scientific reports

OPEN

Evidence of inter-state coordination amongst state-backed information operations

Xinyu Wang✉, Jiayi Li, Eesha Srivatsavaya & Sarah Rajtmajer✉

Since 2018, Twitter has steadily released into the public domain content discovered on the platform and believed to be associated with information operations originating from more than a dozen state-backed organizations. Leveraging this dataset, we explore inter-state coordination amongst state-backed information operations and find evidence of intentional, strategic interaction amongst thirteen different states, separate and distinct from within-state operations. We find that coordinated, inter-state information operations attract greater engagement than baseline information operations and appear to come online in service to specific aims. We explore these ideas in depth through two case studies on the coordination between Cuba and Venezuela, and between Russia and Iran.

The current reach of social media platforms and their affordances for cheap and easy content dissemination, profiling, and targeting, have established social media as a primary avenue for information operations—efforts to manipulate public opinion by intentionally altering the information environment[1]. A substantial literature has emerged studying the tactics and strategies of information operations, particularly on Twitter, where data has been widely available[2–4]. These studies have focused on campaigns attributed to individual states or state-backed organizations. To the best of our knowledge, no prior work has looked at collaboration amongst states in these efforts. Yet, examples of international collaboration for the dissemination of propaganda date back to the first and second World Wars[5–8]. Our research explores evidence of inter-state coordination amongst state-backed information campaigns, operationalized through the following two research questions.

RQ1 Do state-backed information campaigns operating on Twitter collaborate across states? If so, what distinguishes these efforts from internal information operations in terms of design, deployment, and impact?
RQ2 Can we categorize strategic and tactical mechanisms underlying inter-state information operations? e.g., specific roles of individual accounts in support of collusion?

We extract interaction networks amongst 13 state-backed campaigns operating on Twitter between 2011 and 2021, perform static and dynamic pairwise analyses of observed activity, and highlight the varied structure of inter-state information operations through two case studies. Our findings indicate that inter-state operations attract greater engagement than intra-state operations. We suggest that the strategies employed by state-backed information operations serve to create and maintain a desirable information habitat, e.g., by engaging in ambient affiliation through common hashtags[9], initiating network expansion for increased exposure[10], and referencing controversial topics to gain attention[11]. Our findings represent the first insights into tactics and strategies underlying global cooperation and collusion amongst states in strategic information operations deployed through social media.

Related work
Strategic information operations. Starbird et al.[1] use the term *strategic information operations* (IO) to refer to efforts by individuals and groups, including state and non-state actors, to manipulate public opinion and change how people perceive events in the world by intentionally altering the information environment. Tracing their roots to TV and radio propaganda in the 20th century, and in some variation even much earlier[12], the modern digital age and in particular today's social media landscape have enabled these efforts with unparalleled efficiency and have raised a unique set of questions around response[13].

A primary subset of information operations aims to disseminate inaccurate, incorrect, or misleading information, so-called *disinformation*. Disinformation operations have been strongly associated with political

College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA, USA. ✉email: xzw5184@psu.edu; smr48@psu.edu

campaigns[14], focused primarily on social media manipulation of public opinion through bot accounts and paid workers[15–18]. In attempt to combat disinformation, scholars have focused efforts on developing technical solutions for automated detection of multimodal disinformation, e.g., fact-checking claims in text, identifying falsified images, and detecting deception through speech and facial expression in video[19–23].

While the vast majority of studies reporting on information operations focus on individual state efforts, see e.g.,[24–26], or events, e.g., presidential elections[27–29], a few have presented a global overview of information operations and therefore are most similar to our work here. Bradshaw and Howard's 2018 report provides an inventory of organized social media manipulation campaigns in 70 countries[16]. Niblock et al. have compiled comprehensive summary statistics and visualizations of all publicly-shared state-backed information operations on Twitter[14]. Our study likewise provides global analyses, but is unique in its focus on inter-state coordination.

Information operations as collaborative work. Recent work in the Human–Computer Interaction (HCI) literature has highlighted the participatory nature of online information operations and interpreted the dissemination of manipulated information via the lens of online communities[1]. Psychological theories, e.g., distributed cognition[30], offer theoretical bases for examining how social environments, such as mainstream social media platforms, impact collective behaviors among social ties and facilitate disinformation[31]. In particular, sociotechnical systems allow information operations to target, integrate with, and leverage the activities of online crowds—resulting in a combination of orchestrated and organic elements and behaviors. Schoch et al.[32] find that online political astroturfing consistently leaves similar patterns of coordination across distinct geographical settings and historical periods. The collaboration amongst state-backed accounts we describe in this work appears to be orchestrated. However, the audience with whom they engage and acts of implicit coordination with native users, e.g., via hashtagging, embedded URLs, is critical to our work.

Idea habitats. Many mechanisms have been proposed to model the diffusion of misinformation and to expose environmental factors that drive false beliefs[33,34]. Current studies capture a number of psychological, demographic, and environmental variables contributing to the acquisition and spread of misinformation[34–37]. Work in cognitive science has suggested that the ways and extent to which individuals recall and distribute information depends on collections of environmental cues, so-called idea habitats[38]. These cues may include social and political context, linguistic characteristics, and topics of conversation. Common practices such as audience segmentation and micro-targeting represent efforts to nurture habitats receptive to particular narratives. Suitable habitats support self-reinforcing information flows, independent of content validity[39]. In fact, studies have shown that false news spreads further, faster, and more broadly than legitimate news, a phenomenon which has magnified challenges to combat mis- and disinformation in recent years[40]. Our work dovetails with fundamental notions of idea habitat. We examine how accounts involved in information operations cooperate to establish context conducive to the distribution of misleading information.

Role analysis in social networks. Prior work has analysed the different roles of actors/nodes in social network graphs using both graph structure-based and content-based approaches[41]. Structural role analysis is often used for influence maximization tasks[42,43] while content-based role analyses have found use in modeling the growth of online communities[41,44,45]. Our work uses a hybrid approach to define roles in inter-state information operations integrating network metrics and content-based analyses.

Dataset

Since October 2018, Twitter has made public the tweets, media, and account-related information of users presumed to be involved in state-linked information operations, provided through the Twitter Moderation Research Consortium (TMRC). The TMRC suggests these users are engaged in *manipulation that can be reliably attributed to a government or state-linked actor*[46]. We aggregate all account activity shared by the TMRC between 2007 and 2021. In total, this represents 23 state-backed information operations consisting of the full activity of 84,262 distinct accounts and approximately 120 million archived tweets.

During preprocessing, we made the following modifications to the complete dataset. We combined accounts designated by Twitter as linked to Egypt and UAE. Twitter's documentation subsequent to the release of these accounts indicated that much of their activity was attributed to an operation managed out of both countries targeting Qatar and Iran with messaging supportive of the Saudi government[47]. We did not include data from one release in March 2020 which Twitter attributed to "Egypt, UAE and Saudia Arabia" because attribution to a single country was not possible. This omitted subset of the data was relatively small (5350 accounts, 6.3% of the total dataset). We did not find evidence of inter-state activity in content originating from Armenia, Bangladesh, Thailand, Tanzania, Mexico, Catalonia, Ghana, Nigeria, or Spain. These nine countries are therefore included in our dataset and analyses but not represented in the results, which focus on inter-state coordination. In sum, the number of tweets represented by these nine countries accounts for less than 0.2% of the data. Dataset statistics are further detailed in the analyses below.

