Abstract

Previous research has suggested online activities can reflect one's ideological position. However, no externally validated measurement method is available in the contexts of nondemocracy, transitional democracy, and flawed democracy, which account for more than 88% of countries and 95.5% of population across the globe. Based on two assumptions that are valid in all types of political systems, this study conceptualizes ideology as a multidimensional latent variable that can be inferred from user-generated content and structure of retweet network and demonstrates a novel approach of mapping ideological positions by recurrent convolutional neural network (RCNN). The model's external validity is established by evaluating its predictive accuracy on other years' data as well as its consistency with the judgements made by an expert panel. The model was applied to illustrate the ideological diversity of random users and newspapers on China's Sina Weibo, revealing the complicated impact of propaganda in the internet age.

*Keywords*: Ideology; Ideological Point Estimation; Online Public Sphere; Microblog; Neural Network

**Mapping Ideological Landscape of Social Media: A Neural Network Approach**

**Yuner Zhu, King-wa Fu**

Communication scholars have long been concerned with ideology. A fundamental mission for communication studies is to analyze the ways people come to understand the world through their interactions with media (Entman & Usher, 2018). These interactions are not occurring in a vacuum but instead are embedded in a society where political and business power are played out (Castells, 2009). Ideology presents a key conceptual link between media interactions and power. Measuring ideology constitutes a critical yet complex task for communication scholars who aim to empirically disentangle the power structure behind media interactions.

Much research in this area has focused almost exclusively on developed democracies and the prevalent measure uses partisanship as the shorthand of ideology, based on the assumptions of well-functioning democracy where center parties are critical anchors of political landscape and can represent the popular beliefs (Bennett & Pfetsch, 2018; Campbell, Converse, Miller, & Stokes, 1980; Converse, 2006). However, in the face of democracy crisis all over the world, such assumptions are being challenged (Bennett & Pfetsch, 2018; Freedom House, 2018; Krastev, 2014). As reported by the Economist Intelligence Unit (2018), out of 167 studied countries, only 20 were developed democracies[i] hosting 4.5% of studied population, while in other 147 countries the democratic systems were either dysfunctional or absent. Furthermore, according to *Freedom in the World 2018* report (Freedom House, 2018), as many as 71 countries suffered declines in democracy level. Some of them, like China and Cambodia, were nondemocracies; others were developed democracies like the United States and the United Kingdom. As pointed out by Bennett and Pfetsch (2018), in the face of hollowing of center parties as well as widespread disconnection of publics from party systems, communication scholars should "avoid the poorly specified assumptions about textbook democracy (p. 246)" and embrace the complex reality of disrupted democracy by admitting that politics extends beyond elite politicians and ideology extends beyond partisanship.

In this paper, we aim to demonstrate an encompassing approach to evaluating ideological position of social media user, which is applicable to not only developed democracies, but also to nondemocracies, transitional democracies and flawed democracies. To avoid above mentioned "ideological reductionism," we give up a narrow conception of ideology and integrate two types of media features into the model, namely linguistic features and relational features. The adoption of these features is built on two loose assumptions that are valid in all types of political systems.

First, linguistic preference entails ideological prejudgment (Ahmed & Xing, 2010; Gentzkow & Shapiro, 2010; Iyyer, Enns, Boyd-Graber, & Resnik, 2014; Therborn, 1980). Language is never neutral but highly selective (Fowler, 2013). Words reveal thoughts. Different word choices reflect various ways of thinking. This is even more true for nondemocractic state where authority censors public information and civic-minded netizens strive to find alternative expressions to circumvent filters of sensitive words. Some used secret jargons and others used homophones (Lim, 2018; Sullivan, 2014). The diversity of ideology among media users can be gauged by the diversity of their word choices.

Second, people are inclined to interact more often with those of similar ideology (Boutet, Kim, & Yoneki, 2012; Colleoni, Rozza, & Arvidsson, 2014; Conover et al., 2011; McPherson, Smith-Lovin, & Cook, 2001; Shin & Thorson, 2017; Sunstein, 2009). Similarity breeds connections (McPherson, Smith-lovin, & Cook, 2001). Research has supported that homophily principle structures virtually every type of network (Conover et al., 2011; Ibarra, 1992; Kalmijn, 1998; Lazarsfeld & Merton, 1954) and catalyzes higher rate of contact between similar people than dissimilar people (McPherson, Smith-lovin, et al., 2001). Thereby, people's ideology can also be inferred from whom they interact with.

Specifically, with the data collected from Sina Weibo, we develop a recurrent convolutional neural network (RCNN) model, feed user-generated content into it, train the model to predict user's ideology with the constraints of the user's retweeting relationships with others. In other words, when measuring ideology, we consider both what users said, whom they retweeted, and who have retweeted them. To validate the model, we evaluate the model predictions with two independent sets of data, i.e. expert survey and the Weibo data in the other five years. Both resulted in high accuracy, assuring robustness of the model. It's worth noting that we make no assumption about the dimensionality of ideological space at work. Whether this space turns out to be unidimensional, two-dimensional, or higher-dimensional is entirely determined by the model that best fitted the data.

Our approach represents the first measurement tool that can automatically evaluate the ideological positions for a large online population in a regime without well-functioning multi-party system. To illustrate its potential utilities, we apply the model to estimate ideological points of a representative sample of online populace as well as a list of 96 general-interest newspapers. Finally, methodological and political implications of this study are discussed.

## Literature Review

### Definition of ideology

Ideology is an inexhaustible topic in academic writings and it carries a variety of connotations. It can refer to the core values in individuals about the conditions of their existence, the moral principles in communities about the proper way they should react to different situations, and the common beliefs in society about what the world should be like in the future.

When Karl Marx first popularized the term in the 1930s, it was merely about class struggles. In Marx's definition, ideology was the false consciousness about the bourgeois society to be a fantasy world which served as spiritual weapons for bourgeois class to suppress proletariat (Marx & Engels, 2009). Althusser (2001) expanded its connotation to encompass all social classes, saying that "man is an ideological animal by nature" (p.171). From his point of view, scientific knowledge is the only thing that is free of ideology. Therefore, he advocated scientific knowledge as a weapon to demystify ideology of bourgeoises, rituals, institutions and other ideological state apparatuses (ISAs).

Starting with a critical review of Althusser's work, Therborn (1980) expressed his disagreement with Althusser in many aspects. Firstly, Therborn (1980) theorized that ideology was the medium through which individuals would recognize "what exists," "what is good" and "what is possible." This definition clearly considers scientific knowledge as an embodiment of ideology, since the missions of descriptive, analytical and predictive research are exactly to make people recognize "what exists," "what is good" and "what is possible." Therefore, he insisted that knowledge was always knowledge from certain position which should never be taken as neutral like in Althusserian theory. Secondly, he dismissed arguments postulating ideology as merely social cement and pointed out the coexistence of multiple ideologies. From his perspective, ideology was plural and dialectical, which was potential to serve the interest of any community. Thirdly, ideology was constituted by dynamic social process with heterogeneous subjects. Finally, discourse, rather than the ISAs, was the critical field of ideological operation.

From Marx to Althusser to Therborn, we can see a shift from traditional approach of class struggle to a new approach of sociolinguistics. Along with this reorientation, the connotation of ideology was broadened step by step and evolved into its contemporary form.

Greatly inspired by Therborn's propositions, the current study conceptualizes ideology as an enduring structure of descriptive (“what exists”), evaluative (“what is good”) and anticipatory (“what is possible”) ideas which will predispose individual to behave selectively. Ideology operates in discourse through various linguistic devices, such as lexical choice, grammatical constructions, metaphors, and exemplars. According to this conception, ideology of a media user can be inferred by a comprehensive inquiry into the relevant discourse and behavior.

**Mapping ideology and the conundrum of dimensionality**

Previous literature often maps an entity's ideology as a position that someone takes in a latent ideological space (Barberá, 2015; Carroll, Lewis, Lo, Poole & Rosenthal, 2009; Jessee, 2009; King, Orlando & Sparks, 2016; Poole, 1997; Shor & McCarty, 2011). However, the actual dimensionality of such latent space is fraught with controversies and dissenting opinions.

Among various typologies, a unidimensional spectrum is most commonly used. Since French Revolution, the terms “left” and “right” are widely used across the world in referring to opposing ideologies. Ideological space is thereby imagined to be a line connecting leftmost and rightmost ideology (Hayek, 2011; Huntington, 1957; Lewis-beck, 1993), with socialism on the left, conservatism on the right, and liberalism in the middle (Hayek, 2011).

Albeit widely used, this unidimensional typology has been criticized as over-simplistic (Carmines & D’Amico, 2015; Zaller, 1992). As Zaller (1992, p.26) put it, “although there are numerous ‘value dimensions’ between which there is no obvious logical connection, many people nonetheless respond to different value dimensions as if they were organized by a common left-right dimension.”

