Lessons in Social Election Monitoring

Thomas N. Smyth
Sassafras Tech Collective
Ann Arbor, MI, USA
tom@sassafras.coop

Amanda Meng
Sam Nunn School
Georgia Tech
Atlanta, GA, USA
amandahmeng@gmail.com

Andrés Moreno, Michael L. Best
United Nations University
Institute on Computing and Society
Macao SAR, China
{amoreno, mikeb}@unu.edu

Ellen W. Zegura
School of Computer Science
Georgia Tech
Atlanta, GA, USA
ewz@cc.gatech.edu

ABSTRACT

Since 2011, our research group, along with numerous local partners, has been building a platform and methodology for monitoring elections using social media. Historically, election monitoring has traditionally been the domain of trained monitors provided by international monitoring groups. But monitoring by domestic groups with fewer resources has been a growing phenomenon, supported in part by the availability of inexpensive digital technologies such as SMS. Social media represents a further, exciting step in this trend. We describe our five years of experience in this endeavor and report a series of key lessons learned. These lessons touch on issues such as source types and curation, collaboration with other election-related groups, human vs. automated analysis, varying stakeholder needs, and the value of falsification. We also share our vision for the next five years of this research.

CCS Concepts

• Human-centered computing ~ Computer supported cooperative work • Human-centered computing ~ Social media • *Human-centered computing ~ Empirical studies in collaborative and social computing*

Keywords

election monitoring; social media; democracy; ICT4D; lessons learned; civic technology; civic participation

1. INTRODUCTION

Election monitoring has emerged as a standard democratic practice, and is especially common in jurisdictions with histories of questionable electoral integrity. The presence of monitors is intended to deter fraud and improve electoral processes. Meanwhile, reports issued by election monitoring groups, as well as stories of thwarted fraud attempts, are thought to help to build a public perception of a trustworthy election [1].

Election monitoring is traditionally the work of trained observers. Drawing on decades of experience, groups such as The Carter Center have distilled the practice into something of a science (e.g. [2]). Observers are trained using standardized curricula, deployed to geographical areas chosen based on statistical methods, and equipped with standardized surveys and purpose-built tools.

Meanwhile, the advent of social media has opened up a whole new opportunity in the practice of election monitoring. Regular citizens are empowered to post to social media about their own experiences on election day. Hash tags and discussion groups are developed and advertised for most major elections. Organized groups have also coalesced around such efforts. Social media campaigns encourage citizens to post and suggest particular things to watch out for and comment on. Such groups may also collaborate with security and election management bodies, and issue their own reports on the character of the election.

ICTD '16, June 03-06, 2016, Ann Arbor, MI, USA
ACM 978-1-4503-4306-0/16/06.
http://dx.doi.org/10.1145/2909609.2909640

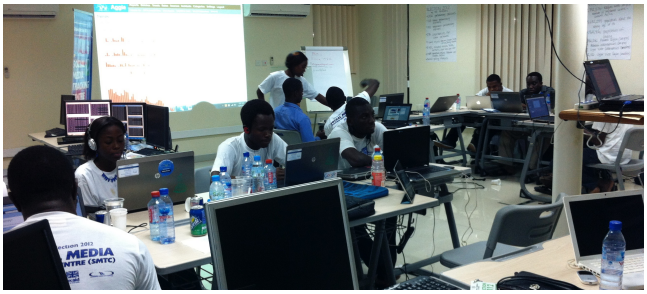

Figure 1. The SMTC (Social Media Tracking Center) for the 2012 Ghana elections. An early version of our Aggie monitoring application is projected on the wall.

Our research team of civic technologists from Georgia Tech and Sassafras Tech Collective has been involved with many such initiatives, beginning in 2011. We employ the term "social election monitoring" to describe the activity, in contrast to "formal" monitoring that is the purview of trained observer teams.

Informed by our experiences on the ground, we have been developing a platform and methodology for social election monitoring, consisting of a purpose-built software tool along with a set of human processes for harnessing the flood of data and citizen energy that is inspired by the electoral process.

In this note, we review related work and describe our practical experience thus far. We then review some of the lessons we have learned over our five years of work in this area, and discuss our vision for the way forward.

2. RELATED WORK

Even before the advent of social media, new digital technologies had begun transforming the field of election monitoring. Starting around 2006, numerous press articles, NGO reports, and blog posts discussed SMS-based election monitoring (e.g. [3]). Schuler [4] reported on the National Democratic Institute's use of SMS as conduit for observer data. The popular Ushahidi crowd-mapping tool has been used in many election contexts (e.g. [5]). In these early days, SMS proved to be an especially useful medium in countries with limited communications infrastructures due to its ubiquity, robustness, low bandwidth requirements, and low cost.

