Computational Analysis of Bot Activity in the Asia-Pacific: A Comparative Study of Four National Elections

Joshua Uyheng and Kathleen M. Carley
CASOS Center, Institute for Software Research, Carnegie Mellon University
{juyheng, kathleen.carley}@cs.cmu.edu

Abstract

Bot-driven electoral disinformation represents a major threat to democracies worldwide. Extant scholarship, however, tends to concentrate around Western contexts. This paper undertakes a comparative computational analysis of bot activity during four recent elections in the Asia-Pacific. Through a systematic, multi-level comparison of bot activity, we contribute novel insights about shared and distinct computational features of the disinformation landscapes within a significant yet understudied geopolitical region. Across case studies in Indonesia, the Philippines, Singapore, and Taiwan, we find non-negligible levels of bot activity: bots engage in higher levels of tweet production; interact with humans especially through mentions; tend to occupy denser and more isolated communities; use simpler and abusive language; and share partisan, irrelevant, or conspiratorial content. We conclude with implications for deepening and utilizing the analysis presented here as well as future directions for further cross-national work.

Introduction

Extensive attention has been accorded to how inauthentic agents - often referred to as bots - manipulate public discourse on social media (Bessi and Ferrara 2016; Shao et al. 2018). Implicated in a wide range of major influence campaigns, bots have been studied on multiple levels, including: their idiosyncratic account features and interactional behaviors, their effect on public narratives and community structures, and the overall quality of information that is propagated through cyberspace (Alizadeh et al. 2020; Grinberg et al. 2019). Growing literature on these topics has accumulated across multiple disciplines, with vital implications for methods and policies for the detection and mitigation of bots and their impacts (Tan 2020; Varol et al. 2017).

Many studies, however, have tended to focus on Western contexts, especially in the US and Western Europe (Bastos and Mercea 2019; Nizzoli et al. 2020). While the global reach of the field has expanded in recent years (Cresci 2020), open empirical questions continue to concern the extent to which various stylized facts about bot activity are applicable to different geopolitical contexts. More broadly, while tool creation and in-depth analysis of particular case studies are well-developed, comparative work that explicitly examines converging and diverging features of bot activity across multiple contexts is relatively scant (Miller and Vaccari 2020; Schuchard et al. 2018). This paper designs and implements a framework for bridging these gaps in the literature in the Asia-Pacific, a diverse geopolitical region with known cases of digital disinformation (Kaur et al. 2018). Broadly, we ask the following research questions:

1. How can we computationally characterize bot activity during elections in the Asia-Pacific?
2. What features of bot behavior are similar or dissimilar between the contexts under study?

The succeeding sections of this paper are organized as follows. First, we offer a brief overview of related work probing digital disinformation during elections in general, as well as the particular role of social bots in enacting disinformation maneuvers through coordinated behavior. We posit the significance and need for comparative computational analysis especially in the Asia-Pacific. Next, we detail the core aims of this work, focusing on distinct analytical levels in order to capture different facets of bot activity and identify holistic contributions to the literature. We then present the datasets and tools we use to facilitate systematic empirical comparisons between online electoral conversations. Finally, we present our analysis of bot prevalence and interaction patterns; their distinct linguistic, structural, and link-sharing behaviors; and discuss the implications of these findings for the broader computational social science of disinformation worldwide.

Related Work

Rich scholarship establishes the importance of examining the activities of bots in propagating disinformation and disrupting democracy. Major elections like the 2016 US Presidential race featured notable proliferation and consumption of fake news, with quantifiable impacts on subsequent online behavior (Bovet and Makse 2019; Grinberg et al. 2019). This spread of low-credibility information has been linked to coordinated influence operations by inauthentic actors, triggering concerns over the security of the digital sphere and state-sponsored cyber warfare (Shao et al. 2018;

Zannettou et al. 2019). Beyond the US, bot activities have been documented in influencing hyperpartisan discourse in relation to Brexit in the UK (Bastos and Mercea 2019; Nizzoli et al. 2020), obfuscating political discourse in the 2017 German elections (Keller and Klinger 2019), and sowing discord in large-scale international events in Western Europe (Uyheng et al. 2019).

Collectively, these findings have suggested the importance of formulating capable responses to detect, monitor, understand, mitigate, and prevent coordinated information operations (Carley 2020). Research efforts built on these motivations have been multifaceted and multidisciplinary (Miller and Vaccari 2020; Shu et al. 2017). By applying computational tools grounded in social scientific theory, researchers in this area have gleaned insights into key behavioral differences between bots and humans (Alizadeh et al. 2020; Gilani et al. 2019). Beyond analyzing aberrant behaviors by individual accounts, more recent work highlights the value of taking a group-based perspective to find coordinated actors (Cresci et al. 2017; Mendoza, Tesconi, and Cresci 2020; Pacheco et al. 2020). These findings inform the development of further techniques which facilitate rapid responses to emergent influence campaigns (Beskow and Carley 2018; Luceri, Giordano, and Ferrara 2020; Yang et al. 2020).

However, outside the US and Western Europe, similar computational evidence remains in its relatively nascent stages. Certainly, in recent years, analysis has steadily grown in relation to information operations in South America, Africa, and the Asia-Pacific region (Ndlela 2020; Recuero, Soares, and Gruzd 2020; Uyheng and Carley 2019). Outside the West, field studies grounded in local contexts have valuably exposed the inner workings of domestic and foreign disinformation efforts, with rich insights into the organizational makeup of content farms, idiosyncratic campaign styles in varied political settings, and the culturally distinct practices of fake news consumption among diverse publics (Das and Schroeder 2020; Hopkins 2014; Kaur et al. 2018; Ong, Tapsell, and Curato 2019). Yet less is generally known about how such operations play out in their actual digital settings. Some prospects for computational analysis are laid out in relation to the unique regulatory landscapes of Asia-Pacific countries (Cha, Gao, and Li 2020; Tan 2020). Implemented empirical investigations, however, appear much more scant. During the 2019 Indian elections, extensive analysis is conducted into the online dissemination of political content, but bot activity does not explicitly feature in this analysis (Agarwal et al. 2020). In Sri Lanka, some bot activity is linked to the 2015 elections, but studied only in relation to a small dataset of about 2000 users (Rathnayake and Buente 2017). Bot-linked campaigning was also found in the 2014 Japanese elections, but researchers primarily focused on accounts which sent duplicate tweets, without consideration for other features and types of bot activity (Schäfer, Evert, and Heinrich 2017)

This work thus aims to build on burgeoning efforts in the Asia-Pacific through comparative computational analysis. On the one hand, this has the benefit of shedding light on the features of bot activity in multiple understudied settings at once (Uyheng and Carley 2020a). On the other hand, we also respond to more fundamental calls for comparative computational research on electoral disinformation (Miller and Vaccari 2020). Much of the significant foregoing work focused on in-depth investigations into a single event. Explicitly comparative approaches facilitate complementary insights into which inauthentic behaviors - as well as subsequent reponses to them - may be shared or anchored in more locally specific contingencies (Humprecht, Esser, and Van Aelst 2020). While comparative work has certainly been undertaken in this area, to the best of our knowledge, they have largely been framed in terms of Twitter conversation in general instead of electoral discourse (Gilani et al. 2019); or focused again on mainly Euro-American settings (Schuchard et al. 2018). Conceptually, existing studies were likewise focused on relatively narrow measurement objectives such as network centrality or tweet popularity, which we expand on here.