Aware of this, many researchers now posit the necessity of having more than one dimension. Chapel Hill Expert Survey (Polk et al., 2017), Asher (1980) and Malka et al. (2017) came up with a two-dimensional space, constituted by economic dimension and sociocultural dimension. Lijphart (1981) suggested seven dimensions of ideology: socio-economic, religious, cultural, ethnic, urban-rural, regime support, foreign policy, and postmaterialism.

Regarding China, Pan & Xu (2018) found a three-dimensional model, composed of political, socio-economic, and nationalism dimension, was the best fit to a national opinion survey data. Using the same data source as Pan & Xu (2018), nevertheless, Wu (2014) came up with a different ideological space consisted of five dimensions, including cultural liberalism, political conservatism, nationalism, traditionalism and economic sovereignty. Nathan and Shi (1996) argued attitudes towards reform, attitudes of grievance, and attitudes towards democracy were the three dimensions of ideology. Lei (2019) argued constitutionalism and universal values were two most important dimensions of ideology. Ma and Zhang (2014) adopted political, economic, and attitudes towards Mao to be the three dimensions of ideology. Huang et al. (2019) opted for eight dimensions, namely attitudes toward Mao, revolutionary legacy, current regime, ideological propaganda, learning from the West, values of liberty and democracy, and preference of economy.

This being stated, no consensus has been reached on the exact dimensionality at work. The studies presented above seemed to measure overlapping yet distinct conceptions of ideology. Since the structure of dimensions as well as the survey design was mostly derived from a priori knowledge, the proposed dimensionality primarily reflects the dimensionality of researcher’s own understanding rather than reflecting the public opinion. The external validity and completeness of the measures used in previous studies are mostly not established. We don’t know whether the researchers have asked the right questions to encompass sufficiently the dimensions of ideology. As pointed out by Dimaggio, Evans and Bryson (1996), since survey by nature tends to drop low-variance items on which the core consensus may be built as well as to cut out a variety of issues in the long tail, it risks a loss in content and factorial validity in measurement.

**From partisanship to ideology**

A popular and expedient solution to the sticky problem of dimensionality is to shift the research focus from abstract "issue position" to concrete "partisanship." This change is made based on an assumption of well-functioning democracy where political party is the anchor of public opinions and is capable of amalgamating scattering opinions into limited dimensionality (Campbell et al., 1980; Converse, 2006; Poole, 1997).

When mapping partisanship, most studies opted for self-report survey. For example, American National Election Studies included one seven-point and one five-point scale for party identification, ranging from "strong Democrat" to "strong Republican." Pew Research Center used a simplified version with only three options, namely "Democrat," "Republican" and "Independent."

Yet the self-report data is known to suffer from the many issues of reliability and validity, such as nonresponse bias (Hill, Roberts, Ewings, & Gunnell, 1997), subjective bias (Adams, Soumerai, Lomas, & Ross-Degnan, 1999; Shrout et al., 2018), and social desirability bias (Donaldson & Grant-Vallone, 2002; Mercer, Deane, & McGeeney, 2016). To overcome such limitations, there emerges a small yet growing body of studies investigating social media data. Some look into user-generated content while others look into relations among users.

Content-based approach assumes that a person's partisanship manifests in the content he or she generates autonomously. Bakshy et al. (2015) targeted 10.1 million U.S. Facebook active users who provided political views in their profile field and mapped the top 500 political designations onto a five-point scale accordingly. Conover et al. (2011) hired two coders to label the partisanship of 1,000 Twitter users from the content they generated.

Boutet, Kim and Yoneki (2012) adopted a semi-supervised approach as the authors first manually identified 684 partisan accounts and then propagated their partisanship labels through retweet path to their neighbors.

On the other hand, relation-based approach assumes that individuals tend to follow, retweet, or reply to those of similar partisanship. To reduce computational cost and noise, most studies limited their research scope to revolve around politicians whose partisanship is known without dispute. For example, Barberá (2015) scaled Twitter users on an ideological dimension based on which political elites they followed. If a person follows far more Republicans than Democrats, then he or she will be closer to conservative ideology. In the same vein, King, Orlando, and Sparks (2016) calculated pairwise distances between 581 political elites based on their similarity of followers and friends, from which they found a hidden dimension of partisanship. Instead of following relations, Faris et al. (2017) used retweet relations. They identified users who retweeted either Clinton or Trump yet seldom retweeted both as partisans.

Which approach is better? When comparing these two approaches, Pennacchiotti and Popescu (2011) found that both approaches showed great promise but relation-based approach generally outperformed content-based approach, as the former obtained 86% accuracy and the latter only got 77%. However, the best performance came from the combination of them, which resulted in as high as 89% accuracy.

As an exemplification of hybrid approach, Shin and Thorson (2017) considered not only user's sentiments towards parties but also their retweet relations with politicians. People who either post positive comments regarding Democratic Party or retweet pro-Democrat politicians are evaluated as Democrats. Another example is the work by Colleoni, Rozza and Arvidsson (2014), which identified users that followed exclusively either Republican or Democrat legislators as partisans and collected their political tweets as the training sets for partisanship classifier. After training, the classifier was applied to categorize users from their sample into Republicans and Democrats.

The above mentioned papers, despite providing valuable references for our study, deployed partisanship as the only indicator of ideology. However, this approach is not applicable to a context where multi-party system is defective or absent, such as nondemocracy, transitional democracies, and flawed democracies. To address this shortfall, our research, following the hybrid approach, aims to analyze entire retweeting network among a wide range of users, including but not limited to politicians, and let dimensionality as well as ideological groupings naturally emerge without pre-imposing politicians as reference points.

# Method

## Data

In this study, we investigate the Chinese Twitter-like microblogging platform Sina Weibo. The primary data, including original posts and retweets on Sina Weibo, were collected by Weiboscope, a data archive managed by a research group at the University of Hong Kong (Fu, Chan, & Chau, 2013; Fu & Chau, 2013). Since 2010, the research group has been collecting the data through Sina Weibo Open API by sampling a list of 72,703 high-profile users and 43,255 random users,[ii] whose posts, including those being censored later (Fu, Chan, & Chau, 2013), were retrieved. The 72,703 high-profile users are known as key opinion leaders, who generate the majority of messages and play a central role in shaping China's public opinions (Fu & Chau, 2013). However, this sample, despite being influential, is not representative of long-tailed Weibo sphere. Therefore, a separate 43,255 randomly chosen users were added and they constitute a representative sample of the Weibo populace.

To gather training data for the model, we conducted snowball sampling in the pool of high-profile users. We started from 20 prominent news media accounts with more than ten million followers as seeds (see Appendix Table 1) and then sorted out a list of accounts that had been retweeted by any of the 20 seed accounts more than one time in 2016. The exclusion criteria, i.e. ignoring one-time retweets, can significantly minimize the chance of accidental retweets (Rabab'ah, Al-Ayyoub, Jararweh & Al-Kabi, 2016), informative retweets, and dissenting retweets.[iii] We added the new accounts to the seed list and repeated snowballing three times using the same exclusion criteria. Finally, we obtained a list of 10,347 users whose posts and retweeting relations in 2016 were obtained from Weiboscope. Overall 204,652 pairs of retweeting dyads were found. We randomly split these 204,652 pairs into training, testing, and validation sets with 60%, 20%, and 20% respectively. In addition to the 2016 data, we repeated the same procedure for all other years from 2014 to 2018 and randomly extracted 20% of the collected data in each year as additional data to validate the longitudinal generalizability of the model.

**Preprocessing**

Our next task is to convert textual data into numeric data that can be processed by computational tools. This process entails three steps: text segmentation, representative sampling, and vectorization.

We first applied a standard preprocessing procedure that segmented text into a set of tokens, removed stop words, single-character words, [iv] and functional symbols.[v] Then we extracted a set of 1000 words for each user, which is representative of their typical discourse and attitudes. It's worth noting that we didn't exclude apolitical content for two reasons. First, according to our definition, ideology refers to a system of descriptive, evaluative, and anticipatory ideas that predispose individual behavior. This definition doesn't assume relation with politics. Opinions that appear apolitical can still be ideological. Besides, being apolitical is also a political decision by nature. It could be a result of political structure and can have an influence over political structure (Scott-Smith, 2003). As demonstrated later, the inclusion of apolitical content does bring upon some unexpected findings that help broaden our understanding of public opinion dynamics. For details of representative sampling, please refer to Note 1 in supplementary document.

