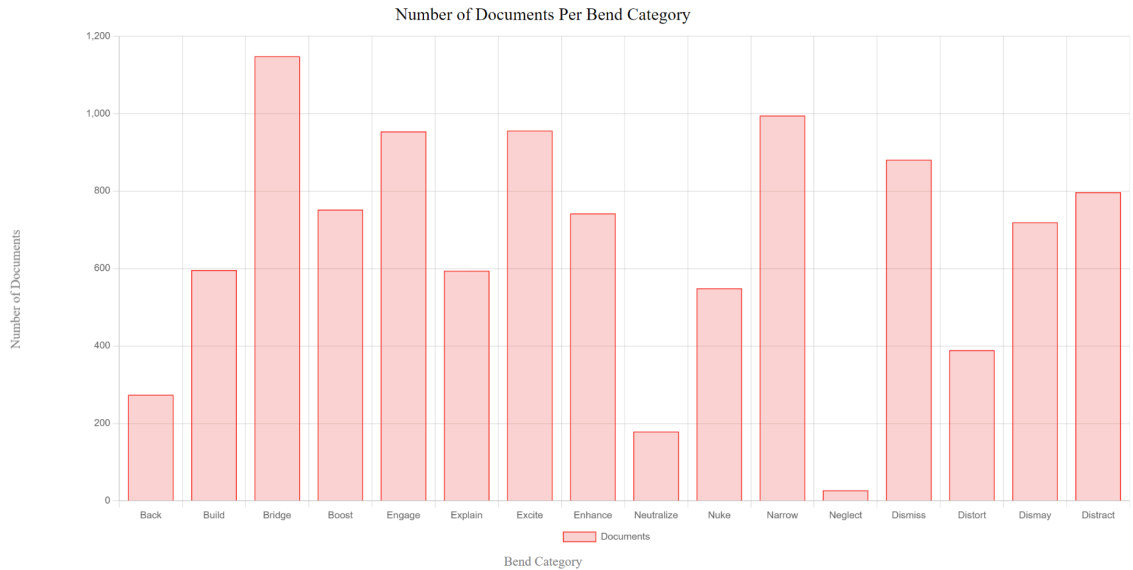


Number of agents	Number of agents maneuvering in at least one BEND category	Number of agents maneuvered-upon in at least one BEND category	Sum of agent maneuverers across all BEND categories	Sum of agents maneuvered-upon across all BEND categories
1,538	838 (54.49% of total agents)	1,462 (95.06% of total agents)	7,138	6,485

Figure 20: BoycottIndianProducts BEND Document Analysis

BEND Documents

The chart below shows the number of documents in each BEND category. A document is computed to be in a BEND category by averaging a collection of metrics to produce a likelihood and strength (or confidence) of category membership. A document with a likelihood value greater than the Mean is counted as a member.