RQ1: Inter-state activity

We use the terms "coordination" and "coordinated operations" to characterize purposeful collaboration in service to shared objectives. Informed by explanations of the dataset provided by the TMRC upon each data release, our analyses make the following assumptions: (1) all activities associated with accounts tagged by Twitter as participating in information operations are part of those operations; (2) all pairwise interactions between state-backed influence operation actors are coordinated information operations/platform manipulation. Evidence of inter-state coordination is informed by static and dynamic network and content analyses across state-linked accounts.

Inter-state interaction network. We build a global inter-state interaction network amongst state-linked accounts (Fig. 1a). Nodes represent accounts and directed edges represent retweets, replies, mentions, and quotes, between 2011 and 2021. Node color corresponds to country and edge color matches source node. We observe two predominant substructures within the network. The first is a radiating pattern, consisting of one or a few central nodes with high out-degree centrality (e.g., Fig. 1b(i)). This motif appears for countries with either dominating in-degree centralization or out-degree centralization. Central nodes function as either content creators or self-promoters surrounded by a substantial number of followers and amplifiers to disseminate content and establish new social ties. The contrasting motif is balanced with similar in- and out-degree (e.g., Fig. 1b(ii)). Figure 1c distills the inter-state interaction network via aggregation by country. Node size is proportional to log-scaled number of associated accounts. Edge width is proportional to log-scaled number of interactions in each pair-wise coordination and edge color matches source node. We note that Cuba, Serbia, and Ecuador appear to
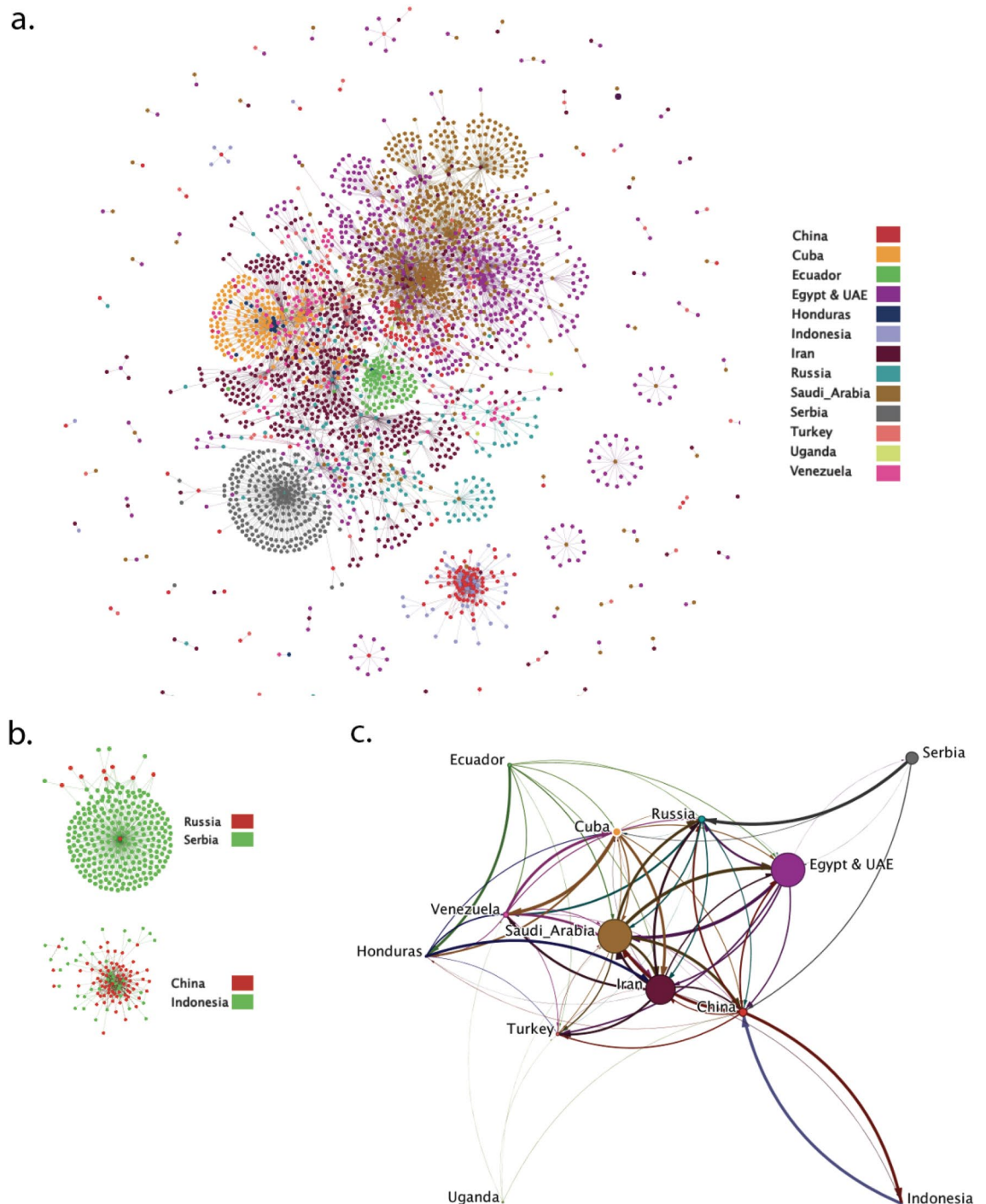
Figure 1. (**a**) Global inter-state interaction network; (**b**) observed network motifs; (**c**) inter-state interactions, aggregated by country.

use retweets and replies to connect with other states for network expansion and content promotion. Whereas, Venezuela, Russia, Turkey, and Iran are predominantly the target of interactions. Accounts linked to these countries disseminate relatively more original content. Indonesia, Egypt & UAE, China, and Saudi Arabia exhibit more balanced structures. We calculate the reciprocity of each state in the weighted network, decomposing dyadic fluxes into a fully reciprocated component and a fully non-reciprocated component[48]. Reciprocity levels are 0.944, 0.924, 0.844, and 0.777 for Indonesia, Egypt & UAE, China, and Saudi Arabia, respectively. With the exception of Honduras, where we observe a comparable number of interactions as source and target, 94.9% of global outgoing interactions are directed to Russia, and 88.3% of global in-coming interactions originate from Ecuador (reciprocity: 0.0079).

Temporal analysis of inter- and intra-state activity. Figure 2 shows the cumulative inter- and intra-state interaction counts over time, along with labeled interaction peaks for each of the 13 states. We observe that initial inter-state activity lags behind intra-state activity (avg. lag approximately 2 years). Notably, a majority of inter-state interactions reach peak activity synchronously in late 2017 and 2018. Comparing inter- and intra-state interactions for each state individually, we find that 10 out of the 13 countries in our dataset have substantially different temporal patterns. That is, in most cases, inter-state operations do not occur concurrently with intra-state operations. Rather, they represent what appears to be a separate strategic operation. Seven peaks in inter-state and intra-state activities occur more than one year apart, three occur three to twelve months apart, two peaks occur within three months of one another, and one occurs simultaneously within the same month).

Measuring engagement with inter-state activity. Observing that one aim of inter-state coordination appears to be increased visibility, we measure differences in engagement statistics between inter- and intra-state activity using a two sample T-test. We perform a priori power analysis to determine the minimum sample size, set the significance level α to 0.05, power to 0.8, and Cohen's d effect size to 0.2. We obtain a minimum effective sample size of 394 to perform target statistical testing. Using this threshold value, we filter out two states (Turkey and Uganda) with less than 394 inter-state interactions. Given the significantly smaller fraction of quotes in the dataset (the number of quote tweets in the dataset is less than the effective sample size), we select likes, retweets, and replies as three indices for comparative study. We perform a Welch's T test to determine if observed differences are significant. Table 2 gives these results. In 7 of 11 states, inter-state interactions receive more likes on average than intra-state interactions, 5 of which are significant. Similar trends hold for retweets and replies. We observe that countries like Venezuela, China, Indonesia, and Egypt and UAE which are more active in inter-state coordination also apply tactics discussed later (RQ2).