Social media as an election monitoring tool has received some attention in various literatures as well. Several empirical pieces have reviewed its use in different regions. Along with Ushahidi, Bani and Sgueo [6] mention both the *TwitterVoteReport* project in

the 2008 U.S. elections and a similar project for Bulgaria's 2011 polls. Fung [7] also touches on *TwitterVoteReport* in his description of MyFairElection.com, a crowdsourced election monitoring site built for the same 2008 U.S. elections.

As with SMS-based monitoring, Sub-Saharan Africa has seen plenty of activity in this area. In his 2013 dissertation, Smyth [8] compared social and formal election monitoring in two West African elections, finding that social monitoring methods made important contributions to the overall monitoring effort. Lazarus and Saraf [9] explored the integration of formal and social monitoring data during 2014 elections in Nigeria, calling the union a success and reporting lessons learned. Meanwhile, Sambuli et al. [10] reported on a post-hoc analysis of several million tweets collected during Kenya's 2013 elections, finding that all newsworthy events were reported on Twitter, usually well in advance of traditional media.

Crowdsourced election monitoring has also been the subject of some helpful theoretical analyses. Writing from a legal standpoint, Tuccinardi and Balme [11] consider crowdsourced election monitoring in the context of international obligations and agreements. For the Electoral Integrity Project (EIP), Grömping [12] proposes a causal model for election monitoring running from public perceptions of integrity and voter actions to official responses, deterrence, and electoral reform. He lays out an ambitious research agenda for the EIP, to study and validate various stages of the model. In a separate work, Grömping [13] also puts forth a set of six "mechanisms", including "accumulation of social capital" and "strengthening of the public sphere", through which election monitoring has the potential to positively impact democracy. This, too, is a useful theoretical lens for the analysis of initiatives such as the one described here.

3. PAST DEPLOYMENTS

As social media began to build critical mass West Africa, where SMS-based monitoring had previously been so prevalent, our lab at [institution] recognized the technology's obvious potential for election monitoring. Working with partner organizations, we took the first steps toward the social election monitoring platform described in this note. Since then we have helped monitor six elections in four countries. Table 1 lists all our deployments to date, including the approximate number of volunteers involved in the monitoring effort.

Table 1. Past uses of our social election monitoring platform.

Location/Year	Volun-teers	Dates	Reports Crawled	Incidents Created
Nigeria 2011	15	Apr 9, 16, 29	398,000	262
Liberia 2011	10	Oct 10, Nov 7	31,500	13
Ghana 2012	35	Dec 7	321,000	358
Nigeria 2014 (Ekiti State)	10	Jun 21	92,400	87
Nigeria 2015	25	Mar 28-29, Apr 11	2,668,000	249
Argentina 2015	10	Nov 22	400,000	60

The 2011 Nigeria deployment was our first foray into this area. We were contacted by the Nigerian youth democracy group "EnoughIsEnough" for help in developing software tools to monitor the extensive amount of election-related Twitter traffic in that country. Aggie, our social media monitoring tool, and the first "Social Media Tracking Center" (SMTC) were the results of this collaboration [14].

Aggie is a web-based application that enables monitoring of large volumes of social media traffic from multiple types of sources including Twitter, Facebook, RSS feeds, and ELMO (an election monitoring system) instances. Figure 2 shows the flow of data through the system. "Reports", or raw items of social media content fetched from the Internet, are scanned by a combination of manual and automated means, and pertinent reports are highlighted. These are organized into trackable "incidents" for further analysis and monitoring. Current versions of Aggie are built for scale, able to handle thousands of reports per minute.

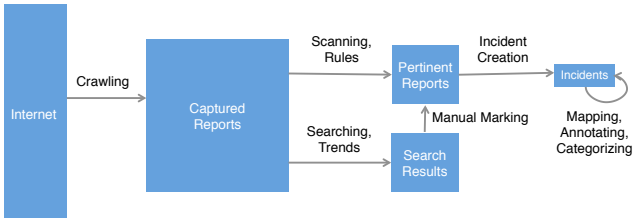

Figure 2. The Aggie workflow.