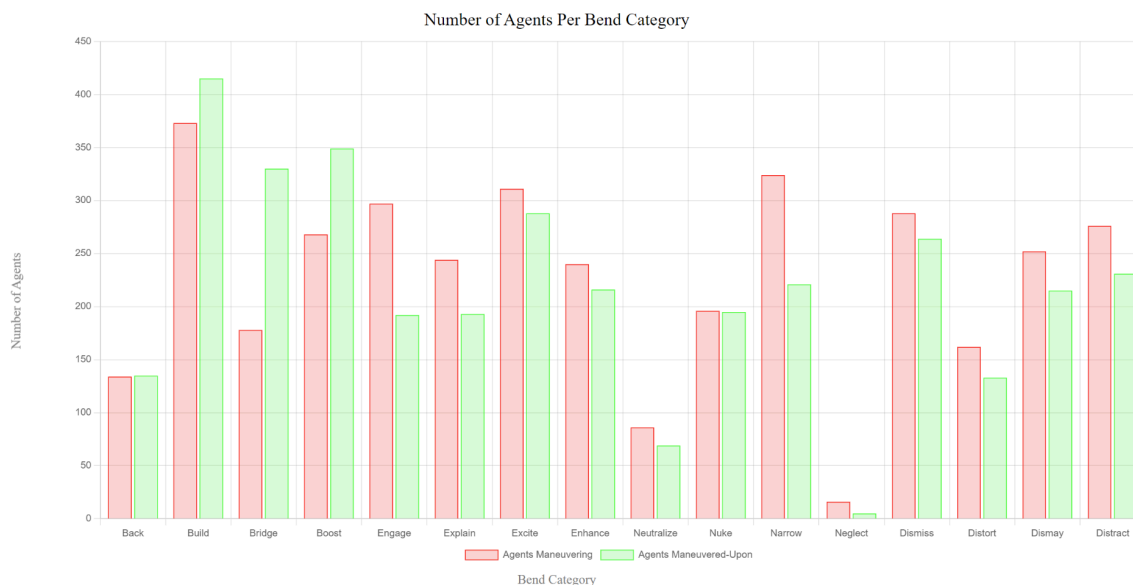


Number of documents	Number of documents in at least one BEND category	Sum of documents per BEND category
1,664	1,491 (89.60% of total documents)	10,552

Figure 21: BoycottIndianProducts BEND Agent Analysis

BEND Agent Maneuvers

The chart below shows the number of agents in each BEND category that maneuvered or were maneuvered-upon. An agent is computed to be in a BEND category by averaging a collection of metrics to produce a likelihood and strength (or confidence) of category membership. An agent with a likelihood value greater than the Mean is counted as a member.



Number of agents	Number of agents maneuvering in at least one BEND category	Number of agents maneuvered-upon in at least one BEND category	Sum of agent maneuverers across all BEND categories	Sum of agents maneuvered-upon across all BEND categories
819	447 (54.58% of total agents)	673 (82.17% of total agents)	3,645	3,451

In both #IndiaOut and #BoycottIndianProducts, there appears to be slightly above 50% agents maneuvering.

Figure 22: IndiaOut BEND Community Assessment - Bot Activity

[illegible]

Figure 23: IndiaOut BEND Community Assessment - Pakistan

COMMUNITY	Degree Centrality	Density	E/I Index	Echo Chamber Index 1	Echo Chamber Index 2	Reciprocity	Specialty Concepts	Total Agent	Total Hashtag	Total Tweet	Total Url
pakistan_i_classification=0	0.016	0.818	-0.685	0.935	1	1	NTK, Seeman, WeHateIndia, IndiaMilitaryOut, GettingBackChagos	778	900	1,657	754
pakistan_i_classification=1	0.019	0.969	0.915	0.990	0	1	Islamic, Economist, PTI, PTIOfficial, AlJazeera	64	138	104	46
pakistan_i_classification=No Value	0.002	6.193e-06	0.999	0.018	0	1	IndiaOut	696	1	1	0

The Pakistan data highlights several key insights into the nature of online discourse and potential interference in the #IndiaOut campaign. The community with a high E/I Index (0.915) and substantial density (0.969) features the tags "PTI" and "PTIOfficial," which are associated with Pakistan's political party PTI, indicating a structured and potentially coordinated effort within these communities to push narratives against Indian influence. The presence of high activity levels, especially in terms of tweets and hashtags, underscores the intensity and reach of these campaigns, reflecting a significant push within digital spaces that could be indicative of orchestrated political interference.

Figure 24: BoycottIndianProducts BEND Community Assessment - Bot Activity

COMMUNITY	Degree Centrality	Density	E/I Index	Echo Chamber Index 1	Echo Chamber Index 2	Reciprocity	Specialty Concepts	Total Agent	Total Hashtag	Total Tweet	Total Url
IS_BO T=0	0.082	0.766	0.968	0.915	0	1	Thailand, Malayasia, HumanRightsViolations, BoycottIndianServices, israelTerrorists	16	57	137	102
IS_BO T=No Value	0.031	0.292	-0.870	0.663	0	1	Boycott, BoycottIndi, RestoreDemocracy, FreeDemocracy, SaveBangladesh	802	585	1,261	738
IS_BO T=1	-	-	-	-	-	-	-	-	-	-	-