We additionally perform engagement comparisons between inter-state interactions and state-backed accounts' interactions with external accounts, i.e., accounts not tagged by Twitter as state-backed actors, presumed "normal" accounts). As the data is imbalanced (see Table 1), we randomly sample external interactions matching the observed number of inter-state interactions. With a similar level of variance, we perform a regular two-sample T-test. Results are provided in Table S3 of "Supplementary Material". We observe a general pattern of more likes and retweets associated with external accounts. This is expected as external accounts have greater visibility, e.g., news outlets, and likes and retweet counts are derived from the original post. However, number of replies associated with inter-state interactions is substantially *greater* than those associated with external interactions, indicating success of inter-state coordination to prompt meaningful engagement (e.g., Cuba, China, and Indonesia).
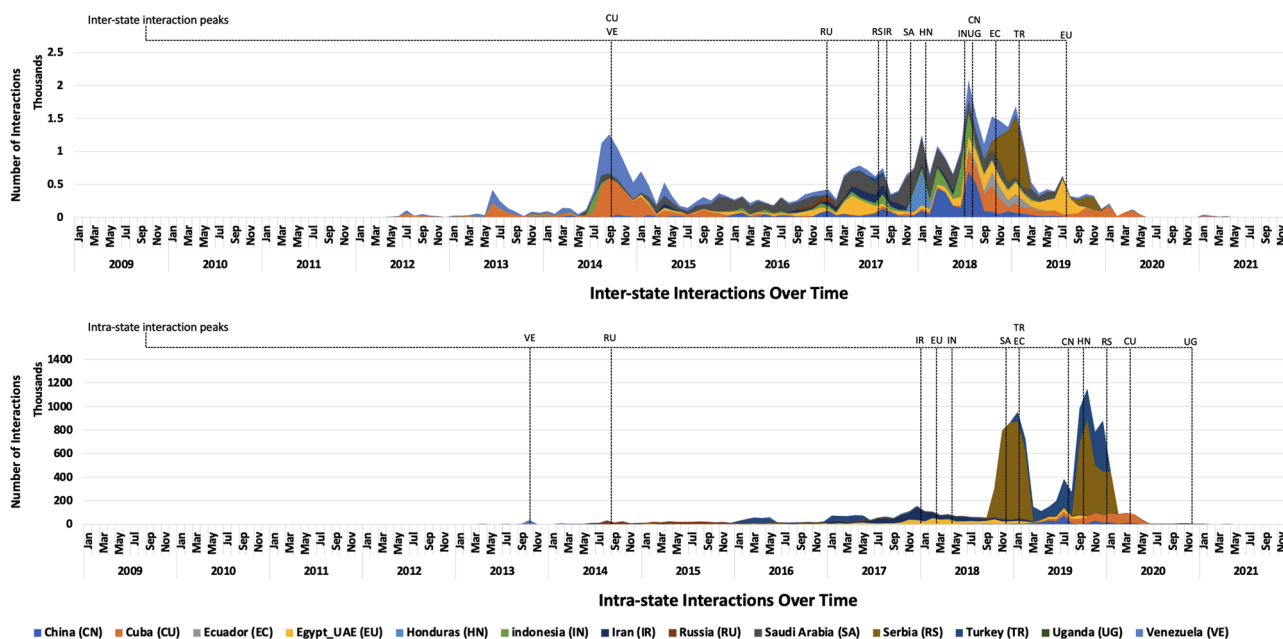

Figure 2. Cumulative intra- and inter-state interactions, over time.

Country	Inter-state actors	Inter-state interactions (source)	Inter-state interactions (target)	Intra-state interactions	External interactions & isolates	Total actors	Total tweets
Cuba (CU)	181	6649	1027	742,654	4,055,338	526	4,805,668
Venezuela (VE)	114	1521	6131	219,973	10,368,229	2261	10,595,854
Iran (IR)	569	1528	3275	905,384	9,524,551	7025	10,434,738
Russia (RU)	143	356	5227	352,943	4,124,305	1741	4,482,831
Serbia (RS)	254	3267	1	5,750,528	7,968,305	8558	13,722,101
Turkey (TR)	72	5	359	2,600,704	12,750,040	7340	15,351,108
Indonesia (ID)	40	2683	2539	28,439	2,669,174	795	2,702,835
Honduras (HN)	23	1064	1059	37,529	1,126,426	3104	1,166,078
Ecuador (EC)	87	973	0	23,456	675,811	1019	700,240
China (CN)	176	4251	4691	294,616	13,964,665	31,119	14,268,223
Saudi Arabia (SA)	649	8650	6813	225,499	32,047,515	5968	32,288,477
Egypt & UAE (EU)	570	5187	5017	706,047	8,764,523	7060	9,480,774
Uganda (UG)	3	5	0	62,857	461,219	418	524,081

Table 1. Count of inter- and intra-state interactions by country. Inter-state: source and target nodes associated with *different* state-linked operations; Intra-state: source and target nodes associated with the same state-linked operation; External & isolates: target node is not identified by Twitter as a state-linked actor, or content does not retweet/mention other actors. External interactions and isolates do not inform the analyses provided in this work; the count is provided for context.

RQ2: Strategic and tactical mechanisms

We study the strategic use of network structure and shared content in service to inter-state coordination. These are explored in detail through two case studies—coordination between Cuba and Venezuela and between Russia and Iran. These examples are selected to highlight the diversity of structural and functional activity we observe across the dataset. In the case of Cuba and Venezuela, we observe relatively bi-directional interaction; both countries serve as source and target of coordinated activity. We also observe administrators playing distinct roles in the campaign. Russia and Iran's coordinated operations, on the other hand, are at a larger scale and structurally very different.

Ambient affiliation. Implicit association among social network actors is facilitated through hashtagging, a phenomenon which has been studied in the sociolinguistics literature as *ambient affiliation*[9]. These indirect interactions enhance visibility of users' discourse through search[49]. The social role of hashtagging is to facilitate the establishment of ad hoc social interaction groupings or subcommunities, which constitute a temporal habitat for information operations. Hashtagging has been employed and proven effective across platforms, from "influencers" and organizations to disinformation operations, for acquiring followers and increasing exposure[1,50–52]. We suggest that inter-state ambient affiliation is used by information operations to create idea habitats conducive to information spread.

We construct the inter-state hashtag network (Fig. 3a). Nodes in the network represent accounts that both engage in hashtagging and are involved in inter-state coordination; Undirected edges indicate use of common hashtags by these accounts. There are 1014 nodes and 61, 010 edges in the network (density = 0.594; diameter = 11). The largest connect component contains 861 nodes and 60, 510 edges (density = 0.0817). Use of ambient affiliation is variable across operations (Fig. 3b). While pervasive within Cuban operations, Russia, Serbia, and Turkey use this tactic only negligibly despite having large-scale operations. Notably, engagement in ambient affiliation appears correlated with greater engagement (see Table 2, e.g., we see greater engagement with content from Cuba and Honduras than from Russia and Serbia).