Finally, we used the pretrained word embeddings by Tencent AI Lab (Song, Shi, Li, & Zhang, 2018) to translate 1000-word representative texts into numeric matrices of 1000 by 200 dimensions. Tencent's embedding is state-of-the-art vector representation for Chinese words. It includes not only conventional words but also internet slang. Besides, based on word2vec structure, it can precisely capture semantic and syntactic similarity between words and can place words of similar connotations to closer positions in a multi-dimensional latent space. Compared with other alternatives, using Tencent embedding can optimize the text representation with maximal contextual information. After vectorization, we obtained a 1000 by 200 embedding matrix for each user.

**Retweet network**

Besides linguistic features, another component of the model is relational features. Much research (Boutet et al., 2012; Conover et al., 2011; Shin & Thorson, 2017) has indicated retweet network represents segregation of political affiliation, even more profound than mention network. Therefore, we used retweeting relationship as the relational features.

We retrieved all retweeting relationships among 10,437 users in 2016 and removed one-time retweeting relationship, which resulted in 204,652 ties. As we ignored one-time retweet, the minimum tie weight is two. Based on raw weights, we calculated another three indicators, namely the normalized weight (Eq.1) and weighted indegree centrality of retweeter (Eq.2) and weighted outdegree centrality of originator (Eq.3). Since the distribution of tie weights is highly right-skewed, we applied log-transformation over tie weights to make them conform more closely to normality. Normalization is not necessary for neural network-based modeling. However, it can help get better initial weights and speed up convergence procedure, as LeCun et al. (2012) suggested.

$$\text{Normalized weight: } \overline{W_{ij}} = \ln\left(\frac{w_{ij}}{\sum_i w_{ij}}\right) + 10 \qquad \text{Eq.1}$$

$$\text{Weighted indegree: } In'_j = \ln\left(\sum_i W_{ij}\right) \qquad \text{Eq.2}$$

$$\text{Weighted outdegree: } Out'_i = \ln\left(\sum_j W_{ij}\right) \qquad \text{Eq.3}$$

Normalized weight (Eq.1) is an indicator of ideology similarity/closeness between users. As some people are more reluctant to make retweets than others, to ensure comparability of users, individual variance in retweeting preference needs to be controlled. Suppose user *j* retweets user *i* $W_{ij}$ times and overall, he or she makes $\sum_i W_{ij}$ retweets. The opinion closeness between user *j* and user *i* is expressed as the percentage of $W_{ij}$ in $\sum_i W_{ij}$, which ranges between 0 and 1 with positive skewness. If we take the logarithm of the percentage to adjust the skewness, it will result in a negative number, which is not easily interpretable. Therefore, we added 10 to the logarithm to scale it between 0 and 10. Normalized weights is the target variable in the model.

Weighted indegree (Eq.2) is an indicator of user *j*'s the outreach efforts of information seeking and need for orientation, whereas weighted outdegree (Eq.3) is an indicator of user *i*'s influence in shaping others' opinions. These two variables were adopted as control variables.

**Implementing the hybrid model with recurrent convolutional neural network**

We assume normalized retweet frequency $\overline{W_{ij}}$ between user *i* and user *j* was negatively correlated with their ideological distance, whilst controlling for retweeter's outreach efforts $In'_j$ and originator's influence $Out'_i$. The mathematical equations are as follows:

$$\overline{W_{ij}} = \alpha \cdot In'_j + \beta \cdot Out'_i - \omega \cdot \left\|\mu_i - \mu_j\right\| + \varepsilon \qquad \text{Eq.4}$$

$$\mu_k = f(C_k) \qquad \text{Eq.5}$$

where $\mu_k$ is the ideological point of user $k$ which is a function of its 1000*200 embedding matrix denoted by $C_k$; $\|\mu_j - \mu_i\|$ is the Euclidean distance between user $j$ and user $i$; $\varepsilon$ is the bias term; $\alpha$, $\beta$, $\omega$ and $\varepsilon$ are four trainable parameters that are optimized through back propagation. We set $\omega$ to be non-negative as we assume negative correlation between retweet frequency and ideological distance.

As for $f(C_k)$, we chose to use a deep neural network combining recurrent layers and convolutional layers. Recurrent layer and convolutional layer are two major building blocks of artificial neural network. Recurrent layer is appropriate for modelling sequential structure as it is equipped with memory for handling sequence-dependency. Whereas, convolutional layer is appropriate for modelling hierarchical structure as it is equipped with filter for summarizing information (Yin, Kann, Yu, & Schütze, 2017). Regarding our research object, textual data presents a combination of sequential and hierarchical structure. It is sequential as the order of words does matter to the making of meaning, while its meaning is also hierarchical as we need to go from the understanding of words, clauses, complex sentences, and to the more complex sequences of sentences (Biber, van Dijk, & Kintsch, 1983; Kintsch & van Dijk, 1978). This duality can be largely captured by a combination of recurrent layers and convolutional layers. Besides, since neural network block supports nonlinear modeling, it can outperform linear models like autoregressive integrated moving average model (ARIMA), principal component analysis (PCA), and structural equation model (SEM), in probing nonlinear relationship (Kohzadi, Boyd, Kermanshahi, & Kaastra, 1996).

Figure 1 shows a flow chart describing the general framework of the model in which ideology estimator is simply illustrated as a black box and Figure 2 shows the detailed inner structure of ideology estimator. Generally speaking, we analyze user-generated content to estimate the user's ideological position and then attached this ideology estimator (Figure 2) to a broader framework (aka *motherboard*, Figure 1) with constraints of the user's retweet relationship with others to adjust the parameters inside the ideology estimator in accordance with model assumptions (Eq. 4). Once the ideology estimator is completely trained, metaphorically it can be *dismounted* from its *motherboard* and used separately to predict ideological position of any user account.

[Figure 1 is inserted here]

[Figure 2 is inserted here]

To be specific, the ideology estimator took embedding matrix as input. Such embedding matrix was processed by a GRU layer, whose output was then sent through a one-dimensional convolutional layer plus a max pooling layer. The output of pooling layer was sequential data conforming to same logic across the sequence. Hence, a simple recurrent layer[vi] is added on to the pooling layer whose output is the estimated ideological point of user. As Figure 1 shows, the estimated ideological point of user $j$ was paired with that of user $i$ to calculate their pairwise distance $\|\mu_j - \mu_i\|$. With the pairwise distance, weighted indegree, and weighted outdegree, the model outputs a predicted normalized weight $\overline{W_{ij}}$ as Eq.4 and improves its predictions by optimizing the trainable parameters through multiple iterations of back propagation.

In order to find the best set of hyperparameters for the model, we varied each of the five hyperparameters' value in a reasonable range (see Note 2 of supplementary document) and conducted an exhaustive grid search through the hyperparameter space. Based on the model performance and the rule of parsimony, we chose the set of hyperparameters that yielded satisfactory performance at the least expense of complexity. The hidden ideological space that best fit the model turned out to be of eight dimensions.

The final model obtained relatively good performance, with mean absolute error (MAE) of 0.56, mean squared error (MSE) of 0.64, and mean percentage error (MPE) of 14.7% on testing set. Please refer to the supplementary document for full description of the model structure, model training, model testing, and hyperparameter tuning.

**Validation**

For a model developed by data-driven method, validation is an essential step to confirm whether the statistical model is generalizable to a wider context and, more importantly, is consistent with prior expertise. In this study, we set up two validation tests: First, examining whether the model is longitudinally generalizable to other year's data; Second, examining whether the model can replicate an expert panel's judgement.

We randomly sampled 20% data from 2014 to 2018 as validation set and evaluated the model accuracy over every year's data. The resulting evaluation metrics in other years were found to be consistent and as satisfactory as the one in 2016 (Appendix Table 2), indicating that the model is generalizable longitudinally and, more intriguingly, that even the ideological landscape varied over the years, the underlying logics of mapping ideology on China's social media did not change substantially in this period.

Besides, we conducted a survey to reach 20 experts, who were with either more than five years' experience in the news industry in China or holding a doctoral degree in political science, communication, or journalism. Experts were asked to rate the degree of ideological similarity between 72 pairs of accounts.[vii] With their responses, pairwise interrater correlations were first calculated to test the degree of agreement among the experts, which results in an average of $r_{\text{intercoder}} = 0.57$ ($SE = 0.015$, $t\text{-}statistic = 36.9$, $p < 0.001$) that was significantly greater than 0, assuring that moderate consensus did exist among experts. Then, we sorted out the most votes for each item and calculated the correlation coefficient between the expert panel result and the model output. It's worth mentioning that the output of the model is ideological distance, which is opposite to ideological similarity. Therefore, the model output is expected to be negatively correlated with the expert responses. The resulting correlation coefficient is $r_{\text{validation}} = -0.72$ ($SE = 0.005$, $t\text{-}statistic = -8.79$, $p < 0.001$), which confirms that the model output is strongly consistent with the responses of the expert panel.