Our monitoring efforts generally take the form of a "social media tracking center" (SMTC), which is a physical space in which volunteers gather, around the clock, to monitor social media traffic via Aggie and other tools. Figure 1 shows the SMTC from the Ghana 2012 election, while Figure 3 depicts the typical SMTC structure and information flow. There are usually three teams. The relevance team monitors trends, scans reports, and creates incidents from those deemed worthy of further attention. The veracity team verifies these new incidents and passes confirmed incidents to the escalation team, which works with the relevant authorities to ensure appropriate action is taken.

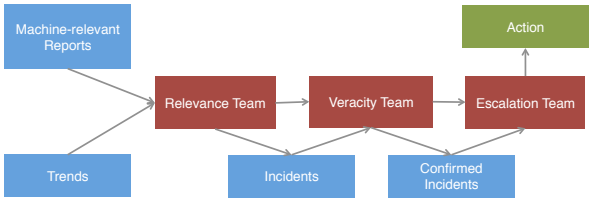

Figure 3. The flow of data and information within the SMTC.

Prior to election day, SMTC preparations are also intricate. They typically include identifying local partners in appropriate sectors (election management, formal observers, law enforcement, etc.), soliciting volunteer staff members, establishment of embedded agents at partner organizations, procurement of physical space and computing infrastructure, training, and software testing.

Our first SMTC, in Nigeria 2011, was widely considered a success. As described by Best [15], the SMTC was routinely "out front of both national security forces and traditional media in identifying and initiating responses to electoral irregularities and conflict."

The SMTC in Ghana 2012—the best-staffed to date at 35 volunteers—was also effective. Meng [16] reports several instances where the SMTC was able to alert authorities of developing situations (e.g. unrest at a polling station, missing election materials, ballot box burning) picked up from social media.

The 2014 deployment in Nigeria's Ekiti state gubernatorial election was notable in that it was our first attempt to incorporate data from formal monitoring operations via the ELMO formal

monitoring tool. Lazarus and Saraf [9] report on this effort, citing multiple successes and challenges.

Most recently, we collaborated with Pol-it, an Argentinean election technology startup, to deploy Aggie in that country's presidential run-off election. This was the first Aggie deployment to integrate reports from a crowdsourced election monitoring application—a web application developed by Pol-it called *Ojo con el Voto*. App users filled out surveys of their voting experience and reported on election logistics as well as irregularities. Aggie received 14,000 reports from *Ojo con el Voto* users in addition to 400,000 social media reports.

Other groups unaffiliated with [institution] or [business] have also used Aggie, which is open source, for election monitoring deployments.

4. KEY LESSONS LEARNED

Over our past five years of experience in social election monitoring, a handful of key lessons have emerged as pivotal to an effective deployment. We discuss them in this section.

Sources vary in prevalence by region.

After our first deployment in Nigeria, it might have seemed sensible to focus the system on Twitter reports only, since the social media scene in Nigeria at the time was dominated by Twitter, and nearly all the captured reports were from that platform. In Liberia, however, Twitter was nearly unheard of, and most of our data came from Facebook, though the total volume was much lower. It is thus important to be prepared for different source mixes in different contexts.

A corollary to this is that the system must be able to handle vastly different volumes of reports, and different levels of signal vs. noise, from different source types. Twitter tends to produce a high level of noise due to the imprecision of global keyword searches. On the other hand, since Facebook and RSS feeds are targeted to specific topical groups or pages, the level of noise is far less. Aggie must prevent one from overwhelming the other. This is one of the main design challenges in the system.

Source curation is essential and best done by local experts.

Curation—by which is meant researching, selecting, and testing—of data sources for the app is an intricate task that should be begun well before election day. The involvement of local experts with deep experience in the local social media scene is imperative. Testing is also important, especially with Twitter keyword sets. Too broad a net results in the monitoring staff being overwhelmed. Too narrow and key information may be missed. We typically set up a test instance of the system some days before the election to get an idea of the report volume our source set might generate. A notable case in Ghana was the inclusion of the northern city name "Tamale" in the keyword set, which matched a large volume of unrelated tweets about the Mexican food item of the same name.

A fourth verb could also be added to the above definition of curation—source development. This entails working with the social media community to encourage reporting by the citizenry on and around election day. Such reporting is typically organized around a common venue, such as a hashtag (e.g. #nigeriadecides) or a Facebook group. In Nigeria 2015, the election commission even requested that tweets include the polling unit code (e.g. "PU 014 hardo dan iya ward") associated with a reported incident. This work may be done by the SMTC itself or by partner groups, but it should certainly be done.

To employ Grömping's taxonomy [12], this last endeavor broadens the engagement from "passive" crowdsourcing, in which existing public communications are merely observed, to "open" crowdsourcing, in which a call is sent out to the general public soliciting reports.