Globally, we identify 1148 unique hashtags that occur within inter-state activity a total of 33, 119 times. Figure 4 lists and categorized the hashtags which appear in at least 200 inter-state interactions. We observe pervasive, intentional exploitation of political controversy within inter-state hashtagging behavior. In the case of several prominent collaborative operations, a majority of inter-state interactions target specific political events (e.g. Honduras and Iran: 2017 Honduras' Election Crisis; Iran and Russia: 2016 US election; Iran and Venezuela: 2017 Venezuelan Protests). Other coordination activities incorporate media outlets associated mostly with unsubstantiated news.

In addition, we observe the use of hashtagging for network expansion, e.g., #syts, #openfollow, #siguemeytesigo (follow me and I'll follow you). This tactic appears to take one of two forms: (1) explicitly requesting followers; and (2) using mentions and tags for penetration into new communities.

Taxonomizing roles within inter-state operations. Within inter-state operations, we observe that different users/accounts appear to have different patterns of behavior. We contextualize these differences through the lens of role analysis, defining primary roles as follows.

- **Administrator** Manages operations of the information campaign. Administrators self-identify as group leaders through profile information and shared content.
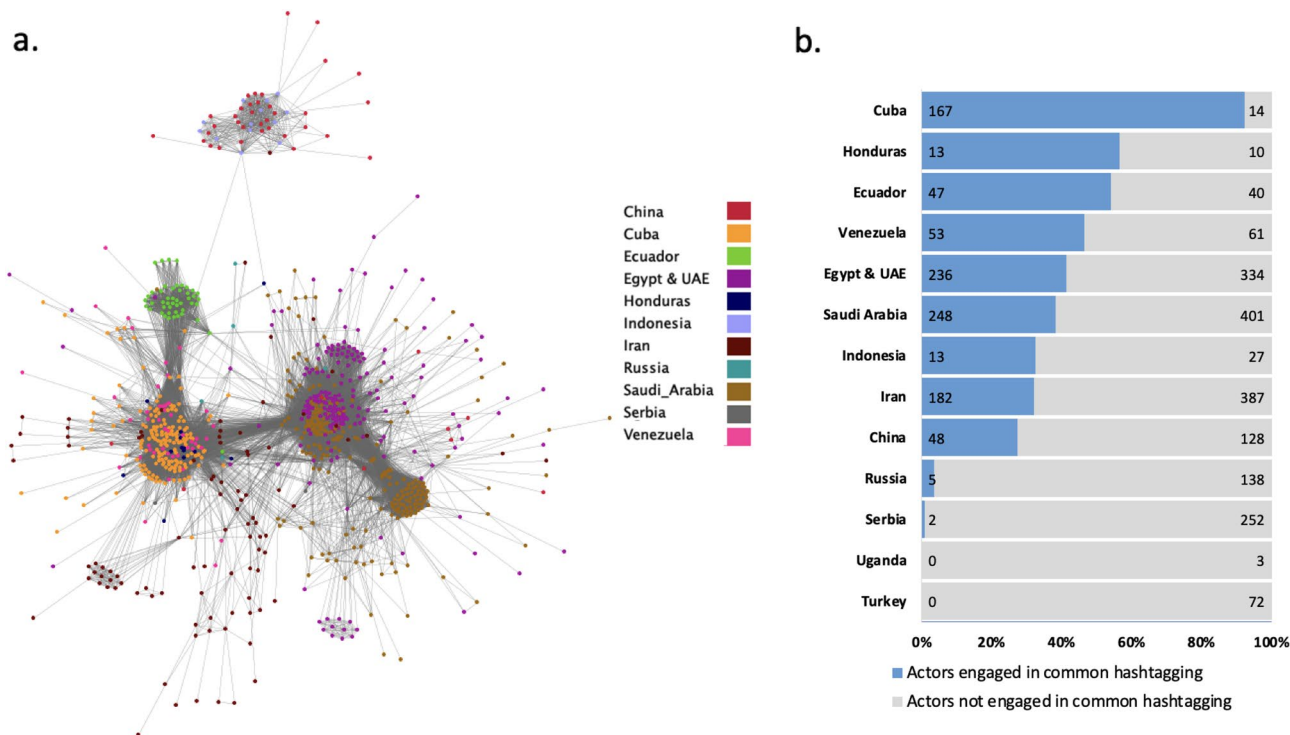
Figure 3. (**a**) Inter-state hashtag network; Components of size ≤ 20 are suppressed for clarity. (**b**) Actor engagement within inter-state hashtag network, by country.

Country	Avg like	p	T	d	Avg retweet	p	T	d	Avg reply	p	T	d
VE (Inter)	0.401	**0.000**	9.218	0.566**	0.440	**0.000**	4.737	0.096	0.128	**0.000**	6.038	0.478*
VE (Intra)	0.039				0.168				0.015			
CU (Inter)	0.539	0.169	1.375	0.057	0.366	0.794	0.261	0.011	0.125	**0.000**	8.420	0.314*
CU (Intra)	0.351				0.332				0.020			
CN (Inter)	0.035	**0.000**	4.438	0.063	0.058	**0.000**	6.608	0.027	0.351	**0.000**	31.720	0.902***
CN (Intra)	0.016				0.016				0.025			
IR (Inter)	0.056	**0.000**	−6.142	−0.034	0.052	0.668	0.429	0.007	0.101	**0.000**	6.937	0.304*
IR (Intra)	0.159				0.044				0.018			
RS (Inter)	0.000	**0.000**	−125.492	−0.053	0.000	**0.000**	−97.538	−0.041	0.000	**0.000**	−192.692	−0.082
RS (Intra)	0.012				0.008				0.012			
RU (Inter)	0.007	**0.000**	−4.570	−0.010	0.000	**0.000**	−10.009	−0.017	0.018	**0.000**	−1.938	−0.084
RU (Intra)	0.034				0.021				0.038			
ID (Inter)	0.034	**0.003**	2.981	0.040	0.072	**0.000**	6.476	0.111	0.559	**0.000**	33.782	1.425***
ID (Intra)	0.020				0.019				0.067			
SA (Inter)	0.037	**0.039**	2.064	0.012	0.068	0.138	1.482	0.007	0.034	**0.000**	−4.935	−0.027
SA (Intra)	0.026				0.039				0.050			
EC (Inter)	0.020	0.995	−0.007	0.000	0.002	0.040	−2.059	−0.035	0.000	**0.000**	−6.803	−0.046
EC (Intra)	0.020				0.007				0.018			
HN (Inter)	0.450	0.611	0.509	0.004	0.619	**0.000**	5.879	0.148	0.030	0.212	−1.250	−0.016
HN (Intra)	0.416				0.182				0.039			
EU (Inter)	0.073	**0.000**	6.683	0.100	0.076	**0.000**	5.774	0.114	0.036	**0.000**	8.463	0.148
EU (Intra)	0.012				0.008				0.012			