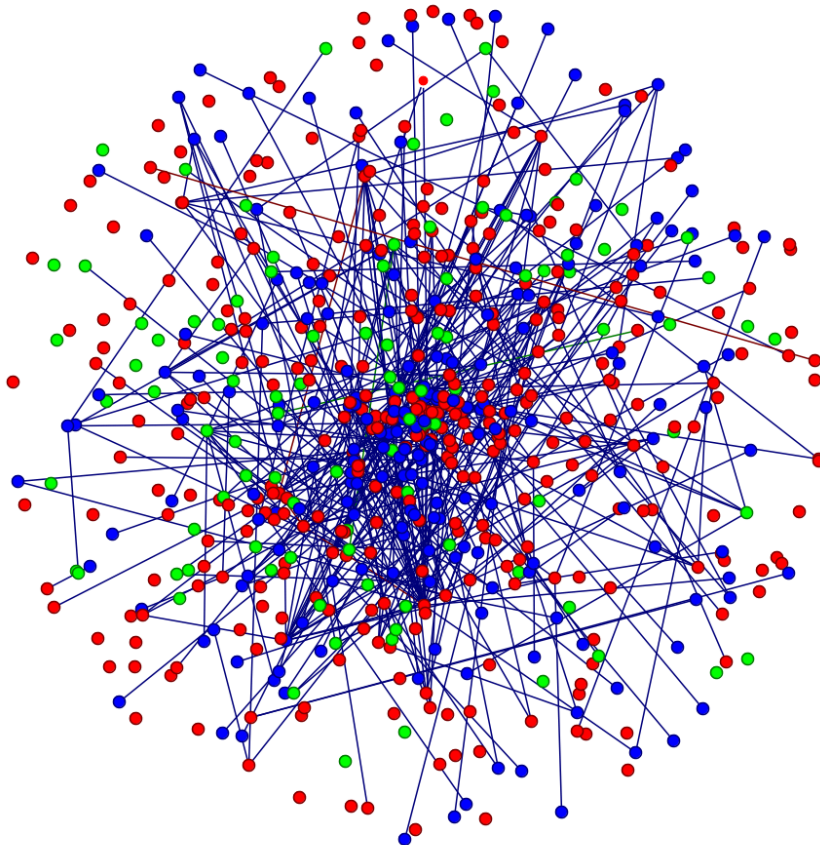
Figure 25: BoycottIndianProducts BEND Community Assessment - Pakistan

COMMUNITY	Degree Centrality	Density	E/I Index	Echo Chamber Index 1	Echo Chamber Index 2	Reciprocity	Specialty Concepts	Total Agent	Total Hashtag	Total Tweet	Total Url
-----------	-------------------	---------	-----------	----------------------	----------------------	-------------	--------------------	-------------	---------------	-------------	-----------

pakistan i_classif ication= 0	0.035	0.995	-0.559	0.998	1	1	TakeBackBangladesh , BoycottIndi, RestoreDemocracy, RestoreCaretakerGov t, FreeDemocracy	394	540	1,220	687
pakistan i_classif ication= 1	0.035	1	0.872	1	0	1	snow, PalestineWillBeFreeS oon, Hamas, israelterroriststate, IsraelCrimesagainstH umanity	54	105	285	239
pakistan i_classif ication= No Value	0	0	1	0	0	1		371	0	0	0

#KashmirSolidarityDay

Figure 26: Network of #KashmirSolidarityDay colored by Pakistan Classification



The red nodes indicate Pakistan accounts, green as not Pakistani and blue indicates no value. Given the presence of red nodes, this indicates Pakistani presence in the network.

Figure 27: KashmirSolidarityDay BEND Community Assessment - Government Account

COMMUNITY	Degree Centrality	Density	E/I Index	Echo Chamber Index 1	Echo Chamber Index 2	Reciprocity	Specialty Concepts	Total Agent	Total Hashtag	Total Tweet	Total Url
government_agency=0	0.017	1	-0.760	1	1	1	Opinion, protest, taxes, electricitybills, KashmirIsPakistan	419	544	864	660
government_agency=1	0.016	1	0.936	1	0	1	UNSC, fbise, federalboard, Khanewal, PakistanStandsWithKashmir	28	81	92	81
government_agency=None Value	0	0	1	0	0	1	—	211	0	0	0