## Findings

### Application One: A Tale of Two Public Spheres on Sina Weibo, Newspapers and the Publics

The developed model is applied to examine the widespread dispute about the state and the trend of opinion diversity in China. Some argued that the high-speed internet has accelerated the dissemination of discontents and boost opinion diversity prominently (Herold & Marolt, 2011; Rauchfleisch & Schäfer, 2015; Rosen, 2010; Wines, M., & LaFraniere, 2011; Yang, 2004, 2009), whereas others claimed that China's internet exemplified only prosperity but not diversity in government's battle for information control (Hernández, 2019; Qin, Strömberg, & Wu, 2018). Either growing or decreasing diversity assumptions are yet tested systematically largely due to the lack of reliable methodology and researcher's overreliance on scattered evidence, anecdotes, and abstract inductions (Sullivan, 2014).

We applied the model to estimate ideological points of 4,182 random samples[viii] as well as 96 general-interest newspapers'[ix] Weibo accounts between January 2015 and December 2018. The average Euclidean distance[x] among random samples and that among newspapers were calculated as an indicator of intra-group diversity. If the average distance is larger, then the members are more dispersed in ideological space and the group is more diverse.

[Figure 3 is inserted here]

As shown in Figure 3, the random samples and the newspapers demonstrated very different patterns of change. The ideological diversity of the random samples was relatively stable, fluctuating around an average level of 0.18 with a moderate bulge between July 2016 and November 2017. Whereas, newspapers were more concentrated and less diverse. With only two exceptional points at September 2015 and April 2016, newspaper sample's ideological diversity was consistently lower than random samples. This finding is understandable given that all newspapers in China are under vigorous ideological control and accountable to different levels of Communist Party of China (CPC) apparatus (Qin et al., 2018; Wang & Sparks, 2019) whereas the random samples are mainly constituted by ordinary users who enjoy relatively more freedom in sharing a wider range of contents.

Moreover, the trend of newspapers appears to be consistent with decreasing diversity assumption. As depicted by Figure 3, the first turning point of the newspapers' diversity came in February 2016, after which it was decreasing for nine consecutive months. The decline reflects a harsh ideological control over newspapers imposed by the Chinese government since February 2016. On February 19, President Xi Jinping paid an unusual visit to three state news organizations, namely People's Daily, China Central Television (CCTV) and Xinhua News Agency, and exhorted that news media must pledge its loyalty to the party (Hornby & Clover, 2016; Wong, 2016), signaling a radical move to further tighten the party's ideological control over the news media. Nine days later, an inspection team was dispatched to Central Propaganda Department and State Radio and Television Administration, the two highest decision-making bodies in charge of media, to scrutinize their personnel and to assure the two organizations firmly implemented the party line in their work (Ministry of Supervision, 2016). This marked the beginning of a series of dispatches of inspection teams to media-related sectors, like China Daily and All-China Journalists Association. Thereafter, the ideological diversity of newspapers started dropping and the opinion space for alternative views appeared to be narrowed. However, newspapers, especially the local ones,[xi] were resilient to the central control. After the scrutiny period, newspapers managed to restore previous level of diversity. Yet it did not last long. On April 11, 2017, in the face of the upcoming National Congress which would see a reshuffle in top leadership, Central Propaganda Department hosted a minister-level meeting to deploy pre-congress propaganda agenda in which the unification of thoughts regained the upper hand as to create a "harmonious and stable atmosphere for the $19^{th}$ Congress" (Xinhua News Agency, 2017). This resulted in a sharp dip of newspaper diversity since April 2017. Such reduction is not observed among the random samples.

To sum up, our analysis reveals that newspapers and ordinary publics exhibited different patterns of change in diversity. The assumption about an overall reduction in China's ideological diversity in online public sphere holds true only for newspaper sector but not so for the general public.

**Application Two: Apolitical Netizens and their Politicalization**

Our next application is aimed to shed light on the observed increase in ideological diversity among the random samples during July 2016 and November 2017. Despite the small effect size, does the rise imply any momentum of democratization among the Chinese public even under a tightening media control? To address this question, random sample's average intra-group distance was compared to its average inter-group distance with newspapers in Figure 4. The inter-group distance measures the extent to which random samples resemble official mouthpiece's rhetoric and their ideological position.

[Figure 4 is inserted here]

We have two major observations. First, people on average were quite distant from the newspapers most of the time. The average intra-group distances between random samples is 0.18 while the average inter-group distances between random samples and newspapers is 0.35, almost double the former one. To contextualize this finding, only 0.5% of the random users and 0.04% of their posts were censored[xii] per month. The overwhelming majority of the random users as well as their posts were not politically sensitive nor contentious. Therefore, the large gap is more likely to imply people's apathy rather than discontent with newspapers. This result is consistent with Colleoni, Rozza and Arvidsson (2014), who found most user-generated contents are apolitical and only 10% of discussions on Twitter is related to political issues. Second, unexpectedly, these two trend lines appear to be a mirror image of each other – when one goes up, the other goes down; vice versa – suggesting a negative correlation between intragroup and intergroup distance. The results of Pearson's correlation test affirmed this proposition. The correlation coefficient is significantly negative ($r = -0.71$, $SE = 0.08$, $p < 0.001$). Theoretically speaking, the intra-group diversity opinion increases only under three circumstances: 1) part of the random samples moves closer to the party line while the others remain unchanged; or 2) some turn away from the party line while the others remain unchanged; or 3) some get closer and some get away. The first circumstance would come out with a negative correlation, as the distance between the party media and the random samples decrease and distance within the random samples increase. Similarly, the second circumstance will produce a positive correlation and the last with no significant correlation. Since the resulting coefficient turns out to be negative, the increase of intra-group diversity is induced by people's move towards party media, rather than the other way around.

These findings suggest that political apathy is pervasive among Chinese netizens and the online opinion space is less involved in issues covered by newspapers (Mou, Atkin, & Fu, 2011; J. C. Wu, 2014), which led to a sharp gulf between newspapers and grassroot netizens. When apolitical are politicalized, the opinion diversity is increased. However, the politicization of apolitical publics on Sina Weibo is often accomplished by their echoing, instead of challenging, official rhetoric and dominant ideology of mouthpiece newspapers. This type of politicization is not sustainable and cannot substantially vitalize the opinion landscape of ordinary publics. After a period of increasing, the diversity of the random samples fell back to its original apolitical level and the distance with the party media was widened again to a degree similar to that in 2015.

## Conclusion

Communication studies often require systematic information on the ideology of social actors. The development of measurement methods to estimate ideological points is typically built on an assumption of well-functioning democracy (Barberá, 2015; Campbell et al., 1980; Converse, 2006; Poole, 1997). However, such an approach may not be fully valid in the contexts of nondemocracy, transitional democracy and flawed democracy, which largely outnumbers developed democracy in 2018 (The Economist Intelligence Unit, 2018). In this study, we show that, beyond the framework of well-functioning democracy, user-generated content and the structure of retweet network can indicate social media user's ideological positions by using the developed model.

The benefits of this novel approach are fourfold. First, it is applicable to all types of political systems. Second, it is built upon a bottom-up investigation into observational data. As opposed to self-report survey that requires strong priori assumption, the approach allows the ideological space to emerge from data, allows the inclusion of a variety of features and avoids researcher's prior position from oversimplifying the complex reality of public opinions. Third, by exploring social media data, it can catch up on the changing social currents and timely issues, which can possibly go unperceived by researchers. Last, it can locate not only high-profile influencers but also ordinary citizens in a shared high-dimensional space that is validated by expert survey.

We applied the model to estimate ideological positions of 4,182 random samples and 96 newspapers on Sina Weibo from 2015 to 2018. Results revealed that newspapers and general public constituted two parallel public spheres and demonstrated different patterns of change across the study period. The public sphere constituted by newspapers are less diverse and its diversity decreased periodically in response to state's media control policy. Whereas, ordinary publics are more diverse, less reactive to state intervention, and distance from newspapers in ideological positions. This result emphasized the importance of not mistaking newspapers for the reflection of mass public opinion. General public, despite not necessarily being contentious or rebellious to newspapers, varied more remarkably than newspapers and they might seldom follow issues on newspapers attentively. However, such variation is widely ignored and the general public is usually underrepresented in the literature on political communication in nondemocracies. As criticized by Sullivan (2014), research about China's public opinions largely drew on ad hoc evidence that were specific to a few issues supportive of author's argument rather than systematic investigation. The method presented here can be leveraged for large-scale systematic exploration of general public's ideological positions.