Table 2. Comparison of inter- and intra-state engagement. p: p-value; T: T-score; d: Cohen's d value; *: small effect size, **: medium effect size, ***: large effect size; p-values in bold indicate significance after FDR correction.
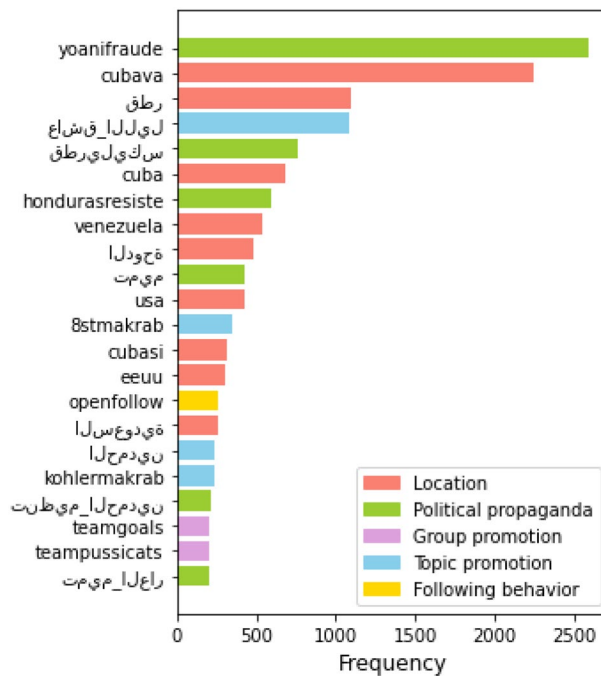
Figure 4. Top hashtags used in inter-state IO. (English translations, from top: Yoani fraud, Cuba goes, Qatar, love the night, Qatarileaks, Cuba, Honduras resist, Venezuela, Doha, Tamim, USA, 8stmakrab*, Cuba yes, USA, openfollow, Saudi Arabia, Al-Hamdan, Kohlermakrab*, organization_Hamdeen, Team Goals, Team Pussicats, Tamim_shame; *Note: Makrab is the abbreviation of the phrase "Malam Keakraban" in Indonesian which means friendship night).

- **Influencer** A hub of the operation with high in-degree centrality. Typically, an influencer is the source of the information who exploits fake news sites and may have multiple similar accounts in the network to avoid takedown. We define users with in-degree greater than 10 as the influencers in the network.
- **Promoter** Primarily promotes content for enhanced visibility and engagement. We define users with out-degree greater than 10 as promoters in the network.
- **Broker** A gatekeeper, connecting multiple communities/organizations with relatively high in-degree and out-degree centrality. We identify users who meet criteria for both promoter and influencer as brokers.
- **Follower** An actor with minor (observable) impact within the operation. Users that are not identified within aforementioned roles are categorized as followers.

We leverage this taxonomy in the case studies which follow. The two case studies are selected for their diversity with respect to structure and content.

Case Study 1: Coordination between state-linked accounts from Cuba and Venezuela.
Network structure. We construct the inter-state interaction network between Cuban and Venezuelan state-linked accounts. We observe 6469 interactions between 56 Cuban and 62 Venezuelan accounts. Notably, the network has two well-connected clusters connected by a single edge (see Fig. 5a). We observe a relatively balanced network structure between Cuba and Venezuela where the interactive pattern is bidirectional (5464/6649 of Cuba's out-degree interactions point to Venezuela, and 1005/1027 Venezuela's out-degree interactions point to Cuba). In each cluster, a small subset of Cuban and Venezuelan nodes dominate activity while remaining nodes connect with them through retweets and mentions. This network structure is a trademark of Cuba's inter-state operations. We observe that Cuba's intra-state interaction network, by contrast, has greater connectivity and more uniform in-/out-degree sequences.

Content analysis. Notably, Cuban and Venezuelan information operations appear to center around structured *teams*, and each team has a self-identified administrator.

Cluster 1. In Fig. 5b, we dive deeply into the inter-state operation structure. Within Cluster 1, we identify five representative nodes, two attributed to the Venezuelan campaign and three from Cuba. Representative nodes are selected as those with the most within-cluster interactions. The two team leaders are also identified though their user profiles and account descriptions. V1 is a Venezuelan node with high out-degree centrality, and the two Cuban nodes with which V1 frequently interacts are designated C1 and C2.

Actors C1 and C2 are administrators in TeamGoal and TeamPussicats, respectively. The majority of unilateral interactions from V1 to C1 and C2 take the form of direct retweets and retweets from others that mention

Figure 5. (**a**) Inter-state interaction network between Cuba and Venezuela; (**b**) schematic structure of inter-state disinformation operations between Cuba and Venezuela (usernames are hashed for accounts with fewer than 5000 followers).

C1 or C2. Manual content analyses suggest that the primary objective of these two Cuban actors is to promote their team and its members through establishing unique sets of emojis and hashtags that symbolize their team identities and consistent mentioning of the team leaders. By retweeting C1 and C2 as well as other members of their teams, V1 serves as a promoter of their content. Venezuelan node V2 has bilateral coordination with both C1 and C2. The majority of V2's activities are replies to users or tweets with direct mentions. These two sets of tweets focus mostly on promoting members of teams including C1 and C2 and others, e.g., GhostBand Team, Orgasmas Team, Incognitos Team, CódiceRasta Team, ElBúnker, and Team Dioses_Míticos. V2 acts as a broker between the aforementioned groups, facilitating communication and collaboration between the teams and the team members.

Cluster 2.   Within Cluster 2, we identify node V3 as an influencer. V3 primarily distributes fake news via the use of URLs and hashtags that receive considerable engagement (65.26% of total tweets from Cuba are connected to V3). We note that numerous accounts that closely resemble this suspended account still exist in the current social network to avoid being taken down by Twitter, as shown in Cluster 2 of Fig. 5b. V3 is linked to C4, C5, and several other Cuban promoter accounts, through both explicit (retweeting) and implicit interactions (common hashtags).

Role analysis of accounts within the inter-state Cuban-Venezuelan operations (see "Supplementary Material") suggests that, broadly speaking, Venezuelan accounts act as influencers, while Cuban accounts primarily promote content shared by Venezuelan accounts.

URL analysis. We collect account profile information and tweets of all accounts engaged in inter-state operations between Cuba and Venezuela. In total, there are 16 accounts (13.56% of actors from both countries) whose profiles contain URLs. Further, we identify 1913 unique URLs within their tweets, occurring over 2553 interactions (39.47% of total interactions) (see "Supplementary Material"). Invalid URLs that include broken links and cannot be manually identified by name are removed. Then, we manually verify the status of each link, categorizing each as active if the link is still functioning and inactive if the content has been removed or the account has been set to private. A substantial number of URLs are invalid (70.10%), indicating that most were temporary. We observe that the majority of valid URLs in profiles redirect to accounts on other social networking sites like Facebook, Instagram, and YouTube (9.27% of total valid URLs, 37.74% of which are still active), and blogs with little or no regulations (28.85% of total valid URLs, 100% of which are still active). URLs within tweets often direct to politically-oriented news (57.00% of total valid URLs), e.g., Telesurtv. Some also point to social media content management applications such as Twitlonger to bypass the character limit and Twitpic for picture archiving (still accessible after the account has been taken down), as well as to manage followers from social media, likely for open follow practices, e.g., Tuitil.

Case study 2: Coordination between state-linked accounts from Russia and Iran. *Dynamic network structure.* We construct a dynamic view of inter-state operations between Russia and Iran through four interaction networks, see Fig. 6a. We aggregate all interactions prior to 2016 and after 2018, since the majority of interactions occur between 2016 and 2017; 2016 and 2017 are represented as snapshots. A total of 54 Russian accounts and 329 Iran accounts are represented in the network. We observe several clusters that contain one Russian node and multiple Iranian nodes, in a radiating pattern. The radiating network topology is well-suited for executing strategic aims such as news media distribution and social network expansion.