Contributions of This Work

In view of the foregoing literature, this paper stakes the following contributions to the literature:

1. Systematic computational analysis of bot activity in recent elections in Indonesia, the Philippines, Singapore, and Taiwan extends the predominantly Western literature and augments valuable insights from field work and policy analysis within the Asia-Pacific region.
2. A comparative framework empirically examines the consistencies and contingencies of various features of bot activity in multiple electoral contexts, including patterns of tweet production, user interaction, abusive language, and low-quality information-sharing.
3. A general methodology using machine learning, network science, and linguistic tools offers a multi-level view of bot-driven disinformation which may aid practical analysis of other online conversations vulnerable to digital disinformation.

Overview of Present Research

To organize the present study, we introduce and employ a multi-level framework. Table 1 summarizes four analytical levels which characterize our approach: agents, interactions, narratives, and networks.

First, an **agent-level analysis** echoes the major, more traditional paradigm in computational disinformation research. Here, our interest is in examining basic questions about the prevalence of bot activity (RQ1). Notwithstanding the weaknesses of individual-based analysis (Cresci 2020; Rauchfleisch and Kaiser 2020), they remain a major first step in identifying emergent online threats based on known properties of social bots based on historical knowledge (Alizadeh et al. 2020; Yang et al. 2019). Contributing findings in this area from the Asia-Pacific thus offers commensurate knowledge with much of the existing work in the West (Luceri, Giordano, and Ferrara 2020; Varol et al. 2017), while recognizing further work needs to be done to address gaps in dominant methodologies to keep up with paradigm shifts in information disorder (Cresci 2020).

Level	Research Questions	Tools
Agents	RQ1a: How prevalent is bot activity in each national election?	Bot Detection
Interactions	RQ2: How do bots interact with other agents in online conversations?	Analysis of Interaction Types
Networks	RQ3: What network structures do bots participate in?	Social Network Analysis
Narratives	RQ4a: What language do bots deploy?	Linguistic Analysis
	RQ4b: What hashtags do bots use?	Hashtag Analysis
	RQ4c: What external links do bots share?	URL Analysis

Table 1: Overview of research questions and hypotheses.

Second, an **interaction-level analysis** considers how the identified social bots engage other accounts in the online conversation (RQ2). The spread of disinformation by bots does not take place as a passive process of message amplification. It involves interactions with other bots as well as human accounts to be successful (Cresci et al. 2017; Starbird 2019). On Twitter, we specifically examine how bots differentially make use of interactive tools like retweets, replies, mentions, and quotes. This may characterize bot objectives, such as artificial accrual of influence through inauthentic interactions with other bots, or disruption of real-world conversations by actively replying to humans.

Finally, we synthesize broader patterns of bot activity in terms of networks and narratives. Going beyond descriptive statistics of bot activity, **network-level analysis** and **narrative-level analysis** correspond to the pillars of information operations (Carley 2020).

We examine networks to probe the structural properties of online electoral discourse in the Asia-Pacific (RQ3). Network structure, which emerges from large-scale patterns in who talks to whom, is significant to both the health and disruption of online conversations by directing information flow (Carley 2020). Bots may manipulate network structure through the coordinated creation of echo chambers or artificial connection of communities to spark conflicts (Uyheng and Carley 2020a). Network-based campaigns have been so central to bot activity that they form the basis for innovative methods of bot detection focusing on coordination activities (Mendoza, Tesconi, and Cresci 2020; Pacheco et al. 2020). Conversely, recent work clarifies that coordination and automation are distinct concepts (Nizzoli et al. 2020), pointing to the need to probe further what community structures feature bots in particular settings.

Meanwhile, narratives are the information that is thematically salient to the online conversation. On the one hand, the language used by bots can be impactful irrespective of the substantive content of discourse (RQ4a). Strategic language use, through emotion and moralization, can be effectively deployed by bots for deception or accelerated information diffusion (Rashkin et al. 2017). On the other hand, Twitter-specific features like hashtags allow more directed analysis of narratives (RQ4b). As hashtags explicitly allow users to define and participate in significant facets of large-scale events like elections, bots may harness these tools to influence aspects of discourse (Cruickshank and Carley 2020). Moreover, noting the multi-platform nature of information operations, bots may also inject new information into the online conversation by linking to other websites (Krafft and Donovan 2020), We therefore also examine how bots make use of link-sharing (RQ4c).

Data and Analytical Tools

In this work, we examine bot activity in four recent elections in the Asia-Pacific: the 2019 Indonesian elections, the 2019 Philippine elections, the 2020 Taiwanese elections, and the 2020 Singaporean elections. To facilitate comparative analysis, we designed a methodology that holistically addressed each research question in our multi-level framework. It is broadly applicable to social media datasets in general, and mindful of their multilingual nature reflecting unique political and cultural backdrops. Combining techniques network science and machine learning, our framework relied on principles of parsimony and interoperability to ensure interpretable and commensurable insights across contexts (Uyheng et al. 2019).

Data Collection

Data for this study was collected using the Twitter Rest API. The collection process began as soon as official campaign periods commenced for each country. However, because the incumbent Singaporean parliament was dissolved about a month before the election day, all datasets considered in this study have been constrained for uniformity to a five-week period where the last week end the week after election day.

Keywords chosen for data collection included the official election hashtags for each country. For instance, we used #Halalan2019 for the Philippines, #Pemilu2019 for Indonesia, #TaiwanVotes for Taiwan, #sgelection2020 for Singapore, as well as variants thereof. We also included the names of key candidates for each country, with validation from local participants in each election as well as regional news coverage. Data and collection scripts will be made publicly available upon publication.

Overall dataset statistics are summarized in Table 2. Despite similar time frames of data collection, notable disparities in numbers of tweets and unique Twitter accounts may correspond to differences in actual population sizes as well as idiosyncratic usage patterns of Twitter as a social media platform in each country.

Bot Detection with BotHunter

To perform bot detection in each dataset, we used the BotHunter algorithm (Beskow and Carley 2018). BotHunter relies

Country	#Tweets	#Users
Indonesia	866K	204K
Philippines	779K	227K
Taiwan	274K	86K
Singapore	240k	42K

Table 2: Dataset statistics of four Asia-Pacific elections.

on a tiered approach to bot detection that cumulatively considers an increasingly expanded set of user features, message features, network features, and temporal features with more advanced tiers. Based on several random forest models trained on labeled datasets of known inauthentic accounts, the algorithm produces probabilistic scores for each user based on the likelihood that the account is a bot.

Given that we aim to examine the interactive, narrative, and network features of bots, we specifically use Tier 1 of BotHunter, which relies primarily on user features. It also uses several content and structural features, but none which are present in our subsequent analyses. While the use of a specific tool for bot detection by definition limits analysis to the subset of bots it can predict, using Tier 1 prevents a circular argument of merely exploring the same features BotHunter used to make its predictions. As we see later, diversity in results points to the utility of BotHunter in capturing various types of bots across contexts.

Moreover, despite its relatively small number of features, BotHunter (even at Tier 1) has been shown to have competitive precision, recall, and generalizability relative to the state-of-the-art in supervised methods (Beskow 2020). The algorithm has also been used in a variety of case studies of bot activity within national and international settings of online conflict and digital disinformation (Uyheng and Carley 2019; 2020a; Uyheng et al. 2019).