Figure 49: Tweet from Pakistani Embassy in USA



The BEND Community Assessment for KashmirSolidarityDay as presented by the Pakistani Embassy in the USA illustrates a strategic use of social media by governmental entities to advocate for Kashmir. The accounts show a strong focus on international and educational themes, specifically highlighting Pakistan's stance on Kashmir through engagements that emphasize UN involvements and the support of educational bodies within Pakistan. The term "PakistanStandsWithKashmir" signals a diplomatic effort to solidify and broadcast Pakistan's support for Kashmir on a global stage, leveraging platforms like the United Nations to frame Kashmir as an international human rights concern rather than merely a bilateral issue. This strategy is aligned with broader public relations efforts to garner international support and sympathy for Kashmir, underscored by social media campaigns promoting solidarity.

Limitations

The data presented faces several limitations that must be considered when interpreting the findings. Firstly, there is missing retweet data from Twitter, which restricts the ability to fully understand the reach and impact of the analyzed tweets. Furthermore, the analysis is confined to certain platforms and does not account for the potential influence of other public military magazines that may contribute to the broader narrative. Another significant limitation is the risk of undetectable obfuscation, where other nation states might masquerade as Pakistani entities to

influence perceptions, adding a layer of complexity to data authenticity. Additionally, the use of LLMs for data annotation could introduce errors if the information is inaccurately labeled, leading to skewed results. Finally, while the focus is on hashtags trending on Pakistani Twitter, it does not encompass those that may be prevalent in other regional Twitter Spheres, possibly overlooking wider trends and sentiments. These limitations suggest that while the data provides valuable insights, it should be interpreted with caution, acknowledging the scope and potential biases present.

Discussion

The analysis of Pakistani influence operations suggests a concerted effort to amplify religious and ethnic divisions within India through the dissemination of exaggerated narratives. These operations predominantly focus on portraying Hindus as extremists and exaggerating incidents of violence against Muslims, which serve not only to justify military actions by Pakistan against India but also to fuel Pakistani nationalism. Moreover, there is significant emphasis on supporting separatist movements in regions like Kashmir, Punjab, and to a lesser extent in Northeast and South India, where the narrative often revolves around local independence and human rights abuses purportedly committed by India.

However, efforts to broadly undermine India's image as a stable and progressive region by highlighting issues like poor sanitation and gender-based violence seem less pronounced. The strategies identified also include stoking polarization between major Indian political parties, particularly through Anti-BJP messaging. While there is evidence of automated bots and possibly fake accounts that echo Pakistani stances, definitively attributing these to direct Pakistani state actions is complex due to the obscured nature of digital identities and the possibility of third-party actors.

The results highlight real-world implications, particularly concerning the perception and international relations of India. Pakistani digital influence operations have strategically amplified specific narratives, creating heightened perceptions of India as a region rife with conflict and instability based on religious zealotry and extremism. This systemic manipulation complicates the global understanding of legitimate security concerns facing India, as distinguishing between exaggerated threats and genuine issues becomes increasingly challenging. Such distorted perceptions can undermine India's image internationally, potentially influencing foreign policy decisions and diplomatic relations unfavorably. The international community may develop skewed views on India's political and social landscape, which could result in biased diplomatic interactions and potentially hinder India's efforts to forge or sustain strategic alliances and partnerships.

In terms of fueling real world violence, these influence operations exacerbate existing Hindu-Muslim tensions within India, potentially inciting sectarian strife. By deliberately highlighting and distorting incidents of communal violence, these campaigns fuel polarization, heightening the sense of identity-based division among the Indian population. This manipulation not only deepens societal fractures but could also escalate to real-world violence and riots, as online narratives incite or justify aggressive actions on the ground. The continuous propagation of such divisive content poses a persistent threat to India's internal security and social harmony, suggesting a critical need for focused research into the direct impact of these operations on communal relations and public safety.

Given these findings, it's crucial for policymakers and security specialists to focus on developing and implementing advanced counter-strategies. This includes improving abilities to track and counteract misinformation efforts, increasing public education on the realities of digital propaganda, and enhancing community resilience against divisive influences. For future research, conducting further scholarly studies to directly link online Pakistani influence operations targeting India with their effects on societal unrest and national security is vital.