Russian and Iranian accounts are dynamically involved in inter-state coordination globally (see "Supplementary Material"). These statistics are calculated over year-long network snapshots, beginning in 2011. We observe temporal uniformity for Iranian inter-state information operations; specifically, accounts reach out-degree interaction peak at around the same time (see "Supplementary Material"). This may suggest that inter-state operations were/are conducted concurrently. Although Iran has more incoming than outgoing interactions (see Table 1), Iran also exhibits particularly many outgoing interactions with Russia. Upon closer inspection, we
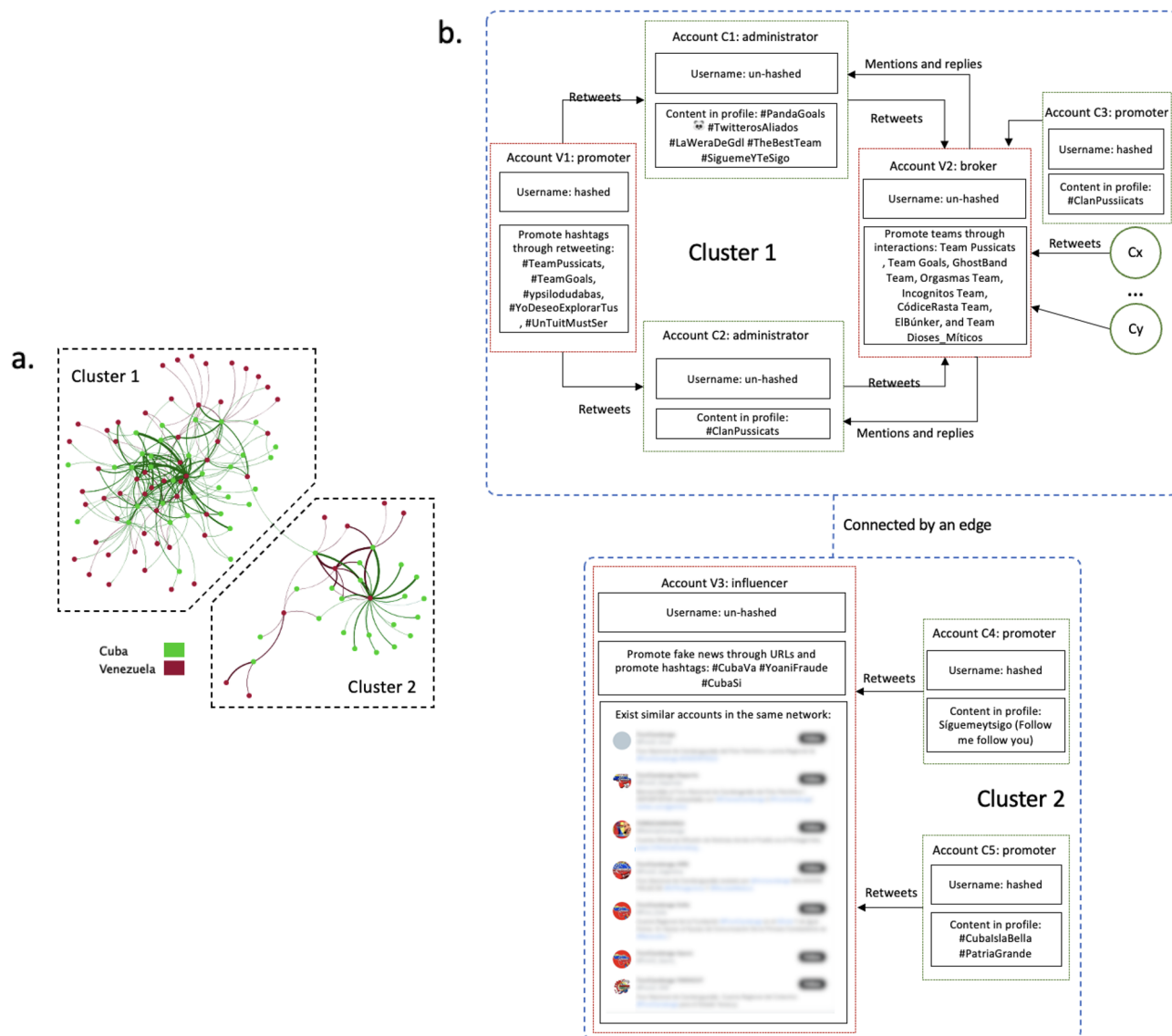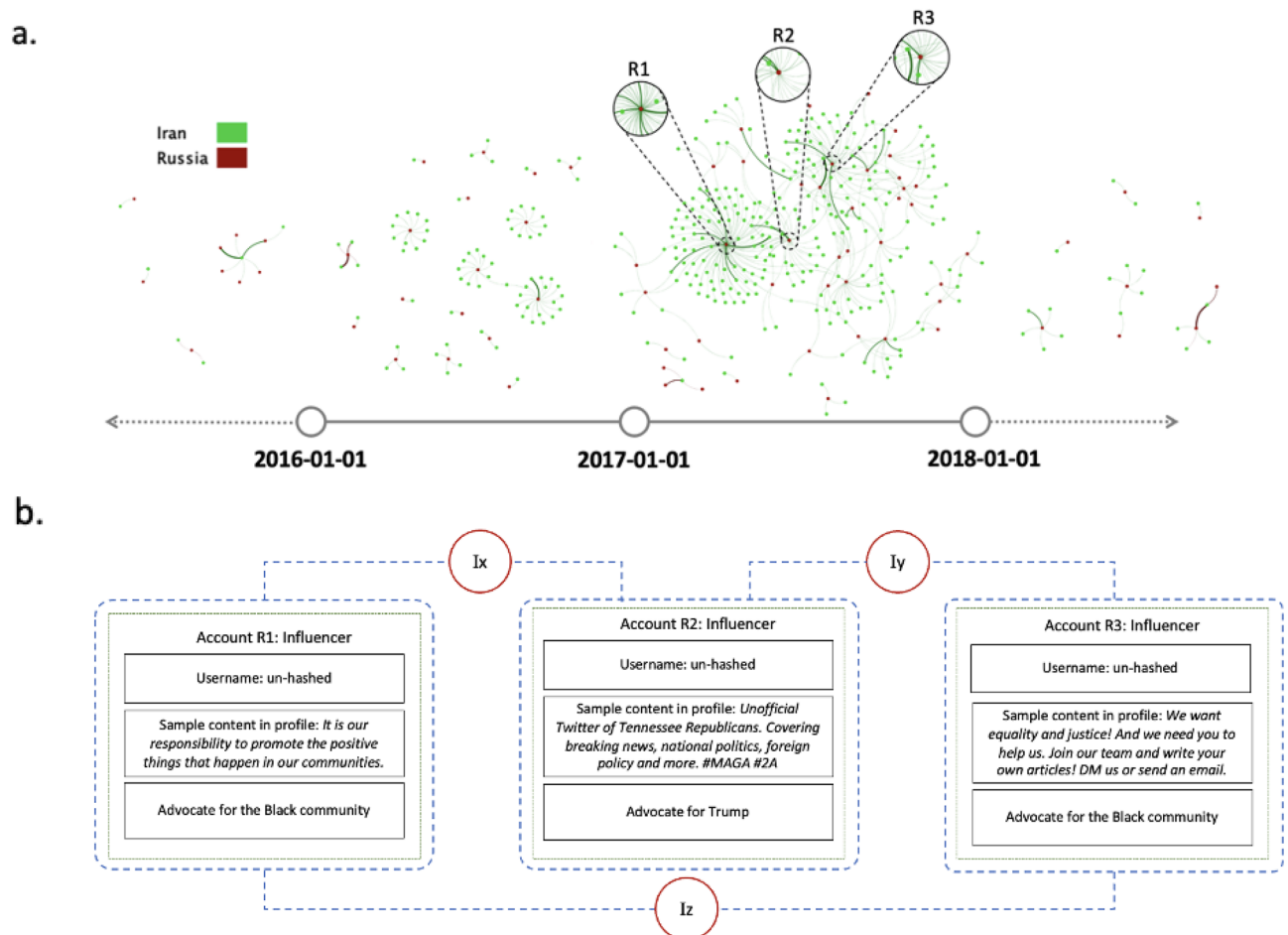