ORA for Social Network Analysis

Finally, to perform a variety of social network analysis tasks, we use the ORA software (Carley, Reminga, and Carley 2018). ORA is an integrated platform for analysis of large-scale, multi-view, and dynamic networks. ORA represents Twitter datasets as complex networks with multiple nodesets and edge types. Nodesets may include user accounts, hashtags, and URLs. Agent by agent networks represent unimodal networks of user accounts connected to each other by directed edges weighted by the number of interactions between users. Different representations are produced for retweet networks, reply networks, mention networks, and quote networks, as well as an all communication network which sums all types of Twitter interaction between users.

Unless stated otherwise, we generally use the all communication network in our analyses. Agent by hashtag or URL networks represent bimodal networks where user accounts are connected to the hashtags they use or the links they share. Here, edges are weighted by the number of times each hashtag or URL is tweeted by the user. Given these network representations, we specifically harness ORA through its community detection functionalities. We use the Leiden grouping algorithm - an improvement with faster run-time and mathematical guarantees over the widely used Louvain algorithm - as a means of automatically detecting network clusters (Traag, Waltman, and van Eck 2019).

In this analysis, we focus on two specific structural features of Leiden clusters: density and the E/I index. Density captures the extent to which users within a cluster are connected to each other. A fully connected cluster (maximum of 1), in which all agents interact with all the others, is maximally dense. On the other hand, the E/I index is a classical network science measure which captures the extent to which members of a group communicate more to out-group members relative to in-group members (Krackhardt and Stern 1988). Higher E/I indices (maximum of 1) indicate more out-group communication; lower E/I indices (minimum of -1) suggest more exclusive in-group communication.

Netmapper for Language Analysis

Finally, to characterize the language used by bots and humans in the datasets examined, we used the Netmapper software (Carley, Reminga, and Carley 2018). The measurement of linguistic cues is based on a rich literature in social psychology linking verbal expressions to various psychological states and social behaviors (Tausczik and Pennebaker 2010). Here, they provide insight into potential maneuvers bots may be engaged in to shift narratives or sow discord through strategic or inflammatory messaging (Rashkin et al. 2017).

For parsimony, we opt for a lexicon-based approach in this work. We specifically focus on measuring each tweet's reading difficulty, and on measuring the use of abusive terms in each tweet. This targets key features of bots characterized by either their use of simple language as automated agents, or their potential use in spreading conflict. Netmapper's measurement of abusive language has previously been validated as a useful tool for detecting hate speech on benchmark datasets (Davidson et al. 2017; Uyheng and Carley 2020b).

Results

RQ1: How Prevalent was Bot Activity in Each National Election?

Our first set of results concerns the prevalence of bot accounts in each electoral conversation across Indonesia, the Philippines, Singapore, and Taiwan. Figure 1 shows the proportion of unique users that BotHunter would classify as bots at different cutoff values, beginning at 50% and ending at 90%, in 10% increments. The 80% cutoff value used in prior applications of BotHunter (Uyheng and Carley 2020b) conservatively suggests that 11% of accounts in the Indonesian and Philippine datasets are bots. In Singapore and Taiwan, much higher proportions of 27% and 25% are observed, respectively.

While these values indicate that larger proportions of the Taiwanese and Singaporean election conversations are bot-driven, the raw numbers of bots remain much higher in the Philippines and Indonesia. These may be linked to asymmetries in country-level usage of Twitter (Kaur et al. 2018).

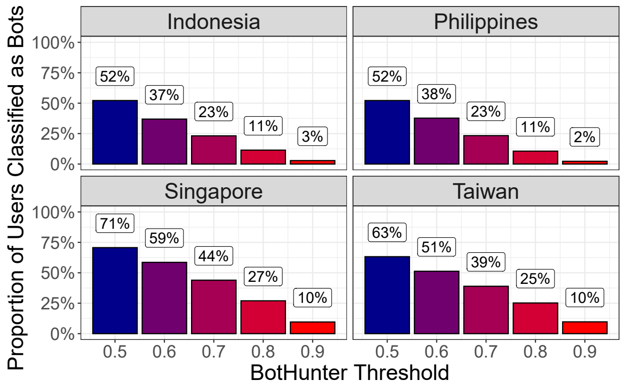

Figure 1: Proportion of users classified as bots at different BotHunter probability thresholds.

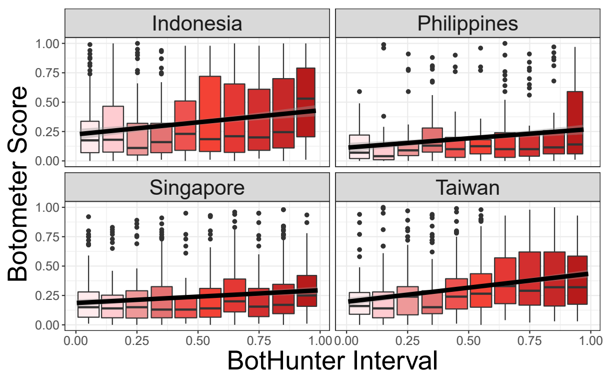

Figure 2: Boxplot comparisons of BotHunter and Botometer predictions on a sample of accounts from each national election. Trendlines depict linear relationships between the two algorithms' assigned bot scores to the sampled accounts.

In settings with fewer human participants, bots account for a larger share of the online conversation even with smaller raw numbers. Furthermore, larger numbers of bots participate where there is likewise more organic participation. But they do not take as large a share of the online conversation.

Validation of Predictions To validate bot predictions, we compared BotHunter predictions with a sample of predictions from Botometer, another supervised model for bot detection (Yang et al. 2020). For each country, we randomly selected 100 accounts having BotHunter scores in the interval from 0 to 0.1, 0.1 to 0.2, and so on. This resulted in a stratified set of 1000 accounts per national dataset. We then used the 'overall universal' scores from the Botometer Pro API[1] as a language-agnostic point of comparison. Figure 2 presents the results of this analysis.

Across datasets, Pearson correlation tests all indicate positive associations ($r = 0.13 - 0.27$) between BotHunter and Botometer predictions, all statistically significant below an $\alpha = .001$ significance level. We note, however, that the relatively low magnitude of the correlation coefficients indicate a weak absolute correspondence in the scores of both models. This aligns with concerns raised in recent reviews of the supervised bot detection paradigm (Cresci 2020; Rauchfleisch and Kaiser 2020). We touch on resulting caveats from this analysis at the end of this paper.

These comparisons, however, are also informative in a practical sense. Given the lack of ground truth with the new data in these electoral contexts, the comparison of two state-of-the-art tools may be informative for deployment in the wild (Yang et al. 2019). First, BotHunter appeared to consistently underestimate Botometer predictions, as evidenced by the outliers featured in the boxplots. This may suggest that BotHunter is less likely to produce false positives; conversely, Botometer may be able to flag more borderline cases. Second, the overall positive relationship between results indicates that there is a non-negligible proportion of bots which are likely bots for both models. For subsequent analysis, this also suggests the utility of considering the entire range of BotHunter scores instead of using absolute cutoff values. This echoes recent research which likewise takes a continuum view of information disorder rather than a binary one (Nizzoli et al. 2020). In this manner, ordinal results instead of binary predictions may still benefit from the observed positive relationship.

Finally, we note that this relationship may also be underestimated due to Twitter suspensions. Whereas BotHunter scores were collected concurrently with the rest of the data (2019-2020), Botometer scores were obtained over a year later (2020-2021). Botometer scores were not accessible for accounts that had been suspended in the intervening period, which is more likely for bot accounts which behave in a manner violating Twitter guidelines. This constitutes the second phase of validation by examining the relationship between suspensions and BotHunter predictions, as summarized in Figure 3.