If this more interventionist path is chosen, establishing a reputation for the SMTC's own social media identities is also helpful. Monitors in past deployments have reported a greater willingness of the public to cooperate when they are already familiar with the SMTC.

There is great potential in social–formal monitor collaboration.

The Ekiti 2014 election was the first in which Aggie was set up to collect data from formal monitoring sources. A post-hoc analysis revealed that this relatively small amount of data seems to have been overwhelmed by the far greater volumes of social media data, and went largely overlooked [9]. However, an unexpected benefit of the collaboration effort was that SMTC staff ended up in frequent phone contact with formal observers on the ground, thanks to a phone number list that was distributed to SMTC staffers in advance of the election. Analysis of Aggie incident records reveals that this triangulation work led to a high percentage of incident closures or resolutions.

That the formal observer data was overlooked reinforces the salience of the above lesson on source prevalence and volume—clearly more design work is needed in this area. Additionally, though, the resourcefulness of the SMTC staffers in resorting to phone communication for veracity checking remains entirely appropriate and is a key example of the type of human process that we intend to include in our platform.

Human and automated analysis are both essential.

The sheer volume of social media data—even with first rate source curation—clearly calls for some measure of automated analysis. We have experimented with several aggregation and visualization techniques based on basic keyword matching. We have also run post-hoc analyses using a specialized algorithm for automated topic extraction [8, p. 102]. Further experimentation with natural language processing techniques is high on our research agenda. However, we believe that a high level of human involvement will always be required in most stages of the system.

Donor needs may conflict with monitoring goals.

Since election monitoring work is not self-sustaining, cooperation with donors is an inevitability. Furthermore, donors often have distinct sets of needs that may be at odds with a strict focus on monitoring effectiveness. For our team, this came to bear on the design of Aggie visualizations. The donor's desire for an aesthetically pleasing visualization well suited to print media did not always align with the design team's focus on efficiency.

Due to the nature of the field, this naturally divergent set of needs should be accounted for early in the design process. We believe that solutions satisfying both parties are achievable in this case.

Trusted embeds make action possible.

An SMTC tends to operate as part of a system of election-focused civic organizations, including the elections management body, security forces, formal monitoring groups, traditional media outlets, and others. We have found that locating SMTC staffers on the premises of these groups has been indispensable to the SMTC's ability to act on information gleaned from monitoring.

To take one example from the Ghana 2012 deployment, an SMTC embed at the Ghana Police headquarters was able to bring about swift action and resolution in response to an incident of unrest at a polling station. The incident was previously unknown to the police. Given the hectic nature of the election day, the team

believes that such prompt action would not have been possible without a physical presence at the headquarters building.

Placing embeds requires diligent planning efforts well in advance of election day and cooperation with local groups who can help establish the trust networks needed to gain entry.

Falsification is important work.

In Ghana 2012, the monitoring team reported that "many [social media] reports of ballot snatching have been falsified by our counterparts at CODEO [a formal monitoring group] and police headquarters." While verification and escalation of reported incidents may seem to be the primary and most rewarding work of the SMTC, we believe that falsification—and especially promulgation of the same—of reports is also crucial to establishing an accurate public perception of the election.

Many false reports of malfeasance may create an inaccurate perception of the election, with deleterious effects on the nation's democracy. By diligently looking into reports of foul play and sharing findings with the community, the SMTC can work against this effect and help create a more accurate perception. The SMTC is uniquely positioned to do this work.

5. FUTURE DIRECTIONS

Like any thriving software project, the feature request list for Aggie is long and growing. As we continue to refine the system, we are also enthusiastic about several new directions.

The first is a public face for Aggie. Though it is a Web-based application, to date it has been accessible to SMTC staff members only. We are interested in exposing a subset of Aggie's data to the public via a dashboard-style interface. While there are ethical concerns—we do not want Aggie to become a clearinghouse for malicious reports, for one—we feel that a carefully composed public view of Aggie's data could be interesting and beneficial. Interestingly, this adds a third stakeholder in the system design. Aside from monitors and donors, the public presents its own set of needs, such as easy digestibility.

A second goal is to greatly extend the automatic processing capabilities of Aggie. Given the availability of domain experts to hand-label a subset of report data as it arrives, we are especially interested in real-time supervised machine learning techniques (e.g. [17]). Topic extraction methods (e.g. [18]) are also compelling. Aggie would be the first tool that we know of to apply advanced techniques such as these to a real-time social media analysis application.