Moreover, as opposed to optimistic views on Internet (Herold & Marolt, 2011; Rauchfleisch & Schäfer, 2015; Rosen, 2010; Wines, M., & LaFraniere, 2011; Yang, 2004, 2009), increasing opinion diversity does not always come with an impetus of democratization and deliberation. Our results revealed that, in a society where the authority keeps a strong grip on public opinions, the increasing diversity is mainly achieved by politicization of apolitical publics in a way closer to dominant ideology. This finding can help shed light on the effects of propaganda on citizens' public expression that would remain elusive and untested otherwise. It indicates that the main effect of propaganda might not lie in how capable it is to brainwash people all the time, but in how capable it is to distract their attention from alternative sources and minimize their possibility of accessing alternative narratives. Therefore, even though publics are distant from newspapers most of the time, at the critical moments when they need a narrative to interpret political issues it is still easier for them to opt for official rhetoric imposed by propaganda.

Some important limitations are worth mentioning. First, regarding relational features, current model only explores retweet relationships. Future research can expand the list of relations to include follow, comment, reply, and group relations. Second, the current study did not include the year 2013, which is argued to be a critical year for China's public sphere as it witnessed the surge and the fall of a rare online collective movement of journalists requesting for press freedom. In the absence of early data, we cannot track the origin of political apathy nor the origin of the centralization of media ideology. Third, data are aggregated monthly to ensure the data size are large enough to indicate one's ideology. However, after aggregation, we cannot trace change and causal relationship that occurs in a smaller temporal unit (e.g. daily or hourly). That would be our next step to improve the granularity of the model and make it sensitive to the daily change, or even hourly change.

**References**

Adams, A. S., Soumerai, S. B., Lomas, J., & Ross-Degnan, D. (1999). Evidence of self-report bias in assessing adherence to guidelines. *International Journal for Quality in Health Care*, *11*(3), 187–192.

Ahmed, A., & Xing, E. (2010). Staying informed: supervised and semi-supervised multi-view topical analysis of ideological perspective. *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, 1140–1150.

Althusser, L. (2001). *Lenin and philosophy and other essays*. New York: Monthly Review Press.

Asher, H. B. (1980). *Presidential elections and American politics: Voters, candidates, and campaigns since 1952*. Dorsey Press Homewood, IL.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. doi: 10.1126/science.aaa1160

Barberá, P. (2015). Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data. *Political Analysis*, *23*(1), 76–91. doi: 10.1093/pan/mpu011

Bennett, W. L., & Pfetsch, B. (2018). Rethinking Political Communication in a Time of Disrupted Public Spheres. *Journal of Communication*, *68*(2), 243–253. doi: 10.1093/joc/jqx017

Biber, D., van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press. doi: 10.2307/415483

Boutet, A., Kim, H., & Yoneki, E. (2012). What's in Your Tweets? I Know Who You Supported in the UK 2010 General Election. *Advances in Social Networks Analysis and Mining (ASONAM)*, 411–414.

Carroll, R., Lewis, J. B., Lo, J., Poole, K. T., & Rosenthal, H. (2009). Measuring Bias and Uncertainty in DW-NOMINATE Ideal Point Estimates via the Parametric Bootstrap. *Political Analysis*, 17(3), 261–275. doi: 10.1093/pan/mpp005

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1980). *The american voter*. University of Chicago Press.

Carmines, E. G., & D'Amico, N. J. (2015). The new look in political ideology research. *Annual Review of Political Science, 18*, 205–216.

Castells, M. (2009). *Communication Power*. New York: Oxford University Press.

Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data.

*Journal of Communication*, *64*(2), 317–332. doi: 10.1111/jcom.12084

Conover, M. D., Ratkiewicz, J., Francisco, M., Gonc¸alves, B., Flammini, A., & Menczer, F. (2011). Political Polarization on Twitter. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media Political* (Vol. July, pp. 89–96). doi: 10.1021/ja202932e

Converse, P. E. (2006). The nature of belief systems in mass publics (1964). *Critical Review, 18*(1–3), 1–74.

Dimaggio, P., Evans, J., & Bryson, B. (1996). Have American's Social Attitudes Become More Polarized? Have Americans' Social Attitudes Become More Polarized?'. *American Journal of Sociology*, *102*(3), 690–755.

Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology*, *17*(2), 245–260.

Entman, R. M., & Usher, N. (2018). Framing in a Fractured Democracy: Impacts of Digital Technology on Ideology, Power and Cascading Network Activation. *Journal of Communication*, *68*(2), 298–308. doi: 10.1093/joc/jqx019

Faris, Ro. M., Roberts, H., Etling, B., Bourassa, N., Zuckerman, E., & Benkler, Y. (2017). Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election. Retrieved from Berkman Klein Center for Internet & Society website: https://dash.harvard.edu/handle/1/33759251

Fowler, R. (2013). *Language in the News: Discourse and Ideology in the Press*. Routledge.

Freedom House. (2018). *Freedom in the World 2018*.

Fu, K. W., Chan, C. H., & Chau, M. (2013). Assessing Censorship on Microblogs in China. *IEEE Internet Computing*, *17*(3), 2–10. doi: 10.1109/MIC.2013.28

Fu, K. W., & Chau, M. (2013). Reality Check for the Chinese Microblog Space: A Random Sampling Approach. *PLoS ONE*, *8*(3). doi: 10.1371/journal.pone.0058356

Gentzkow, M., & Shapiro, J. M. (2010). What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica*, *78*(1), 35–71. doi: 10.3982/ECTA7195

Hayek, F. A. (2011). Why I am Not a Conservative. In *The Constitution of Liberty: The Definitive Edition* (pp. 517–533). Chicago: University of Chicago Press.

Hernández, J. C. (2019, March 16). Why China Silenced a Clickbait Queen in Its Battle for Information Control. *The New York Times*.

Herold, D. K., & Marolt, P. (2011). *Online society in China: Creating, celebrating, and instrumentalising the online carnival* (Vol. 25). Taylor & Francis.

Hill, A., Roberts, J., Ewings, P., & Gunnell, D. (1997). Non-response bias in a lifestyle survey. *Journal of Public Health*, *19*(2), 203–207.

Hornby, L., & Clover, C. (2016, April 2). China media: Pressed into service. *Financial Times*.

Huang, R., Gui, Y., Sun, X., Huang, R., Gui, Y., & Sun, X. (2019). Beyond the Left-Right Spectrum: A Typological Analysis of Ideologues in China' s Weibo Space Beyond the Left-Right Spectrum: A Typological Analysis of Ideologues in China' s Weibo Space. *Journal of Contemporary China*, doi: 10.1080/10670564.2019.1580423

Huntington, S. P. (1957). Conservatism as an Ideology. *The American Political Science Review*, *51*(2), 454–473.

Ibarra, H. (1992). Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Administrative Science Quarterly*, 422–447.

Iyyer, M., Enns, P., Boyd-Graber, J., & Resnik, P. (2014). Political Ideology Detection Using Recursive Neural Networks. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, 1113–1122.

Jessee, S. A. (2009). Spatial Voting in the 2004 Presidential Election. American Political Science Review, 103(01), 59–81. doi: 10.1017/S000305540909008X

Kalmijn, M. (1998). Intermarriage and homogamy: Causes, patterns, trends. *Annual Review of Sociology*, *24*(1), 395–421.

King, A. S., Orlando, F. J., & Sparks, D. B. (2016). Ideological Extremity and Success in Primary Elections: Drawing Inferences from the Twitter Network. *Social Science Computer Review*, *34*(4), 395–415. doi: 10.1177/0894439315595483

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363–394.

Kohzadi, N., Boyd, M. S., Kermanshahi, B., & Kaastra, I. (1996). A Comparison of Artificial Neural Network and Time Series Models for Forecasting Commodity Prices. *Neurocomputing*, *10*(2), 169–181. doi: 10.11648/j.ajtas.20160502.13

Krastev, I. (2014). Introduction. In *Democracy disrupted: The politics of global protest* (pp. 1–6). Philadelphia: University of Pennsylvania Press.

Lazarsfeld, P. F., & Merton, R. K. (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and Control in Modern Society*, *18*(1), 18–66.

LeCun, Y. A., Bottou, L., Orr, G. B., & Müller, K. R. (2012). Efficient Backprop. In *Neural networks: Tricks of the trade* (pp. 9–48). Berlin, Heidelberg.: Springer.

Lei, Y.-W. (2019). Extending Liberalization from the Press to the Internet. In *The Contentious Public Sphere* (pp. 104–128). doi: 10.2307/j.ctvc77jgh.10

Lewis-beck, M. S. (1993). Anchoring the French Voter: Ideology versus Party. *The Journal of Politics*, *55*(4), 1100–1109. doi: 10.2307/2131950

Lijphart, A. (1981). Political Parties: Ideologies and Programs. In D. Butler, H. R. Penniman, & A. Ranney (Eds.), *Democracy at the Polls: A Comparative Study of Competitive National Elections*. Washington D.C.: American Enterprise Institute.