Here, we see a positive relationship between BotHunter scores and suspension proportions for Indonesia, the Philippines, and Singapore ($r = 0.52 - 0.77, p = .006 - .12$).

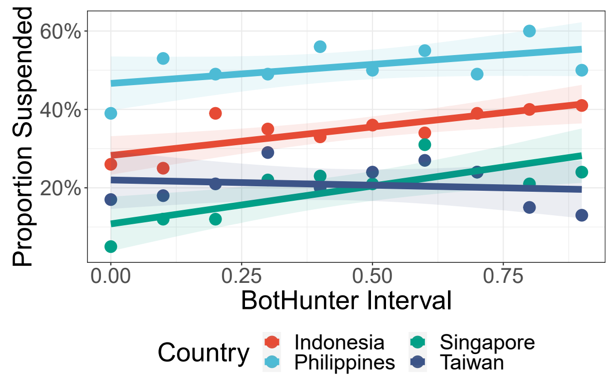

Figure 3: Analysis of account suspensions in relation to BotHunter scores. Trendlines are fitted using OLS regressions. Shaded areas depict 95% confidence intervals

[1]https://rapidapi.com/OSoMe/api/botometer-pro

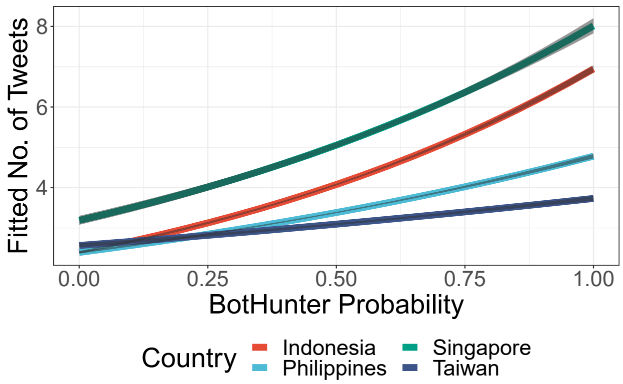

Figure 4: Fitted negative binomial models of tweets by users with different BotHunter probability scores. Shaded areas depict 95% confidence intervals.

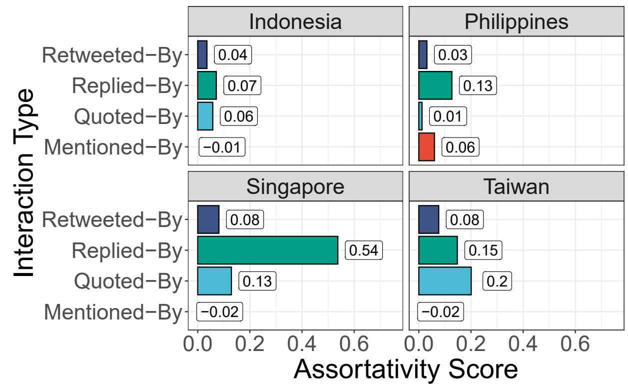

Figure 5: Assortativity scores for retweets, replies, quotes, and mentions.

This offers evidence for the predictive utility of BotHunter in relation to accounts which eventually get suspended by Twitter. Interestingly, the Taiwan dataset has a negative, though non-significant, correlation with suspension probability ($r = -0.15, p = .67$). This may suggest that bot accounts in Taiwan are not as massively engaged in behavior in violation of Twitter guidelines.

RQ2: How Do Bots Interact with Other Agents in Online Conversations?

Negative binomial regression analysis was used to estimate the relationships between BotHunter probabilities and tweet production in our datasets. The results in Figure 4 show that more bot-like accounts generally produced more tweets on average in all four countries.

In the Philippines and Singapore, the most bot-like accounts - with BotHunter scores between 0.9 and 1 - produced twice as many tweets as the least bot-like accounts - with BotHunter scores between 0 and 0.1. In Indonesia, the most bot-like accounts produce quadruple the number of tweets produced by the least bot-like accounts. Interestingly, this trend was relatively weak in Taiwan. While the most bot-like accounts did produce more tweets than less bot-like accounts - around 50% more - the increase was not as smooth as in the other three countries.

We also zeroed in on levels of retweets, replies, mentions, and quotes across different BotHunter scores for sources and targets. To analyze these behaviors, we used the classical measure of (continuous) assortativity over BotHunter scores for each interaction type (Newman 2002). Assortativity is a normalized measure of homophily in interaction patterns. Positive assortativity indexes more exclusive bot-to-bot and human-to-human interactions, while substantial bot-human interactions are marked by negative assortativity. Figure 5 depicts these results.

For most countries, retweets, replies, and quotes are characterized by positive but relatively low assortativity (i.e., close to zero). This indicates that most humans tend to retweet, reply to, and quote other humans. Conversely, most bots tend to interact with other bots. However, there remains a non-negligible proportion of bot-to-human retweet, reply, and quoting behavior. Indonesia and the Philippines, for instance, feature no assortativity scores above 0.15, suggesting high bot exposure across all interaction types. Furthermore, the negative assortativity scores for mentions in Indonesia, Singapore, and Taiwan suggest that mentions, in particular, enable bots in these settings to reach human audiences. Meanwhile, replies feature among the highest assortativity scores in all four countries, suggesting that bots and humans are unable to carry out reply-based conversations.

RQ3: What Network Structures Do Bots Participate In?

Going beyond dyadic relationships, we now consider the structural features of the social networks in which bots are embedded. Using Leiden grouping over all communication networks, we derive network clusters and calculate the average BotHunter scores of all agents in each group. Intuitively, this quantifies the average participation of more bot-like accounts in localized sections of the online conversation. This accounts for their more immediate impacts in online exchanges. Figure 6 visualizes the relationship between the average BotHunter scores in each cluster and their features of density and E/I index.

Broad, joint patterns in community-level density and isolation are consistent across the four countries in the form of a crossover effect. In Indonesia, the Philippines, and Singapore, bots appear in the densest, most isolated (negative E/I index) communities. In contrast, communities which openly interact with other communities (positive E/I index) also tend to have higher numbers of bots when the community is also less dense. In short, heavy bot activity appears to bifurcate in terms of either groups which are open and not dense, or extremely dense and isolated. This may mean that bots join mainstream communities to be part of the broad conversation, or they engage in highly insulated behavior on the fringe of the broader conversation.

We briefly note that Taiwan departs from this pattern to an extent, since denser communities are associated with lower average BotHunter scores for all levels of isolation. Yet it

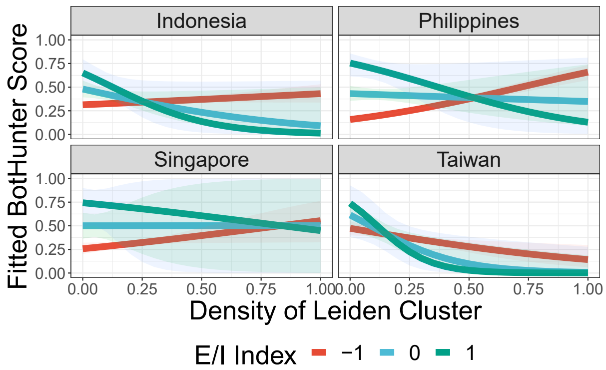

Figure 6: Structural features of Leiden clusters distinguishing between communities with more bot-like and less bot-like accounts. Trendlines are fitted with multiple regression with a logit link function and an interaction effect. Shaded areas depict 95% confidence intervals.

