
Social Media and Repressive Regimes: The Case of Cuba and Venezuela

Michaelanne Dye

Georgia Institute of Technology
85 5th St.
Atlanta, GA 30332 USA
mdye@gatech.edu

Eric Gilbert

Georgia Institute of Technology
85 5th St.
Atlanta, GA 30332 USA
gilbert@cc.gatech.edu

Annie Antón

Georgia Institute of Technology
85 5th St.
Atlanta, GA 30332 USA
aa16@gatech.edu

Amy Bruckman

Georgia Institute of Technology
85 5th St.
Atlanta, GA 30332 USA
asb@cc.gatech.edu

Abstract

Social media played a significant role in shaping the recent revolutionary wave of demonstrations and riots that ultimately led to rulers being forced out of power in Egypt, Tunisia, Libya, and Yemen. However, in other parts of the world, repressive regimes use technical and non-technical methods to suppress what their citizens

Paste the appropriate copyright/license statement here. ACM now supports three different publication options:

- **ACM copyright:** ACM holds the copyright on the work. This is the historical approach.
- **License:** The author(s) retain copyright, but ACM receives an exclusive publication license.
- **Open Access:** The author(s) wish to pay for the work to be open access. The additional fee must be paid to ACM.

This text field is large enough to hold the appropriate release statement assuming it is single-spaced in Verdana 7 point font. Please do not change the size of this text box.

Every submission will be assigned their own unique DOI string to be included here.

say via social media. While much work has been conducted in other parts of the world, we know little about the mechanisms and effects of state-sponsored social media control in Latin America. In this work, we seek to address the technical, conceptual, and operational challenges that social media presents by conducting a mixed-methods study of social media in two countries governed by repressive regimes: Cuba and Venezuela.

Author Keywords

Social media; censorship; Latin America; Cuba; Venezuela.

ACM Classification Keywords

H5.3. Group and Organization Interfaces;
Asynchronous Interaction; Web-based interaction.

Introduction

Traditional narratives paint a picture of the democratizing nature of social media, however, in certain parts of the world, repressive regimes use technical and non-technical methods to suppress what their citizens say via social media. By examining social media activity in Cuba and Venezuela, we seek to understand what persuasive power social media has in these countries, and what power the state has to control it.

Research Questions

Specifically, we are seeking to determine whether: (a) machine learning techniques can predict which individuals sharing content online are likely simply recycling government positions, and which are individual citizens acting independently; and (b) whether it is possible to infer patterns of influence and changes in public opinion within a stream of content from independent citizens. To answer these questions, we are using a combination of machine learning algorithms, geographic inference, paraphrase detection and social network analysis—all informed and validated by qualitative interviews with citizens on the ground. In addition to illuminating the complex nature of information propagation in repressive regimes, this research has broader implications for the role of online social networks as tools of persuasion and intelligence gathering in a global setting.

Background

Cuba

Cuba has been called the second most isolated country in the world, particularly due to its tightly controlled Internet [8]. Reasons for this Internet stagnation include lack of resources, the US embargo against Cuba, and the Cuban government's fear of the implications of the freedom of information [3, 4]. However, changes during the last five years have placed Cuba in the midst of a technological upheaval, pitting government control against the desire for economic success [4]. More Cubans are accessing the Internet due to some key developments such as a high-speed ALBA-1 fiber optic cable to Venezuela [1] and new policy allowing fully owned foreign business in Cuba [2]. In June 2014, Google's Executive Chairman Eric Schmidt visited Cuba along with other Google employees in an effort to learn more about the state of technology on the Island, promote Internet access, and meet with open Internet advocates [7]. In sum, access to the Internet has increased recently in Cuba, and

dramatic change appears to be possible, making this research timely.

Venezuela

The situation in Venezuela has interesting similarities and differences to Cuba (see Table 1). Recent political and economic changes in Venezuela have put the country in a state of unrest causing an outcry of citizens, particularly via social media. The Venezuelan government has increased its surveillance and censorship online in response to the growing usage of social media by the Venezuelan people especially those that seek to oppose the current ruling party [5]. The government has been known to use various tactics to censor opposing information, including site blocking and service disruptions especially during times of heightened political sensitivity (e.g. elections). Further, the government has begun to persecute individuals (particularly bloggers) who utilize social media to speak out against the government. In addition to silencing those who try to speak out against the government, the ruling party is also utilizing social media outlets as tools to control popular opinion through cyber attacks on media outlets, hijackings of Twitter profiles of political activists, and an increased use of anonymous Twitter accounts that support the government [6].

Cuba	Venezuela
15% Internet penetration	44% Internet penetration
Obstacles to Internet access: lack of resources & government control.	Higher levels of access than Cuba. Still limited due to lack of resources & government control.
The government actively blocks certain social media sites & ICT apps.	Social media and ICT apps are not blocked.

The government is assumed to actively monitor Internet activity.	The government is assumed to actively monitor Internet activity.
Activity online is picking up. Much less dissident behavior as in Venezuela.	More activism online; people continue to speak out against government.
The government controls all media forms & works diligently to silence opposing voices online. It's difficult for Cubans to find alternate sources of news.	Media penetration from outside countries is much more prominent in Venezuela, allowing for more diverse opinions.
Bloggers & online activists persecuted & arrested.	Bloggers & online activists persecuted & arrested.
Government uses social media sites to spread propaganda.	Government uses social media sites to spread propaganda.