Lim, L. (2018, March 2). China's #MeToo censorship bypassed through netizens' creative use of language. *South China Morning Post*.

Ma, D., & Zhang, S. (2014). 中國網民的「左」與「右」 [Chinese netizens' "left" and "right]. *Twenty-First Century*, *142*, 86–103.

Malka, A., Lelkes, Y., & Soto, C. J. (2017). Are Cultural and Economic Conservatism Positively Correlated? A Large-Scale Cross-National Test. *British Journal of Political Science*, 1–25. doi: 10.1017/s0007123417000072

Marx, K., & Engels, F. (2009). *The economic and philosophic manuscripts of 1844*. Prometheus Books.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, *27*(1), 415–444.

Mercer, A., Deane, C., & McGeeney, K. (2016). *Why 2016 election polls missed their mark*. Washington, USA. Retrieved from Pew Research Center website: http://www.pewresearch.org/fact-tank/2016/11/09/why-2016-election-polls-missed-their-mark/

Ministry of Supervision. (2016, August 26). Central propaganda department's report of results of supervision and rectification (中央宣传部关于巡视整改情况的通报). Retrieved from http://www.ccdi.gov.cn/special/zyxszt/djlxs_zyxs/zgls_18jzydjl_zyxs/201608/t20160829_85961.html

Mou, Y., Atkin, D., & Fu, H. (2011). Predicting Political Discussion in a Censored Virtual Environment. *Political Communication*, *28*(3), 341–356. doi: 10.1080/10584609.2011.572466

Muchnik, L., Pei, S., Parra, L. C., Reis, S. D. S., Andrade, J. S., Havlin, S., & Makse, H. A. (2013). Origins of power-law degree distribution in the heterogeneity of human activity in social networks, *Scientific Reports*, *3*: 1783. doi: 10.1038/srep01783

Nathan, A. J., & Shi, T. (1996). Left and Right with Chinese Characteristics: Issues and Alignments in Deng Xiaoping's China. *World Politics*, *48*(04), 522–550.

Pan, J., & Xu, Y. (2018). China's Ideological Spectrum. *The Journal of Politics*, *80*(1), 254–273. doi: 10.2139/ssrn.2593377

Pennacchiotti, M., & Popescu, A.-M. (2011). A Machine Learning Approach to Twitter User

Classification. Paper presented at *Fifth International AAAI Conference on Weblogs and Social Media*. San Diego, California, USA. Retrieved from: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/viewPaper/2886

Polk, J., Rovny, J., Bakker, R., Edwards, E., Hooghe, L., Jolly, S., … Zilovic, M. (2017). Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 Chapel Hill Expert Survey data. *Research & Politics*, *4*(1), 1–9. doi: 10.1177/2053168016686915

Poole, K. T. (1997). *Congress: a political-economic history of roll call voting*. New York: Oxford University Press.

Qin, B., Strömberg, D., & Wu, Y. (2018). Media Bias in China. *American Economic Review*, *108*(9), 2442–2476. doi: 10.1257/aer.20170947

Rabab'ah, A., Al-Ayyoub, M., Jararweh, Y., & Al-Kabi, M. N. (2016). Measuring the controversy level of Arabic trending topics on Twitter. In *2016 7th International Conference on Information and Communication Systems (ICICS)* (pp. 121–126). doi: 10.1109/IACS.2016.7476097

Rauchfleisch, A., & Schäfer, M. S. (2015). Multiple public spheres of Weibo: a typology of forms and potentials of online public spheres in China. *Information, Communication & Society*, *18*(2), 139–155. doi: 10.1080/1369118X.2014.940364

Rosen, S. (2010). Is the Internet a Positive Force in the Development of Civil Society, a Public Sphere and Democratization in China? *International Journal of Communication*, *4*, 509–516.

Scott-Smith, G. (2003). *The Politics of Apolitical Culture: The Congress for Cultural Freedom and the Political Economy of American Hegemony 1945-1955*. Routledge.

Shin, J., & Thorson, K. (2017). Partisan Selective Sharing: The Biased Diffusion of Fact-Checking Messages on Social Media. *Journal of Communication*, *67*(2), 233–255. doi: 10.1111/jcom.12284

Shor, B., & McCarty, N. (2011). The Ideological Mapping of American Legislatures. *American Political Science Review*, *105*(3), 530–551. doi: 10.1017/S0003055411000153

Shrout, P. E., Stadler, G., Lane, S. P., McClure, M. J., Jackson, G. L., Clavél, F. D., … Bolger, N. (2018). Initial elevation bias in subjective reports. *Proceedings of the National Academy of Sciences*, *115*(1), E15–E23. doi: 10.1073/pnas.1712277115

Song, Y., Shi, S., Li, J., & Zhang, H. (2018). Directional Skip-Gram: Explicitly Distinguishing Left and Right Context for Word Embeddings. In *Proceedings of NAACL-HLT 2018*, 175–180. doi: 10.18653/v1/N18-2028

Sullivan, J. (2014). China' s Weibo: Is faster different? *New Media & Society*, *16*(1), 24–37.

Sunstein, C. R. (2009). Extremism: why and when. In *Going to Extremes: How Like Minds Unite and Divide* (pp. 26–92). Oxford, UK: Oxford University Press.

The Economist Intelligence Unit. (2018). *Democracy Index 2018: Me too? Political participation, protest and democracy*. Retrieved from http://www.eiu.com/Handlers/WhitepaperHandler.ashx?fi=Democracy_Index_2018.pdf&mode=wp&campaignid=Democracy2018

Therborn, G. (1980). *The Ideology of Power and the Power of Ideology*. London, UK: Villiers Publications. doi: 10.4135/9781452233932.n8

Wang, H., & Sparks, C. (2019). Chinese newspaper groups in the digital era: The resurgence of the party press. *Journal of Communication, 69*(1), 94–119. doi: 10.1093/joc/jqy061

Wines, M., & LaFraniere, S. (2011, July 28). In baring facts of train crash, blogs erode China censorship. *The New York Times*.

Wong, E. (2016, February 22). Xi Jinping's News Alert: Chinese Media Must Serve the Party. *The New York Times*.

Wu, A. X. (2014). Ideological polarization over a China-as-superpower mind-set: An exploratory charting of belief systems among Chinese internet users, 2008-2011. *International Journal of Communication*, *8*(1), 2650–2679.

Wu, J. C. (2014). Expanding civic engagement in China: Super Girl and entertainment-based online community. *Information, Communication & Society, 17*(1), 105–120. doi: 10.1080/1369118X.2013.833276

Xinhua News Agency. (2017, April 11). Solidly implement propaganda work for the Nineteenth Congress of Party, Integrate cohesive positive engergy (扎实做好迎接党的十九大宣传工作 更好凝聚团结奋进的强大正能量). *People's Daily*.

Yang, G. (2004). The Co-evolution of the Internet and Civil Society in China. *Asian Survey*, *43*(3), 405–422. doi: 10.1525/as.2003.43.3.405

Yang, G. (2009). *The power ofthe Internet in China: Citizen activism online*. New York, NY: Columbia University Press.

Yin, W., Kann, K., Yu, M., & Schütze, H. (2017). Comparative Study of CNN and RNN for Natural Language Processing. doi: 10.14569/IJACSA.2017.080657

Zaller, J. R. (1992). *The nature and origins of mass opinion*. Cambridge university press.

# Figures

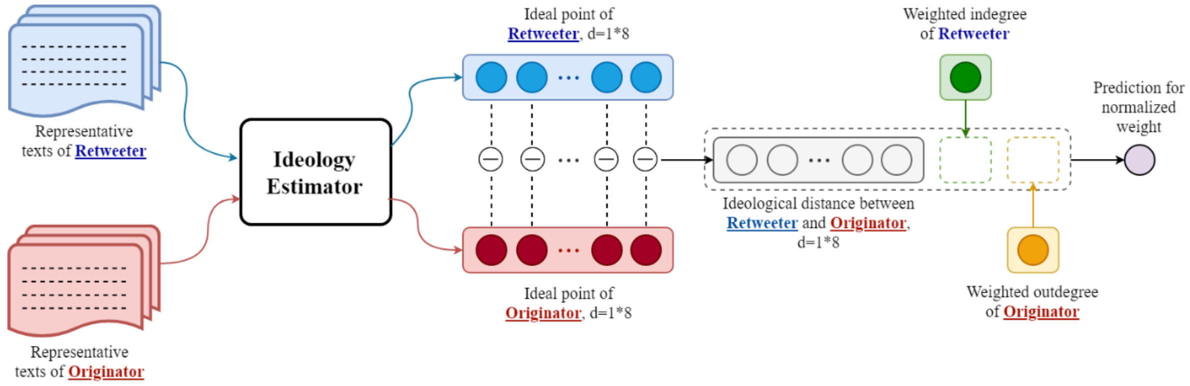


*Figure 1*. General framework of the hybrid model. Input data is a pair of 1000*200 embedding matrices, one from retweeter and the other from originator. Ideology estimator takes a user's embedding matrix as input and returns an estimated ideological point of the user, based on which the pairwise ideological distance is calculated. Weighted indegree of retweeter and weighted outdegree of originator are added as control variables. Model's weights are optimized in order to get better predictions for normalized weights.