Conclusion

Pakistani influence campaigns, particularly those that amplify religious and ethnic divisions within India, have the potential to pose significant threats to Indian sovereignty and security. These operations, frequently underscored by entities linked to the Pakistani political party PTI and its affiliates, utilize tactics like hashtag hijacking to spread their narratives. This indicates a strategic effort to sway public opinion and manipulate social discourse, which necessitates a broader examination across various platforms to fully understand the scope and impact. Monitoring and analysis need to be intensified, focusing on the textual data from social media campaigns and the specific political agendas being pushed. Creating better-annotated datasets will be crucial for distinguishing between authentic trending topics and those artificially boosted by influence operations. This enhanced scrutiny will aid in identifying and mitigating the potentially destabilizing effects of such campaigns on India's internal affairs and its image abroad.

References

- Akhtar, N. (2009). Polarized politics: the challenge of democracy in Pakistan. *International Journal on World Peace*, 31-63.
- Altman, N., Carley, K. M., & Reminga, J. (2020). ORA user's guide 2020. Carnegie-Mellon Univ. Pittsburgh PA Inst of Software Research International, Tech. Rep, 2, 2.
- Bergh, A. (2020). Understanding influence operations in social media. *Journal of Information Warfare*, 19(4), 110-131.
- Bernays, E. L. (2005). *Propaganda*. Ig publishing.
- Beskow, D. M., & Carley, K. M. (2018, July). Bot-hunter: a tiered approach to detecting & characterizing automated activity on Twitter. In Conference paper. SBP-BRiMS: International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation (Vol. 3, No. 3).
- Carley, K. M. (2020). Social cybersecurity: an emerging science. *Computational and Mathematical Organization Theory*, 26(4), 365-381.
- Carley, L. R., Reminga, J., & Carley, K. M. (2018, July). Ora & Netmapper. In International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation. Springer (Vol. 1, No. 3.3, pp. 2-1).
- Chen, W., Castillo, C., & Lakshmanan, L. V. (2022). *Information and Influence Propagation in Social Networks*. Springer Nature.
- Church, K. W. (2017). Word2Vec. *Natural Language Engineering*, 23(1), 155-162.
- Cook, C. L., Schaafsma, J., Antheunis, M. L., Shahid, S., Lin, J. H. T., & Nijtmans, H. W. (2021). Trolls without borders: A cross-cultural examination of victim reactions to verbal and silent aggression online. *Frontiers in Psychology*, 12, 549955.
- Fair, C. C. (2014). *Fighting to the End: The Pakistan Army's Way of War*. Oxford University Press, USA.
- Fallis, D. (2015). What is disinformation? *Library Trends*, 63(3), 401-426.
- Flew, T., & Iosifidis, P. (2020). Populism, globalization, and social media. *International Communication Gazette*, 82(1), 7-25.
- Haqqani, H. (2010). *Pakistan: Between Mosque and Military*. Carnegie Endowment.

Harshe, R. (2003). Cross-border terrorism: Road-block to peace initiatives. *Economic and Political Weekly*, 3621-3625.

Hussain, S., Bostan, H., & Qaisarani, I. (2022). Trolling of female journalists on Twitter in Pakistan: an analysis. *Media International Australia*, 1329878X221145977.

Hussain, S., Shahzad, F., & Saud, A. (2021). Analyzing the state of digital information warfare between India and Pakistan on Twittersphere. *SAGE Open*, 11(3), 21582440211031905.

Inam, U. H., Basharat, H., & Muzammil, S. (2021). The role of humor in understanding the trolling behavior of social media users in Pakistan. *International Journal of Media and Information Literacy*, 6(2), 321-328.

Kanwal, G. (1999). Proxy war in Kashmir: Jihad or state-sponsored terrorism? *Strategic Analysis*, 23(1), 55-83.

Kiessling, H. (2016). *Faith, Unity, Discipline: The Inter-Service-Intelligence (ISI) of Pakistan*. Oxford University Press.

Majid, A., Hamid, A., & Habib, Z. (Majid, Hamid, & Habib, 2014)