Table 1. Similarities and Differences Between Social Media in Cuba and Venezuela [3, 4, 5]

Methods

We are using a combination of qualitative and quantitative methods. These complementary methods will work in tandem through two phases: the qualitative work proposed in Phase 1 lays the groundwork for the machine learning and social network analysis proposed in Phase 2.

Phase 1: Interviews with Cuban and Venezuelan social media users.

In Phase 1, we are conducting qualitative interviews with citizens living in Cuba and Venezuela about their use of social media. In our preliminary fieldwork, we observed that citizens are surprisingly willing and eager

to chat online and share details about their lives. Authors 1 and 3 are bilingual Americans of Cuban descent and are familiar with the culture and customs of this population.

Phase 1 asks questions such as:

- What online sites do Cuban and Venezuelan citizens use?
- What do they post there?
- Do they self censor what they post?
- How do they receive information about their respective countries and the rest of the world?
- What issues concern them, and do they post about those issues on social media?

We are using purposeful, snowball sampling, recruiting initial participants through personal contacts and by approaching individuals who post online. Interviews are conducted primarily by text chat and email; however, audio and video is used when possible. In addition to interviewing participants, we are also examining their activity online to determine their frequency of use, the type of information they post, and the types of interactions they have.

Phase 2: Algorithmic categorization of organic and manicured accounts.

Based on the ground truth obtained in Phase 1, we propose to train algorithms to differentiate *organic* social media accounts (ones that reflect ordinary citizens' viewpoints) from *manicured* accounts (ones that recycle known government positions). We will use a mixed-methods quantitative approach that leverages social network, temporal and linguistic signals simultaneously in a single algorithmic framework.

Data.

The outcomes of Phase 1 will drive the sources and the types of data we collect. We will leverage custom software previously built by our labs to collect data from Twitter, Facebook, Reddit, Imgur, etc. In addition, we will collect documents, news reports and blog posts

known to contain or echo official government positions. These sources will be derived initially from interview informant responses; afterward, we will snowball sample from this initial set via recursively crawling hyperlinks to a broader set of documents representing official government positions.

Analytics.

With these two main data sources, we propose to apply analytics and machine learning to distinguish *organic* from *manicured* accounts. The first step in this process is separating the two types of accounts to establish a training corpus over which we can iteratively learn predictive features (i.e., co-training). In particular, we plan to apply state-of-the-art topic models and paraphrase detection (both of which are resilient to languages other than English) to the accounts we harvest to determine their similarity to known government sources. We will apply this model to the remaining accounts to iteratively develop new classifications of *organic* and *manicured* accounts, refining the model iteratively in each step. The overall contribution of this analytical approach is a summary of narratives told by ordinary citizens in repressive regimes versus those told by people acting in the government's interest.

Challenges

This research presents a number of important technical challenges. First, we have recently experimented with adapting standard topic models to handle social media, and this work will require us to apply, adapt and extend those models for the current context. We have elected to use *language-agnostic* machine learning techniques to cope with the multi-lingual environment in which these algorithms must operate. Second, the blind application of these algorithms to foreign contexts would likely result in discovering shallow or flawed findings. To address this risk, we have coupled our analytical, machine learning-based approach with domain experts, authors 1 and 3. They will regularly examine held-out predictions from these models,

thereby guiding the technical feature development and model construction.

References

1. Biddle, E. R. Rationing the Digital: The Politics and Policy of Internet Use in Cuba Today. *Internet Monitor Special Report Series No. 1*, The Berkman Center(2013), 1-13.
2. CNCB.com Staff. Cuba to Allow Wholly Owned Foreign Companies." *CNBC* (2014). <http://www.cnb.com/id/101515488>.
3. Francheschi-Bicchierai, L. The Internet in Cuba: 5 Things You Need to Know. *Mashable* (2014). <http://mashable.com/2014/04/03/internet-freedom-cuba/>.
4. Freedom on the Net: Cuba. *Freedom House* (2013). http://freedomhouse.org/report/freedom-net/2013/cuba#.U_eYpbyOy6Q.
5. Freedom on the Net: Venezuela. *Freedom House* (2013). <https://freedomhouse.org/report/freedom-net/2013/venezuela>.
6. Minaya, E. and Vyas, K. When Chávez Tweets, Venezuelans Listen. *The Wall Street Journal* (2012). <http://online.wsj.com/news/articles/SB10001424052702303990604577366123856023452>
7. Risen, T. Google Makes Internet Freedom Visit to Cuba. *U.S. News & World Report* (2014). <http://www.usnews.com/news/articles/2014/06/30/googles-eric-schmidt-went-to-cuba-for-internet-freedom-visit>.
8. Voeux, C. & Pain, J. Going online in Cuba: Internet under surveillance. *Reporters Without Borders* (2006). http://www.rsf.org/IMG/pdf/rapport_gb_md_1.pdf.