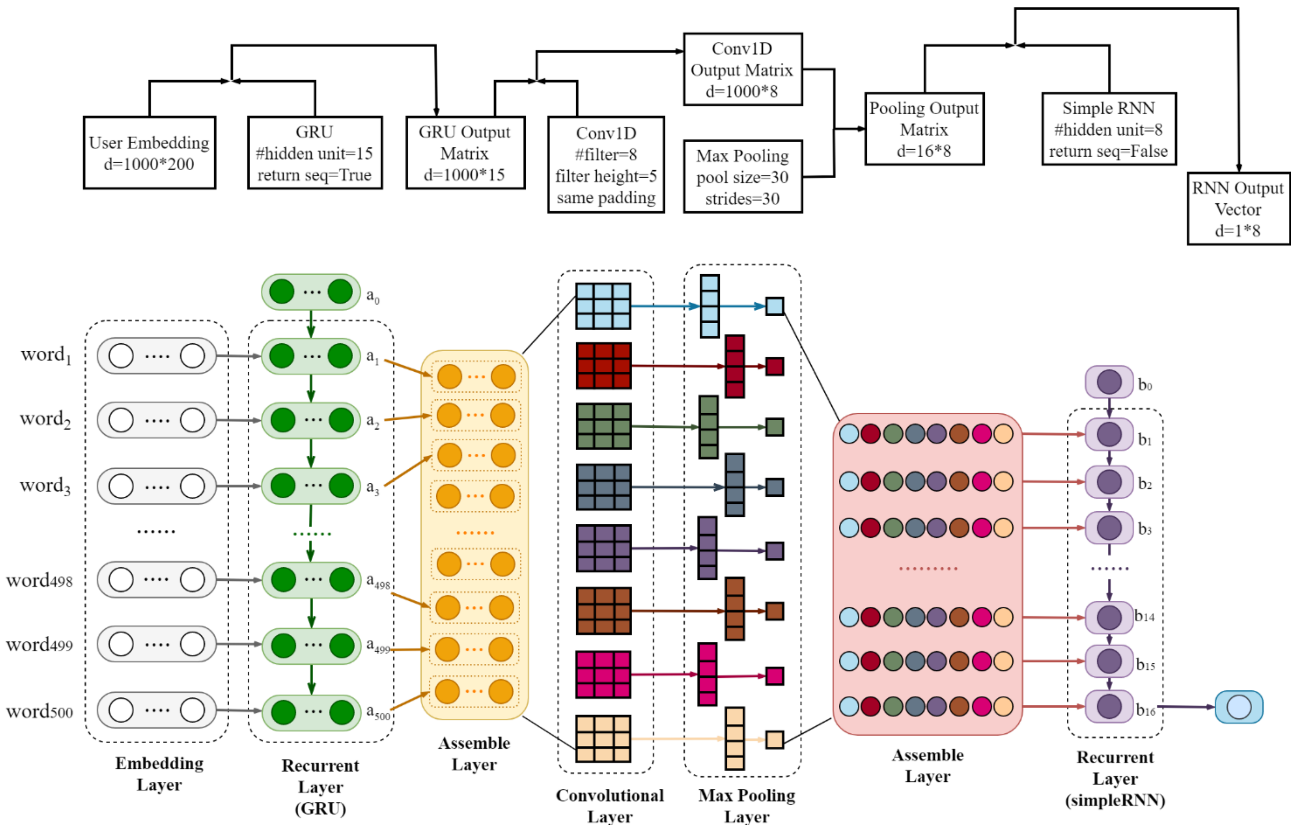


*Figure 2*. Inner structure of the ideology estimator. The estimator has three hidden layers, one gated recurrent units (GRU) layer, one one-dimensional (1D) convolutional layer and one simple recurrent layer. The number of hidden units in GRU, the number and the height of filters in the 1D convolutional layer, the pool size, and the number of hidden units in the simple recurrent layer are five hyperparameters whose values were chosen according to results of hyperparameter tuning.

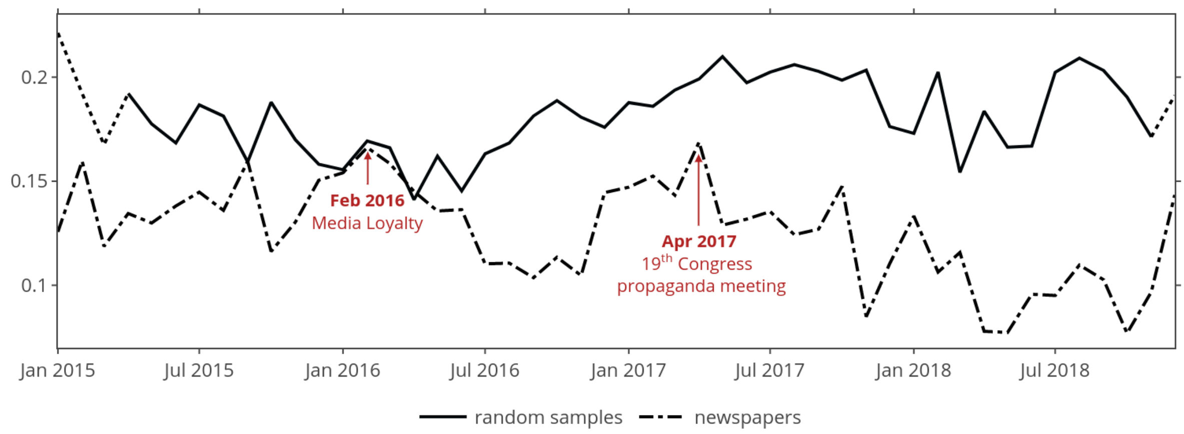


*Figure 3*. Ideological diversities of the random samples and the newspapers during 2015-2018. Solid line represents the trend of opinion diversity of the random samples while the dash-dotted line represents that of the newspapers. Due to the technical issue in Weiboscope, the number of random samples in January 2015, February 2015, and December 2018 were relatively small and thereby the data of these three points were less reliable than those in other months. We dashed these three points as well as the trend lines connecting to them.

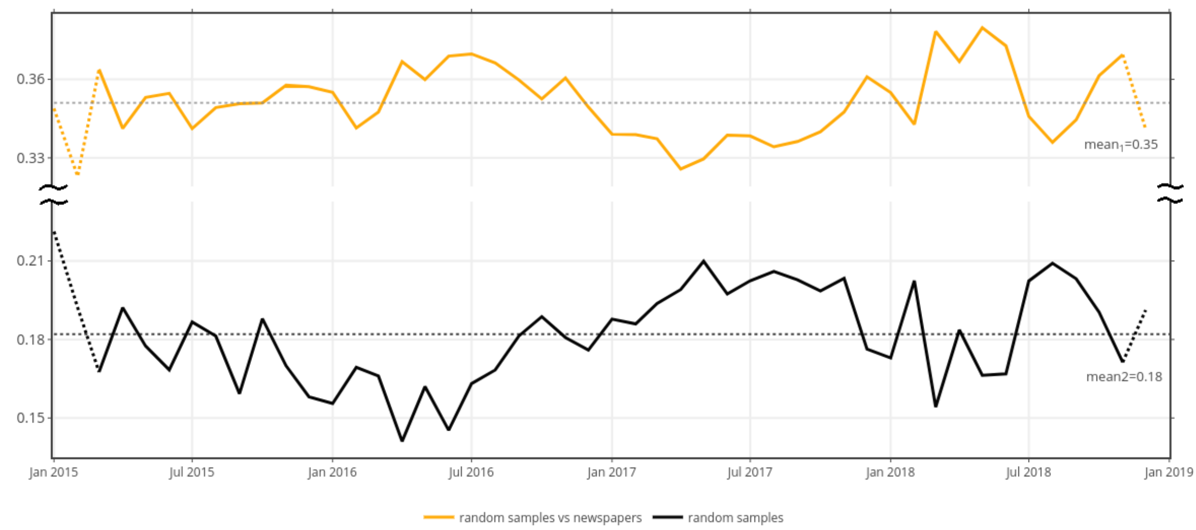


Figure 4. Ideological diversity of the random samples and their average distance with the newspapers during 2015-2018. Black line represents opinion diversity of the random samples, while the orange line represents the average distance between random samples and newspapers. Same as in Figure 3, we dashed three points of January 2015, February 2015 and December 2018 as well as the trend lines connecting to them.