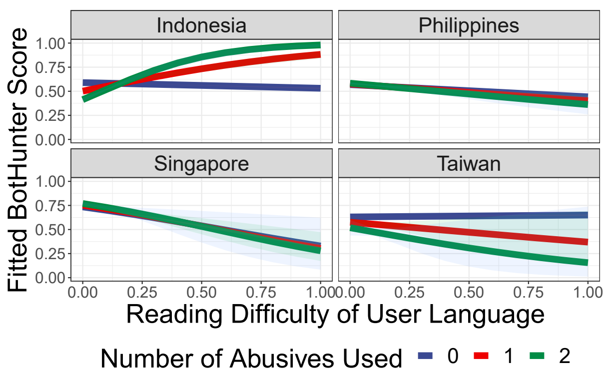

Figure 7: Linguistic measures of abuse and reading difficulty distinguishing more bot-like from less bot-like accounts. Trendlines are fitted with multiple regression with a logit link function and an interaction effect. Shaded areas depict 95% confidence intervals.

still features a crossover effect since the decrease in bot activity associated with higher density takes place at a slower rate for isolated communities.

RQ4a: What Language Do Bots Deploy?

This section now shifts to analyzing bot activity in terms of the language they use to shape online narratives. With Netmapper, we derived measures of abusive language and reading difficulty for all tweets in our datasets. Using a logit link function over individual BotHunter scores, we then ran a regression analysis with lexical abuse and reading difficulty measures as the independent variables. This model included an interaction effect. This quantified the extent to which each feature was associated with users of different levels of bot-likeness. Figure 7 presents the results of this analysis.

Our analysis revealed that low reading difficulty tended to predict higher BotHunter scores across all contexts under study. This suggests that bots are primarily distinguishable based on their use of systematically simpler language relative to humans. This is reasonably in line with prior expectations, given that bots tend to rely on automated software for message-generation; more sophisticated texts would be expected from authentic human accounts.

It is the role of abusive language that differentiates between the four national elections. In the Philippines and Singapore, no difference is observed regardless of whether accounts employ abusive language or not; reading difficulty appears to more sufficiently explain differences between bot-like and human-like accounts. This may indicate that abusive language is used by bots and humans at similar levels in these two countries.

Meanwhile, in Indonesia, resorting to abusive language clearly marks more bot-like activity, reversing the predictive effect of reading difficulty. Because of an interaction effect, Indonesian bots are characterized by their simultaneous use of both abusive and more complex language, possibly pointing to more sophisticated hate tactics.

Counterintuitively, in Taiwan, this pattern is reversed. Here, the use of abusives appears to relate to more human activity, especially among those employing simpler language. This suggests that bots in Taiwan generally use simple language whether or not they are abusive; sophisticated hate is associated here with humans.

RQ4b-RQ4c: What Hashtags and URLs Do Bots Use and Share?

We now deepen our analysis of bot narratives by examining their deployment of hashtags and external links. To this end, we devise an analytical strategy for associating bots and humans with hashtags and URLs.

Computation Strategy For a given country dataset, let G be the bipartite graph corresponding to an agent by hashtag or an agent by URL network. For the purposes of this analysis, both hashtags and URLs may be analyzed equivalently, so henceforth, we refer to hashtags and URLs generically as information. Additionally, we treat all URLs from the same domain as equivalent for parsimony. This prevents us, for instance, from identifying specific YouTube videos salient to bot activity; however, it allows us to focus on broader patterns of information flow, and identify important external websites on the domain-level.

Let e_{ij} represent the weight of the edge connecting agent i, where $i \in \{1, 2, \ldots, m\}$, to information j, where $j \in \{1, 2, \ldots, n\}$. Here m refers to the number of unique user accounts and n is the number of unique pieces of information. We then specifically consider agents $i \in B$, where B is the set of users with BotHunter scores above 0.8, and agents $i \in H$, where H is the set of users with BotHunter scores below 0.2. These two thresholds are chosen complementary to each other, but may be adjusted as a parameter. For given information j, its bot usage score is given by $U_{j,b}$ while its human usage score is given by $U_{j,h}$. These are computed using the equations below:

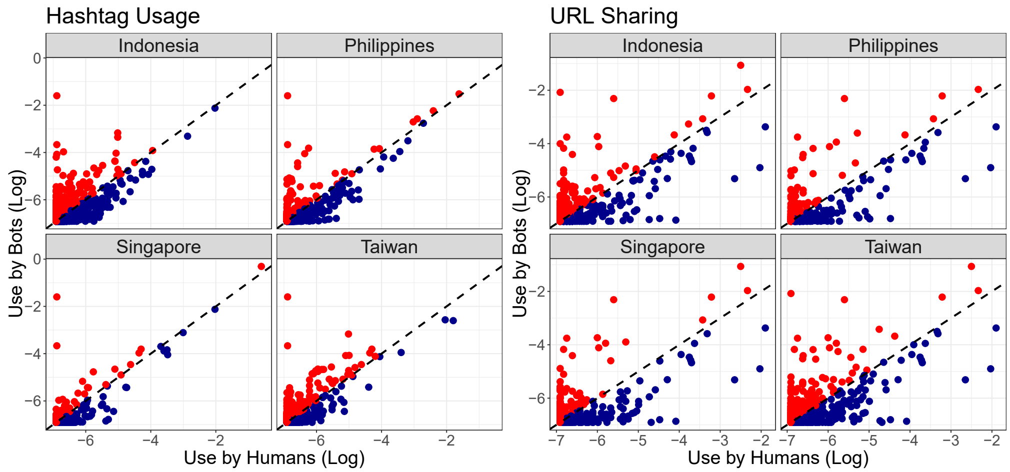

Figure 8: Scatterplot of hashtags and web domains with x-coordinate given by human usage score $U_{k,h}$ and y-coordinate given by bot usage score $U_{k,b}$, both shown in log scale. Broken line depicts $y = x$, such that points located above the line (red) are shared relatively more by bots, while points located below the line (blue) are shared relatively more by humans.

$$U_{j,b} = \frac{\sum_{i \in B} e_{ij}}{\sum_{j=1}^{n} \sum_{i \in B} e_{ij}} \tag{1}$$

$$U_{j,h} = \frac{\sum_{i \in H} e_{ij}}{\sum_{j=1}^{n} \sum_{i \in H} e_{ij}}. \tag{2}$$

Correlation and Outlier Analysis Using these measures, we obtain normalized measurements of how much bots and humans tend to share certain information over others. Figure 8 depicts each hashtag and domain on a scatterplot with positions determined by bot and human usage scores.

Correlation analysis suggests that bot usage score and human usage scores are positively correlated for hashtags ($r = 0.47 - 0.96, p < .001$) and for domains ($r = 0.50 - 0.52, p < .001$). This indicates a relative correspondence between bots and humans in their usage of hashtags and external domains in participating in online electoral information exchange.

However, many pieces of information were also shared disproportionately more by bots than by humans, and vice versa. Visually, these are depicted by points which are farthest from the line denoted by $y = x$. To probe the information used specifically by bots, Table 3 lists the top 30 hashtags and domains j based on their bot usage scores $U_{j,b}$ and for which $U_{j,b} > U_{j,h}$.