Figure 6. (**a**) Dynamic inter-state interaction networks between Russia and Iran; (**b**) Schematic structure of inter-state information operations between Russia and Iran.

observe that the reason behind Iran's high in-degree centrality is that several central news outlets, e.g., Iranian state-controlled Hispantv, is frequently and consistently retweeted by accounts linked to Cuba, Venezuela, and Honduras (99.01%, 99.58%, and 99.7% of the total tweets, respectively). Thus, Iran may be passively engaged in those interactions. However, in its coordination with Russia, Iranian actors actively disseminate content created by Russian actors, e.g., targeting the US 2016 election.

Content analysis. Figure 6b explores the primary structure of inter-state operations between Russia and Iran. We observe that Iranian accounts actively initiate interaction with Russian actors; approximately 96.81% of activity in the inter-state interaction network are retweets of Russian accounts, the majority of which are affiliated with the Internet Research Agency (IRA). Manual analyses of tweets and profile descriptions indicate that content is primarily focused on political and politicized topics. This is in line with prior work examining Russian information operations targeting the 2016 US election[53,54]. For instance, the profile description of the central node R2 is:

> Unofficial Twitter of Tennessee Republicans. Covering breaking news, national politics, foreign policy and more. #MAGA #2A

In addition to US election politics, inter-state activity also focuses on racial issues (e.g., R1 and R3). Iranian actors participate in these processes by disseminating material previously posted by Russian actors with high centrality , and connecting all three satellite communities.

As in Case Study 1, we identify account roles in the Russia-Iran inter-state campaign (see "Supplementary Material"). Russia and Iran differ significantly in terms of number of influencers and followers; Russian actors serve primarily as source content for the Iranian community.

URL analysis. Also as in Case Study 1, we collect profile information and tweets from all accounts engaged in inter-state operations between Russia and Iran. In total, there are just 7 accounts (1.83% of actors from both countries) that contain URLs within their profile descriptions. We identify 35 unique URLs within shared tweets over 77 total interactions (15.28% of total interactions). We categorize these URLs by type/platform and active status (see "Supplementary Material"). Among valid URLs, we see extensive reference to international political news from the United States, the United Kingdom, and Iraq (31.43%). While the majority of linked social media accounts are now defunct (75% for profile URLs and 100% for content URLs) URLs to most news outlets remain active and complement tweet content (68.75% for content URLs).

Discussion and conclusions

Rapid and accurate detection of information operations remains a hard problem for social media platforms globally. Evidence that state-linked operations may collaborate or collude to improve the efficacy of their campaigns adds complexity to that challenge. Our work provides such evidence, and therefore informs ongoing efforts to detect and mitigate the impact of information operations deployed through social media. We have highlighted some recurring strategies and tactics employed by inter-state information operations on Twitter. We have observed that a substantial subset of coordinated inter-state activity can be identified as supportive of explicit aims, e.g., targeting high-stakes political events or seeking additional visibility. Regardless of motivation, it appears that inter-state activities are carried out separately from intra-state operations, resulting in a distinctive information ecosystem, or idea habitat. Relatedly, we discover that a majority of inter-state operations exploit ambient affiliation through hashtagging, and that individual accounts in the network may be tasked with distinct roles in some operations architectures. Overarchingly, our findings suggest that information operations represent collaborative work, not only at the individual level but also at the state level. Notably, our analyses also reveal that country size is not necessarily a determinant of the scale of observed inter-state activity. Smaller countries demonstrate the ability to engage in systematic coordination, strategically expanding their internal operations.

The scope of the current work is constrained in several ways. Our analyses make assumptions about the accuracy of identified accounts and about their activity, e.g., that all account activity serves ongoing operations. We assume that substantial observed interaction indicates strategic coordination. Ultimately, the emergence of inter-state coordination and the ways in which observed activities are moderated through planning remains an open question. Furthermore, the extent to which current insights can be leveraged for the advancement of automated approaches for detection of inter-state information operations will likely depend on the uniqueness of inter-state interaction patterns vs., e.g., standard content promotion strategies employed by traditional organizational accounts seeking visibility and influence. Future work would benefit from designing studies to compare these phenomena.

The inter-state activity we uncover here is likely a very small segment of much larger-scale, dynamic, cross-platform, multi-media, partially-observable coordinated operations. Our findings raise many more questions regarding the offline interactions which underlie observed activity, the ways in which they are facilitated, and the broader political agenda which they serve. The answers to these questions will be distinct across countries and over time. Our work therefore highlights the critical role of policy and international relations in this space. It also suggests that whether and how states cooperate to respond to information operations on social media will require a transdisciplinary research and policy agenda bringing together computational and social scientists, policy makers, and stakeholders.

Data availability

The datasets described herein are readily available and publicly accessible through Twitter's Moderation Research Consortium at the following https://transparency.twitter.com/en/reports/moderation-research.html. Datasets can be downloaded instantly upon providing a valid email address.