Finally, we are deeply interested moving beyond elections by using Aggie as a means to more broadly engage the public. In a pathbreaking report on technology and citizen participation, the National Democratic Institute found that the act of crowdsourcing alone is not enough to bring about increases in citizen political influence [19]. They found that only when technology was used to bring citizens into deeper political activities, such as education and contact with public officials, did it have more political impact.

Citizens making election reports on social media represent a self-selected group of individuals interested in their country's democracy. Once identified through Aggie, social media could be used to invite such citizens into deeper political involvement programs crafted by local partners. Survey data could also be collected at this point, furthering research goals around technology and civic participation. The heightened energy surrounding an election could be channeled into the longer-term political activities where so much of the important work of democracy is done. Aggie and the SMTC methodology are in a unique position to achieve this goal.

6. ACKNOWLEDGMENTS

We are deeply grateful to our partner groups, countless SMTC volunteers, and the support of the MacArthur Foundation.

7. REFERENCES

[1] S. D. Hyde, *The Pseudo-Democrat's Dilemma: Why Election Monitoring Became an International Norm*. Cornell University Press, 2011.

[2] D. J. Carroll and A. Davis-Roberts, "The Carter Center and Election Observation: An Obligations-Based Approach for Assessing Elections," *Election Law J. Rules Polit. Policy*, vol. 12, no. 1, pp. 87–93, 2013.

[3] K. Verclas, "Texting It In: Monitoring Elections With Mobile Phones," MobileActive.org, 2007.

[4] I. Schuler, "SMS As a Tool in Election Observation," *Innov. Technol. Gov. Glob.*, vol. 3, no. 2, pp. 143–157, 2008.

[5] C. S. Bailard and S. Livingston, "Crowdsourcing Accountability in a Nigerian Election," *J. Inf. Technol. Polit.*, vol. 11, no. 4, pp. 349–367, 2014.

[6] M. Bani and G. Sgueo, "We-Transparency," in *Human Rights and the Impact of ICT in the Public Sphere*, C. M. Akrivopoulou, Ed. IGI Global, 2014, pp. 156–174.

[7] A. Fung, "Popular Election Monitoring," in *Race, Reform, and Regulation of the Electoral Process*, C. E. Guy-Uriel, H. K. Gerken, and M. S. Kang, Eds. Cambridge University Press, 2011, pp. 192–208.

[8] T. N. Smyth, "Social media, elections, and democracy in West Africa," Dissertation, Georgia Institute of Technology, Atlanta, Georgia, USA, 2013.

[9] P. O. Lazarus and T. A. Saraf, "Social Media Fostering Social Good: A Case of Election Monitoring in Nigeria," in *Proc. ICTD*, New York, NY, USA, 2015, pp. 47:1–47:4.

[10] N. Sambuli, A. Crandall, P. Costello, and C. Orwa, "Viability, Verification, Validity: 3Vs of Crowdsourcing," iHub Research, 2013.

[11] D. Tuccinardi and F. Balme, "Citizen Election Observation Towards a New Era," *Election Law J. Rules Polit. Policy*, vol. 12, no. 1, pp. 94–99, Mar. 2013.

[12] M. Grömping, "Domestic monitoring," *The Electoral Integrity Project*. [Online]. Available: https://sites.google.com/site/electoralintegrityproject4/projects/election-monitoring/crowdsourcing. [Accessed: 27-Jan-2016].

[13] M. Grömping, "Many eyes of any kind? Comparing traditional and crowdsourced election monitoring and their contribution to democracy," in *Proc. Int'l Relations and Development*, 2012.

[14] J. B. Asuni and J. Farris, "Tracking Social Media: The Social Media Tracking Centre and the 2011 Nigerian Elections," Shehu Musa Yar'Adua Foundation, 2011.

[15] M. L. Best, "Peacebuilding in a Networked World," *Commun ACM*, vol. 56, no. 4, pp. 30–32, 2013.

[16] Meng, Amanda, "Tracking Social Media: The Social Media Tracking Centre (SMTC) and the 2012 Ghana Elections," Georgia Institute of Technology, 2013.

[17] G. Carpenter, S. Grossberg, and J. Reynolds, "ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network," *Neural Netw.*, vol. 4, no. 5, pp. 565–588, 1991.

[18] B. O'Connor, M. Krieger, and D. Ahn, "TweetMotif: Exploratory Search and Topic Summarization for Twitter," in *Proc. ICWSM*, 2010.

[19] National Democratic Institute, "Citizen Participation and Technology: An NDI Study," 2014.