Blogs and Misinformation One distinct set of information more closely associated with bots than with humans included blogs and other non-official sources of political coverage of elections. This included domains like `sekseeh` in Indonesia, `getrealphilippines` in the Philippines, `historyogi` in Singapore, `theconservativetreehouse` in Taiwan, as well as `wordpress` and `reddit` in all four countries. While these are not all indicative of misinformation per se, a number of objects in this category may propagate sensationalized, malinformed, or hyperpartisan content (Bastos and Mercea 2019; Shao et al. 2018).

In the Philippines, for instance, the hashtag 'BBMTheRealVP' alleges a conspiracy about a stolen election in 2016 to attack the local Liberal Party running as an opposition party in the 2019 elections. A manual inspection of the most commonly shared Reddit link in the Philippines further contained conspiracies alleging connections between Liberal Party senatoriables and major drug cartels. Websites like `anonymous-post` have likewise been documented for their implication in content farms in related studies of disinformation in the Asia-Pacific (Kaur et al. 2018).

Miscellaneous Entertainment and Commercial Content While not directly related to electoral discourse per se, we also found several domains associated with bot-like activity. Some entertainment domains, for instance, like `allkpop`, were associated with bot accounts in Taiwan likely due to the mass sharing of a story of a k-pop idol visiting Taiwan. Echoing prior work (Mendoza, Tesconi, and Cresci 2020), k-pop content in all four countries was also salient in hashtags specifically around members of the boy group AB6IX and the girl group TWICE.

Besides k-pop, Singapore in particular also featured commercial content unrelated to the elections in an abundance of hashtags related to buying and selling jewelry. This result aligns with prior findings on spam bots (Cresci et al. 2017), specifically in terms of their opportunistic marketing activities during large-scale events to increase public reach for commercial endeavors (Uyheng et al. 2019). Be-

Country	Hashtags	URLs
Indonesia	JokowiMarufProWongCilik, JokowiProUlama, SatukanIndonesia, IndonesiaAmanDamaiSejuk, bidhumaspoldaaceh, salam2periode, jokowiterdepan, government success*, AB6IX, Bandungkab, 01JokowiPresiden, Cimahi, KBB, 01IndonesiaMajuTerus, 01JokowiMenang, pantaucom, 02MainCurang, nawacita, pemiludamaiaceh, HumasPolsekkualasimpang, Netralnews, kubuhoax, HoaxPanwaslu, JokowiBanjirDukungan, jokowijk, police chief*, action*, appeal*, polressragen, PR Police Sragen*	pemilu, google, straitstimes, pilpres, *wordpress*, asiaone, instagram, menitpertama, todayonline, netralnews, *sekseeh*, gonews, indopos, akurat, econ, substack, aljazeera, time, indonesiainside, gosumbar, goaceh, kanalsatu, merahputih, justonly, aktual, beritaislam, solopos, matamatapolitik, wikipedia, setkab
Philippines	BisayaMovement, BisayaForMar, AbanteDigital, HomeofOPM, CPRM6yearsna, NanangIsabel, ParanaqueforTolentino, TGDKapalit, AbanteVideos, *BBMTheRealVP*, BREAKINGNEWSNOW, NoToFatherOfTrainLaw, NoToTrapo, infosec, ToniteNews, GenZ, GiletsJaunes, HalalanDaily, KMJSDiumano, 1allfilipinonewswebsite, The National Tabloid*, LeagueOfProvinces, *trendingtopic*, TropangGadon, **WWG1WGA**, 26Sabalota, 26JVEjercito, BeNice, AngaraInMindanao, JVGood	google, abante, straitstimes, politics, *wordpress*, abogado, nikkei, livedoor, instagram, menitpertama, manilainformer, *getrealphilippines*, newsbitsph, econ, *reddit*, aljazeera, *showbiztrends*, time, goriau, snippetmedia, *getrealpundit*, *bongbongmarcos*, philippinefails, rmn, thestar, washingtonpost, reuters, pna, wikipedia, pcij
Singapore	WeLoveYouDrChee, Cambodia, Masagone, MasagosZulkifli, GenZ, Singapura, **WWG1WGA**, wednesdaymorning, carryminati, PotongPasirHosay, TampinesTogether, GreenParty, civicspace, PotongPasir, SingaporePeoplesTown, jewellery, BuildBackBetterFairer, DearSingapore, Tibet, famousJamus, gems, GEMSCOMPANY, gemstone, gemstonejewelry, gemstones, Channel8, progresswithcompassion, socialmediamarketing	google, ifttt, straitstimes, politics, wordpress, asiaone, nikkei, beritaharian, instagram, todayonline, aecnewstoday, allsingaporestuff, econ, yhoo, epigrambooks, substack, easybranches, aljazeera, time, emmanuel-maria, localnewsingapore, malaymail, *historyogi*, ws-jp, inside, tradeforprofit, newpaper24, wa, thestar, washingtonpost
Taiwan	AB6IX, Invoicing*, Legislators*, Tzuyu*, Taiwanese politics*, Vigorous*, Jaejoong, topNews, blood type*, Wengui live broadcast*, politics*, Wu Dunyi, news*, drama club acting without stage can perform*, I can't do it if I embrace Korean fans*, Please Kaohsiung people to deal with it by themselves*, the only vote is Tsai Ing-wen*, Absix*, Uncover the truth*, Lim Young-min*, infosec, 2020 Taiwan, Finance, Shanghai*, Shanghai, Gordon*, International*, garbage*, parenting*, resulttaiwan*	pemilu, google, ifttt, *wordpress*, abogado, livedoor, *anonymous-post*, instagram, newsbitsph, upmedia, eyeontaiwan, *theconservativetreehouse*, allkpop, buzztaiwan, *reddit*, substack, aljazeera, moeruasia, time, rb99, nextmedia, lvv2, inews, fsight, nooho, inside, tradeforprofit, newpaper24, rocketnews24, will-news

Table 3: Top 30 hashtags and domains used more by likely bot accounts compared to likely human accounts. Translations are generated where relevant (*). Other categories of interest are also reformatted: italics for blogs, misinformation, and hyperpartisan content; underlined for miscellaneous entertainment and commercial content; and bolded for QAnon.

sides these, other domains consistently associated with bot accounts were `googl` and `ifttt`, which upon inspection of individual links, pointed to automated sharing of photos and content from other websites to Twitter accounts.

QAnon Sympathies in the Asia-Pacific? Finally, we remark on the presence of a QAnon hashtag in our Philippine and Singapore datasets. It is minor in its actual use: in the Philippines, it is the 25th outlier hashtag; in Singapore, it is the 7th. However, manual inspection of collected tweets indicated that in the Philippines, the hashtag's usage incorporated the predominantly Roman Catholic culture of the country. Here, the elections were framed in terms of an act of religious consecration, while expressing sympathy with the far-right in the West. In Singapore, on the other hand, the QAnon hashtag was paired with hashtags related to vaccines and the Covid-19 pandemic. One tweet also contained a link to a YouTube video which has since been taken down. We document these observations here not to overstate the impact of the QAnon conspiracy on Asia-Pacific elections, but to highlight resonances with prior work warning of the transnational nature of conspiracy theories (Pyrhönen and Bauvois 2020). Information operations - whether ultimately successful or not - engage in interplay with local social conditions to help in spreading them (Starbird 2019).