References

1. Starbird, K., Arif, A. & Wilson, T. Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. *Proc. ACM Hum.–Comput. Interact.* **3**, 1–26 (2019).
2. Pierri, F., Artoni, A. & Ceri, S. Investigating Italian disinformation spreading on twitter in the context of 2019 European elections. *PloS One* **15**, e0227821 (2020).
3. Soares, F. B. & Recuero, R. Hashtag wars: Political disinformation and discursive struggles on twitter conversations during the 2018 Brazilian presidential campaign. *Social Media + Society.* **7**, 20563051211009070 (2021).
4. Xia, Y. *et al.* Disinformation, performed: Self-presentation of a Russian IRA account on twitter. *Inform. Commun. Society* **22**, 1646–1664 (2019).
5. Casey, R. D. The Goebbels experiment: A study of the Nazi propaganda machine (1944).
6. Berelson, B. & De Grazia, S. Detecting collaboration in propaganda. *Public Opin. Quart.* **11**, 244–253 (1947).
7. Waddington, L. L. The anti-komintern and nazi anti-bolshevik propaganda in the 1930s. *J. Contemp. History* **42**, 573–594 (2007).
8. Cabrera, M. G. International propaganda in Spain during the first world war: State of the art and new contributions. *Commun. First World War*. 188–218 (2020).
9. Zappavigna, M. Enacting identity in microblogging through ambient affiliation. *Discourse Commun.* **8**, 209–228 (2014).
10. Zhang, Y. *et al.* Assembling the networks and audiences of disinformation: How successful Russian IRA twitter accounts built their followings, 2015–2017. *J. Commun.* **71**, 305–331 (2021).
11. García-Perdomo, V., Salaverría, R., Brown, D. K. & Harlow, S. To share or not to share: The influence of news values and topics on popular social media content in the United States, Brazil, and Argentina. *J. Stud.* **19**, 1180–1201 (2018).
12. Posetti, J. & Matthews, A. A short guide to the history of 'fake news' and disinformation. *Int. Center J.* **7**, 2018-07 (2018).
13. Rajtmajer, S. & Susser, D. Automated influence and the challenge of cognitive security. in *Proceedings of the 7th Symposium on Hot Topics in the Science of Security*, 1–9 (2020).
14. Niblock, I., Wallis, J. & Zhang, A. Understanding global disinformation and information operations (2022).
15. Bradshaw, S. & Howard, P. N. The global disinformation order: 2019 global inventory of organised social media manipulation (2019).
16. Bradshaw, S. & Howard, P. N. Challenging truth and trust: A global inventory of organized social media manipulation. *Comput. Propaganda Project* **1**, 1–26 (2018).
17. Arnaudo, D. Computational propaganda in Brazil: Social bots during elections (2017).
18. Ong, J. C. & Cabañes, J. V. A. Architects of networked disinformation: Behind the scenes of troll accounts and fake news production in the philippines. in *Architects of networked disinformation: Behind the scenes of troll accounts and fake news production in the Philippines* (2018).
19. Vo, N. & Lee, K. The rise of guardians: Fact-checking URL recommendation to combat fake news. in *The 41st international ACM SIGIR conference on research & development in information retrieval*, 275–284 (2018).
20. Sathe, A., Ather, S., Le, T. M., Perry, N. & Park, J. Automated fact-checking of claims from wikipedia. in *Proceedings of the 12th Language Resources and Evaluation Conference*, 6874–6882 (2020).
21. Volkova, S., Ayton, E., Arendt, D. L., Huang, Z. & Hutchinson, B. Explaining multimodal deceptive news prediction models. *Proc. Int. AAAI Conf. Web Social Media* **13**, 659–662 (2019).
22. Garimella, K. & Eckles, D. Images and misinformation in political groups: Evidence from Whatsapp in India. *arXiv preprint.* arXiv:2005.09784 (2020).
23. Krishnamurthy, G., Majumder, N., Poria, S. & Cambria, E. A deep learning approach for multimodal deception detection. *arXiv preprint* arXiv:1803.00344 *(2018)*.
24. Dawson, A. & Innes, M. How Russia's internet research agency built its disinformation campaign. *Polit. Quart.* **90**, 245–256 (2019).
25. Myers, S. L. & Mozur, P. China is waging a disinformation war against Hong Kong protesters. *N. Y. Times.* **13** (2019).
26. Merhi, M., Rajtmajer, S. & Lee, D. Information operations in turkey: Manufacturing resilience with free twitter accounts. *arXiv preprint* arXiv:2110.08976 *(2021)*.
27. Ferrara, E. Disinformation and social bot operations in the run up to the 2017 french presidential election. *arXiv preprint* arXiv:1707.00086 *(2017)*.
28. Faris, R. *et al.* Partisanship, Propaganda, and Disinformation: Online Media and The 2016 US Presidential Election Vol. 6 (Berkman Klein Center Research Publication, 2017).
29. Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B. & Lazer, D. Fake news on twitter during the 2016 US presidential election. *Science* **363**, 374–378 (2019).
30. Hutchins, E. Cognition in the wild MIT press. *Cambridge, MA* **15** (1995).
31. Sarcevic, A. *et al.* "beacons of hope" in decentralized coordination: Learning from on-the-ground medical twitterers during the 2010 haiti earthquake. in *Proceedings of the ACM 2012 conference on computer supported cooperative work*, 47–56 (2012).
32. Schoch, D., Keller, F. B., Stier, S. & Yang, J. Coordination patterns reveal online political astroturfing across the world. *Sci. Rep.* **12**, 1–10 (2022).
33. Del Vicario, M. *et al.* The spreading of misinformation online. *Proc. Natl. Acad. Sci.* **113**, 554–559 (2016).
34. Wang, R., He, Y., Xu, J. & Zhang, H. Fake news or bad news? Toward an emotion-driven cognitive dissonance model of misinformation diffusion. *Asian J. Commun.* **30**, 317–342 (2020).
35. Chen, X., Sin, S.-C.J., Theng, Y.-L. & Lee, C. S. Deterring the spread of misinformation on social network sites: A social cognitive theory-guided intervention. *Proc. Assoc. Inform. Sci. Technol.* **52**, 1–4 (2015).
36. Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G. & Rand, D. G. Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychol. Sci.* **31**, 770–780 (2020).
37. Pennycook, G. & Rand, D. G. The psychology of fake news. *Trends Cognit. Sci.* **25**, 388–402 (2021).
38. Berger, J. A. & Heath, C. Idea habitats: How the prevalence of environmental cues influences the success of ideas. *Cognit. Sci.* **29**, 195–221 (2005).
39. Penney, J. *The Citizen Marketer: Promoting Political Opinion in the Social Media Age* (Oxford University Press, 2017).
40. Jakubowski, G. What's not to like? Social media as information operations force multiplier. *Joint Force Quart.* **3**, 8–17 (2019).
41. Lee, A. J., Yang, F.-C., Tsai, H.-C. & Lai, Y.-Y. Discovering content-based behavioral roles in social networks. *Decis. Supp. Syst.* **59**, 250–261 (2014).

42. Han, S., Zhuang, F., He, Q. & Shi, Z. Balanced seed selection for budgeted influence maximization in social networks. in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 65–77 (Springer, 2014).
43. Lorrain, F. & White, H. C. Structural equivalence of individuals in social networks. *J. Math. Sociol.* **1**, 49–80 (1971).
44. Forestier, M., Stavrianou, A., Velcin, J. & Zighed, D. A. Roles in social networks: Methodologies and research issues. *Web Intell. Agent Syst. Int. J.* **10**, 117–133 (2012).
45. Rundin, K. & Colliander, J. Multifaceted influencers: Toward a new typology for influencer roles in advertising. *J. Advertising* **50**, 548–564 (2021).
46. Twitter moderation research consortium. https://transparency.twitter.com/en/reports/moderation-research.html. Accessed: 2023-03-20.
47. Disclosing new data to our archive of information operations. https://blog.twitter.com/en_us/topics/company/2019/info-ops-disclosure-data-september-2019. Accessed: 2023-03-20.
48. Squartini, T., Picciolo, F., Ruzzenenti, F. & Garlaschelli, D. Reciprocity of weighted networks. *Sci. Rep.* **3**, 2729 (2013).
49. Zappavigna, M. Ambient affiliation: A linguistic perspective on twitter. *New Media Society* **13**, 788–806 (2011).
50. Erz, A., Marder, B. & Osadchaya, E. Hashtags: Motivational drivers, their use, and differences between influencers and followers. *Comput. Hum. Behav.* **89**, 48–60 (2018).
51. Conway, B. A., Kenski, K. & Wang, D. Twitter use by presidential primary candidates during the 2012 campaign. *Am. Behav. Scientist* **57**, 1596–1610 (2013).
52. Martín, E. G., Lavesson, N. & Doroud, M. Hashtags and followers: An experimental study of the online social network twitter. *Social Netw. Anal. Mining* **6**, 1–15 (2016).
53. Benkler, Y., Faris, R. & Roberts, H. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (Oxford University Press, 2018).
54. Berghel, H. Oh, what a tangled web: Russian hacking, fake news, and the 2016 us presidential election. *Computer* **50**, 87–91 (2017).

Acknowledgements

Author contributions

X.W. conceived and conducted experiments, analyzed results, and drafted the paper, J.L. conducted experiments, analyzed results, and contributed to writing the paper, E.S. conducted experiments and analyzed results, S.R. conceived research questions and experiments, and contributed to writing the paper. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-34245-1.

Correspondence and requests for materials should be addressed to X.W. or S.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.