## Appendix

Table 1

*20 News organizations that have been used as seeds in snowball sampling*

| Newspaper | Weibo account | Follower count [a] | Political leaning [b] |
|---|---|---|---|
| People’s Daily | @人民日报 | 92 million | Conservative |
| CCTV News | @央视新闻 | 86 million | Conservative |
| Xinhua News Agency | @新华视点 | 73 million | Conservative |
| Sina Breaking News | @头条新闻 | 66 million | Neutral |
| China Newsweek | @中国新闻周刊 | 52 million | Neutral |
| China Daily | @中国日报 | 43 million | Conservative |
| Beijing News | @新京报 | 36 million | Liberal |
| Caijing.com.cn | @财经网 | 34 million | Neutral |
| Global Times | @环球时报 | 21 million | Conservative |
| Guangming Daily | @光明日报 | 21 million | Conservative |
| People's Liberation Army Daily | @军报记者 | 20 million | Conservative |
| New Weekly | @新周刊 | 19 million | Liberal |
| Phoenix Weekly | @凤凰周刊 | 18 million | Liberal |
| Vista Story | @ Vista 看天下 | 17 million | Neutral |
| Southern Metropolis Daily | @南方都市报 | 17 million | Liberal |
| 21st Century Business Herald | @ 21 世纪经济报道 | 16 million | Neutral |
| Life Week | @三联生活周刊 | 16 million | Neutral |
| Southern Weekly | @南方周末 | 13 million | Liberal |
| China Youth Daily | @中国青年报 | 13 million | Neutral |
| Beijing Youth Daily | @北京青年报 | 10 million | Neutral |

a. Retrieved time: 2019/08/08
b. Determined by the most labels given by expert panel. This column is only used for initiating snowball sampling and has no direct impact over ideology estimator output. As for reducing bias in snowball sampling, we need to ensure the diversity of seeds. Therefore, we opt for the most popular, and probably also the least sophisticated, typology of ideology, i.e. the liberal-neutral-conservative typology, to ensure that every political fraction does take up a substantial portion in the list of seeds.

Table 2.

*Evaluation metrics of model validation*

| Year | MAE | MSE | MPE |
|---|---|---|---|
| *Validate* | | | |
| 2013 | 0.57 | 0.67 | 11.1% |
| 2014 | 0.66 | 0.89 | 15.5% |
| 2015 | 0.67 | 0.95 | 16.0% |
| 2017 | 0.61 | 0.82 | 13.1% |
| 2018 | 0.62 | 0.87 | 13.2% |
| *Train* | | | |
| 2016 | 0.56 | 0.64 | 14.7% |
| *Test* | | | |
| 2016 | 0.61 | 0.78 | 15.2% |

## Endnote

[i] In the original report (The Economist Intelligence Unit, 2018), the authors name it "full democracy", a synonym of developed democracy, which is defined as a country where the political freedoms and civil liberties are respected and a political culture conducive to the flourishing of democracy is underpinned. In addition, the other three categories are "flawed democracy," "hybrid democracy" and "authoritarian regime," the last two of which are equal to the "transitional democracy" and "nondemocracy" in this article.

[ii] Since its establishment in 2010, the data collection scheme of Weiboscope has changed in response to new restrictions imposed by Sina Weibo. Before 2013, Weiboscope experienced fewer restrictions and was able to retrieve a much larger set of user data. However, since 2013 towards, Sina gradually introduced new restrictions to the use of API and then Weiboscope adopted a new sampling pool consisted of over 50,000 high-profile users. In recent years, the project keeps adding high-profile accounts to the system in order to capture the most up-to-date online discussions. Until now, the number of high-profile users has reached 72,703. Besides, it included 43,255 random users who were sampled by random digit generation over user ids so as to obtain a representative sample of Weibo populace.

[iii] Accidental retweets are retweets that happen unintentionally. Informative retweets are retweets that are made to share information neutrally with no value judgement. Negative retweets are retweets that indicate disagreement with original post or originator. The sentiment of retweet is nontrivial as many users emphasize that retweet is not endorsement. Moreover, Sina Weibo allows users to add comments to their retweets, similar to the quoted tweets on Twitter, which introduces more disagreement to be expressed in retweets. We randomly selected 1,693 retweets and manually classify them into three categories, i.e. positive, neutral, and negative. Then we used the 1,693 labeled retweets to train a supervised model to predict the sentiment of 103,121 retweets. We found that 64%, 21% and 15% of one-time retweets were positive, negative and neutral respectively. But if we aggregate the retweets by originators and retweeters, we found that only 3% of the more-than-one-time links were negative. Statistically speaking, an event happening with less than 5% probability is a rare event. So, this supports our assumption that retweeting more than one time is very unlikely to represent disagreement. Thus, in our study, we ignore accounts that are retweeted only once.

[iv] Stop words are extremely common words like "的(of)", "地(-ly)" and "得(to)." Single-character words are words that are composed of only one character. Most stop words are single-character words. Both stop words and single-character words are generally considered to be less informative for text analysis.

[v] Functional symbols include hyperlinks, hashtags and mention/retweet symbols (@XX or RT@XX).

[vi] We didn't choose fully connected (FC) layer, because FC layer assumed entries at the same position have identical meaning while those at different positions were heterogeneous items. In FC layer, first word/entry in all samples will always be multiplied with same weight, which is different from the weight of any other words/entries. However, the pooling output was sequential data that conformed to same rule across the sequence. For example, "I saw Bill on Saturday" and "on Saturday I saw Bill" have same meaning, even though their words are ordered in different ways. If we use FC layer, the first words in both sentences, i.e. "I" and "on", will be assumed to carry same function and will be weighted in the same way. So do the other words. This will produce significantly different outputs for two sentences. As mentioned earlier, sequential data can be better processed with recurrent layer whose weights were shared across sequence. All words in the sentence, regardless of their position, will be weighted in the same way one by one and the final output is determined by their cumulative effects.

[vii] These 72 pairs of accounts were selected based on criterion of popularity and unambiguousness, which means the account as well as its ideology was supposed to be known by experts without much dispute. To assure this, we conducted a pilot survey with 96 pairs of accounts from 36 accounts, asked three assistants who hold a master's degree in journalism or communication to fill the survey and provide their feedback on the clarity of questions and the popularity of accounts after completing the survey. Also, we reviewed the responses for each account and identified a few accounts receiving disparate responses. After the pilot survey, 6 accounts that were either less popular or of ambiguous ideological positions were removed and consequently the 30 remaining accounts constituted 72 pairs that were used in the final validation survey. In case some experts might be unfamiliar with part of accounts, we prepared a website demonstrating representative posts of each account in 2016 to which experts can easily access by clicking the hyperlink underneath survey questions. The link of the survey: https://surveymonkey.com/r/RG7VMMB. The home page showing representative posts of each account: https://juniorworld.github.io/Opinion-Corpus/.

[viii] As mentioned earlier, Weiboscope has been retrieving posts by 43,255. However, online activity is highly asymmetric and power-law distributed with extremely small portion of users contributing to almost all contents on the internet (Muchnik et al., 2013). As suggested by Fu & Chau (2013), on Sina Weibo, 57.4% of random users' timelines were empty and 86.9% of the others with non-empty timelines did not make any original post in a 7-days study period. So, to avoid data sparsity, we considered only random users that made at least 10 original posts in any month within the study period, which amounts to 4,182 users.

[ix] The list of 96 general-interest newspapers is compiled from 2014 China Journalism Yearbook (p.835-839). It includes newspapers at central, provincial and prefectural level. To ensure the relevance of research objects, we removed those newspapers with less than 5% regional market share. The market share is gauged by circulation per issue. For example, Shanghai has an average circulation of 1.6 million per issue in 2014. Therefore, we only include those Shanghai newspapers with circulations of more than 0.13 million per issue.

[x] Since ideological point estimation is produced by simple RNN where tanh is used as activation function, the coordinates of ideological point are limited between 0 and 1. Given that the validated ideological space to be of eight dimensions, the theoretical maximum pairwise distance will be $\approx 2.83$.

[xi] We don't have space for showing related results, yet, put it simply, we found that it was local news outlets contributed to the most of diversity. The average distance between local news outlets is 0.13 while that of central news outlets is 0.02, implying that central news outlets were far more unified and less scattering in their ideal positions.

[xii] In Sina Weibo, posts missing for various reasons are marked differently and can be manifested through API responses. According to Fu et al. (2013), Weiboscope repeatedly makes calls to Weibo API to check the status of posts. Posts that are censored by the platform are marked "permission denied." Posts that the user voluntarily deletes are marked "weibo does not exist." Here we only looked into censored posts, not including those posts deleted by generators themselves.