Conclusions and Future Work

This paper undertook a comparative computational analysis of bot activity in four elections in the Asia-Pacific region. Using a general methodology consisting of network science and machine learning tools, we developed a systematic analytical framework for key features of bot activity

General Finding	Indonesia	Philippines	Singapore	Taiwan
Bots participate in online discourse.	Higher Number	Higher Number	Higher Proportion	Higher Proportion
Bots are more likely to be suspended.	Yes	Yes	Yes	No
Bots produce more tweets than humans.	Yes	Yes	Yes	Yes
Bots interact with humans via mentions.	Yes	Yes	Yes	Yes
Bot clusters are denser and more isolated.	Yes	Yes	Yes	No
Bots use simpler, more abusive language.	No and Yes	Yes and Yes	Yes and Yes	Yes and No
Bots share partisan or irrelevant content.	Yes	Yes	Yes	Yes

Table 4: Summary of key findings about bot activity in four Asia-Pacific elections.

across multiple, largely understudied contexts. Recognizing the complex socio-technical nature of digital disinformation, we linked bot activity to various measures associated with individual agents, social interaction, community formation, and information sharing.

To integrate our key findings, we present Table 4. This synthesis illustrates the strength of our multi-level, comparative approach. Here, we indicate consistent features of bot activity across contexts, while also identifying divergent cases. Overall, the features which do persist paint a picture that is consistent with the broad understanding of bot behavior in the West. Bots engage in patterns of communication and structure formation which inject artificial views into the online conversation, sow discord and conflict, alter networks of information flow, and elevate non-mainstream or polemical sources of information (Keller and Klinger 2019; Varol et al. 2017; Zannettou et al. 2019).

Exceptions to these general patterns are likewise noteworthy. On the one hand, Taiwan broke from several national patterns observed across multiple levels of analysis. In many ways, these pointed to a healthier online sphere: overall lower tweet production by bots, less association with automation and suspension, and mostly entertainment-related bot-shared content. In Taiwan (and to an extent, Singapore), the evidence examined suggests the incursion of artificial forces in online discourse, but they do not particularly appear to be of the same political malignancy seen in the Philippines and Indonesia. Conversely, some countries also amplified the observed effects of concern. In particular, the abundance of low-credibility information sites and the presence of QAnon sympathies in Philippine bots point to the distinct seriousness of information disorder in the nascent democracy. The Asia-Pacific thus collectively constitutes an important, albeit understudied, setting for understanding common dynamics of online bot activity, but it is also an non-monolithic geopolitical region featuring its own diversity premised on online and offline inequalities (Tan 2020).

Taken together, many of the features uncovered point to troubling effects inauthentic actors may have on large-scale national democratic deliberations. Our findings thus also affirm the significance of computational analysis for understanding bot-driven disinformation across regions. Observing information disorder in the Asia-Pacific comparable to more well-studied phenomena in the West also reiterates the urgency of more global approaches to understanding digital disinformation, as well as platform and policy design, for which computational approaches like those shown here may offer important insights. This does not replace but would ideally work hand-in-hand with more idiographic field work contextually attuned to local political and cultural forces.

Several limitations nuance the interpretations of our findings. First, by adopting a comparative perspective, we limit opportunities for more in-depth analyses of particular maneuvers in individual national settings. By casting a wide net, however, our work hopefully points to several worthwhile directions for further exploration in the four countries presented here, as well as in other electoral contexts and information operations more broadly (Humprecht, Esser, and Van Aelst 2020; Miller and Vaccari 2020). Second, the goal of this work was not to improve on existing models but rather to apply them in an integrated fashion. That said, inasmuch as real-time implementation of existing tools often invokes similar reliance on off-the-shelf methodologies (Alizadeh et al. 2020; Yang et al. 2019), important caveats remain. Understanding previously studied bot types like those encoded in supervised models may be useful even for identifying unanticipated bot-related phenomena (e.g., differences in narrative and network dynamics, emergence of QAnon sympathies), but static tools cannot totally address dynamically moving targets like information disorder (Cresci et al. 2017; Rauchfleisch and Kaiser 2020). Moreover, while our analysis does include bot-to-human interactions, we do not make deterministic claims about influencing human cognitions or emotions especially within a culturally diverse setting (Starbird 2019).

Acknowledging the foregoing limitations, future work may therefore extend the insights gleaned from this research in various ways. Numerous observations made here may precipitate fruitful drill-down analysis into more country-specific case analysis. From these standpoints, researchers may consider deep dives into any of the phenomena we point to here, including country-level expressions of abuse, the factors accounting for the structural properties of bot-dominated clusters; and the specific misinformation shared in each election. New tool development may also consider our work as a starting point. For instance, the distinguishing bot behaviors found here may guide augmentation of bot detection datasets and models. More comprehensive lists of trustworthy and untrustworthy websites in the Asia-Pacific may also draw on our exploratory findings. Finally, we posit that the strands of inquiry invoked here may be readily revisited in light of ongoing innovations in bot detection (Cresci 2020; Nizzoli et al. 2020; Pacheco et al. 2020). Given the general, flexible frameworks presented here - both conceptually and methodologically - key aspects of this work may readily be extended across more diverse geopolitical settings and in tandem with analytical developments in the field (Carley 2020).

Acknowledgments

This work was supported in part by the Knight Foundation and the Office of Naval Research grants N000141812106 and N000141812108. Additional support was provided by the Center for Computational Analysis of Social and Organizational Systems (CASOS) and the Center for Informed Democracy and Social Cybersecurity (IDeaS). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Knight Foundation, Office of Naval Research or the U.S. government.

References

Agarwal, P.; Garimella, K.; Joglekar, S.; Sastry, N.; and Tyson, G. 2020. Characterising user content on a multilingual social network. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, 2–11.

Alizadeh, M.; Shapiro, J. N.; Buntain, C.; and Tucker, J. A. 2020. Content-based features predict social media influence operations. *Science Advances* 6(30):eabb5824.

Bastos, M. T., and Mercea, D. 2019. The Brexit botnet and user-generated hyperpartisan news. *Social Science Computer Review* 37(1):38–54.

Beskow, D. M., and Carley, K. M. 2018. Bot-hunter: A tiered approach to detecting & characterizing automated activity on Twitter. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, volume 8.

Beskow, D. 2020. *Finding and Characterizing Information Warfare Campaigns*. Ph.D. Dissertation, Carnegie Mellon University.

Bessi, A., and Ferrara, E. 2016. Social bots distort the 2016 US Presidential election online discussion. *First Monday* 21(11-7).

Bovet, A., and Makse, H. A. 2019. Influence of fake news in Twitter during the 2016 US presidential election. *Nature Communications* 10(1):1–14.

Carley, L. R.; Reminga, J.; and Carley, K. M. 2018. ORA & NetMapper. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*. Springer.

Carley, K. M. 2020. Social cybersecurity: An emerging science. *Computational and Mathematical Organization Theory* 26(4):365–381.

Cha, M.; Gao, W.; and Li, C.-T. 2020. Detecting fake news in social media: An Asia-Pacific perspective. *Communications of the ACM* 63(4):68–71.

Cresci, S.; Di Pietro, R.; Petrocchi, M.; Spognardi, A.; and Tesconi, M. 2017. The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race. In *Proceedings of the 26th International Conference on World Wide Web Companion*, 963–972.

Cresci, S. 2020. A decade of social bot detection. *Communications of the ACM* 63(10):72–83.

Cruickshank, I. J., and Carley, K. M. 2020. Characterizing communities of hashtag usage on twitter during the 2020 COVID-19 pandemic by multi-view clustering. *Applied Network Science* 5(1):1–40.

Das, A., and Schroeder, R. 2020. Online disinformation in the run-up to the Indian 2019 election. *Information, Communication & Society* 1–17.

Davidson, T.; Warmsley, D.; Macy, M.; and Weber, I. 2017. Automated hate speech detection and the problem of offensive language. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 11.

Gilani, Z.; Farahbakhsh, R.; Tyson, G.; and Crowcroft, J. 2019. A large-scale behavioural analysis of bots and humans on Twitter. *ACM Transactions on the Web* 13(1):1–23.

Grinberg, N.; Joseph, K.; Friedland, L.; Swire-Thompson, B.; and Lazer, D. 2019. Fake news on Twitter during the 2016 US presidential election. *Science* 363(6425):374–378.

Hopkins, J. 2014. Cybertroopers and tea parties: Government use of the internet in Malaysia. *Asian Journal of Communication* 24(1):5–24.

Humprecht, E.; Esser, F.; and Van Aelst, P. 2020. Resilience to online disinformation: A framework for cross-national comparative research. *The International Journal of Press/Politics*.

Kaur, K.; Nair, S.; Kwok, Y.; Kajimoto, M.; Chua, Y. T.; Labiste, M. D.; Soon, C.; Jo, H.; and Lin, L. 2018. *Information Disorder in Asia: Overview of Misinformation Ecosystem in India, Indonesia, Japan, the Philippines, Singapore, South Korea, and Taiwan*. The Journalism & Media Studies Centre, The University of Hong Kong.

Keller, T. R., and Klinger, U. 2019. Social bots in election campaigns: Theoretical, empirical, and methodological implications. *Political Communication* 36(1):171–189.

Krackhardt, D., and Stern, R. N. 1988. Informal networks and organizational crises: An experimental simulation. *Social Psychology Quarterly* 123–140.

Krafft, P., and Donovan, J. 2020. Disinformation by design: The use of evidence collages and platform filtering in a media manipulation campaign. *Political Communication* 37(2):194–214.

Luceri, L.; Giordano, S.; and Ferrara, E. 2020. Detecting troll behavior via inverse reinforcement learning: A case study of Russian trolls in the 2016 US election. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, 417–427.

Mendoza, M.; Tesconi, M.; and Cresci, S. 2020. Bots in social and interaction networks: Detection and impact estimation. *ACM Transactions on Information Systems* 39(1):1–32.

Miller, M. L., and Vaccari, C. 2020. Digital threats to democracy: Comparative lessons and possible remedies. *The International Journal of Press/Politics*.

Ndlela, M. N. 2020. Social media algorithms, bots and elections in Africa. In *Social Media and Elections in Africa, Volume 1*. Springer. 13–37.

Newman, M. E. 2002. Assortative mixing in networks. *Physical Review Letters* 89(20):208701.

Nizzoli, L.; Tardelli, S.; Avvenuti, M.; Cresci, S.; and Tesconi, M. 2020. Coordinated behavior on social media in 2019 UK general election. *arXiv preprint arXiv:2008.08370*.

Ong, J. C.; Tapsell, R.; and Curato, N. 2019. Tracking digital disinformation in the 2019 Philippine midterm election. *New Mandala*.

Pacheco, D.; Hui, P.-M.; Torres-Lugo, C.; Truong, B. T.; Flammini, A.; and Menczer, F. 2020. Uncovering coordinated networks on social media. *arXiv preprint arXiv:2001.05658*.

Pyrhönen, N., and Bauvois, G. 2020. Conspiracies beyond fake news. produsing reinformation on presidential elections in the transnational hybrid media system. *Sociological Inquiry* 90(4):705–731.

Rashkin, H.; Choi, E.; Jang, J. Y.; Volkova, S.; and Choi, Y. 2017. Truth of varying shades: Analyzing language in fake news and political fact-checking. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2931–2937.

Rathnayake, C., and Buente, W. 2017. Incidental effects of automated retweeting: An exploratory network perspective on bot activity during Sri Lanka's presidential election in 2015. *Bulletin of Science, Technology & Society* 37(1):57–65.

Rauchfleisch, A., and Kaiser, J. 2020. The false positive problem of automatic bot detection in social science research. *Berkman Klein Center Research Publication* (2020-3).

Recuero, R.; Soares, F. B.; and Gruzd, A. 2020. Hyperpartisanship, disinformation and political conversations on Twitter: The Brazilian presidential election of 2018. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, 569–578.

Schäfer, F.; Evert, S.; and Heinrich, P. 2017. Japan's 2014 general election: Political bots, right-wing internet activism, and Prime Minister Shinzō Abe's hidden nationalist agenda. *Big Data* 5(4):294–309.

Schuchard, R.; Crooks, A.; Stefanidis, A.; and Croitoru, A. 2018. Bots in nets: Empirical comparative analysis of bot evidence in social networks. In *International Conference on Complex Networks and their Applications*, 424–436. Springer.

Shao, C.; Ciampaglia, G. L.; Varol, O.; Yang, K.-C.; Flammini, A.; and Menczer, F. 2018. The spread of low-credibility content by social bots. *Nature Communications* 9(1):1–9.

Shu, K.; Sliva, A.; Wang, S.; Tang, J.; and Liu, H. 2017. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter* 19(1):22–36.

Starbird, K. 2019. Disinformation's spread: Bots, trolls and all of us. *Nature* 571(7766):449–450.

Tan, N. 2020. Electoral management of digital campaigns and disinformation in East and Southeast Asia. *Election Law Journal: Rules, Politics, and Policy*.

Tausczik, Y. R., and Pennebaker, J. W. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology* 29(1):24–54.

Traag, V. A.; Waltman, L.; and van Eck, N. J. 2019. From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports* 9(1):1–12.

Uyheng, J., and Carley, K. M. 2019. Characterizing bot networks on Twitter: An empirical analysis of contentious issues in the Asia-Pacific. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, 153–162. Springer.

Uyheng, J., and Carley, K. M. 2020a. Bot impacts on public sentiment and community structures: Comparative analysis of three elections in the Asia-Pacific. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*. Springer.

Uyheng, J., and Carley, K. M. 2020b. Bots and online hate during the COVID-19 pandemic: Case studies in the United States and the Philippines. *Journal of Computational Social Science* 3(2):445–468.

Uyheng, J.; Magelinski, T.; Villa-Cox, R.; Sowa, C.; and Carley, K. M. 2019. Interoperable pipelines for social cybersecurity: Assessing Twitter information operations during NATO Trident Juncture 2018. *Computational and Mathematical Organization Theory* 1–19.

Varol, O.; Ferrara, E.; Davis, C. B.; Menczer, F.; and Flammini, A. 2017. Online human-bot interactions: Detection, estimation, and characterization. In *Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017)*, 280–289.

Yang, K.-C.; Varol, O.; Davis, C. A.; Ferrara, E.; Flammini, A.; and Menczer, F. 2019. Arming the public with artificial intelligence to counter social bots. *Human Behavior and Emerging Technologies* 1(1):48–61.

Yang, K.-C.; Varol, O.; Hui, P.-M.; and Menczer, F. 2020. Scalable and generalizable social bot detection through data selection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, 1096–1103.

Zannettou, S.; Caulfield, T.; De Cristofaro, E.; Sirivianos, M.; Stringhini, G.; and Blackburn, J. 2019. Disinformation warfare: Understanding state-sponsored trolls on Twitter and their influence on the web. In *Companion Proceedings of the 2019 World Wide Web Conference*, 218–226.