Journal of Mason Graduate Research
Volume 5 Number 2 Year 2018

ISSN: 2327-0764
Pages: 85-107

The Geography of Anonymous Communications: Predicting Escalation of Anonymity Networks During Events of Civil Unrest

BRIAN SANDBERG

George Mason University

Civil unrest can trigger uptake of anonymous communication to protect user identity and location or to circumvent censorship. Anonymity networks such as the Tor Network can support planning, orchestrating, or responding to protest events. This research aimed to understand this relationship between protest events and Tor usage. A methodology was developed to discover the best supervised learning method for predicting Tor usage in response to protest events. Twelve classification algorithms were evaluated using data representing the different conflict event types, number of events per day, source and target actor categories, fatalities, and Tor usage. Experiments were conducted using over five years of conflict event data from nine countries selected from the Africa-based Armed Conflict Location and Event Dataset (ACLED). This research produced unique quantitative results predicting Tor escalation during conflict events with an F1 Score of over 86%. Results are significant given the multitude of use cases for Tor, with the strongest signal occurring in authoritarian regimes.

Keywords: crisis event data, ACLED, anonymous communications, anonymity networks, tor usage metrics, supervised machine learning, classification methods

INTRODUCTION

The geography of cyberspace changes the traditional analytic landscape of the natural and built environment to a virtual environment. As with geographic analysis in physical space, analysis of cyberspace is equally important to understand patterns of activity and behavior that are emerging, evolving, or changing over time (Lee & Chan, 2016). In cyberspace, many geographic-based assumptions and estimates no longer hold true. Patterns and relationships can emerge independent of geographic location. For instance, the idea that near things are more related than distant things has been a core concept in geographic analysis (Tobler, 1970). Transportation systems, communication systems, and the Internet have altered this assumption and have created new challenges and opportunities for spatial analysis. Janelle (1969) described how innovations in transportation systems had a major impact on human movement and interaction. Cyberspace further removes the dependence on distance and geographic space for connections and collective behaviors to converge. Social cliques can include participants with similar interests from diverse geographies (Nicholls Walter, 2008). Social networks and cybercrime networks bring together participants from distant places (Leukfeldt, Kleemans, & Stol, 2017). Periods of civil unrest unite participants from near and far, and in these cases, the physical and virtual worlds intersect (Braha, 2012).

Political events, including protests and riots, are supported via social networks, mobile communications, and anonymization tools. It is often desirable to measure these virtual activities to understand patterns relating to real world events. For example, brand marketers are interested in online consumer trends and habits to understand audience engagement and advertising effectiveness. Governments and militaries need to understand how social networks are used by terrorist groups to facilitate propaganda, radicalization, and recruitment activities (Klausen, 2015). In this latter case, actors often obscure their identities to conduct online activities anonymously. There are also legitimate reasons to obscure online identity, including digital privacy and consumer protections to avoid scams and identity theft (Rika Butler, 2007). Military, law enforcement, investigative journalists, and political dissidents also have strong motivations to use anonymizing technology to protect their informants, sources, and their own identity. The exclusion of user identity and physical location can present some challenges when studying the geography of anonymous communication. While analyzing individual usage patterns of anonymity networks is generally not possible, it is feasible to analyze and draw conclusions on aggregate behaviors of anonymous users.

On a daily basis, various news media report on different types of civil unrest occurring throughout the world. These collective action events range from peaceful protests to violent uprisings. Every day people are also using anonymization technology while communicating and sharing information online to hide their real identity, location, and activity. The Tor Network is a large anonymity network operated by a group of volunteers worldwide and uses virtual tunnels to improve privacy and safety on the Internet. This study analyzes and measures patterns of collective behavior involving the use of anonymity networks to facilitate social movements. Through this research, I aimed to determine if there is a relationship between protest events and the use of anonymization technology. In other words, what is the effect of civil unrest on the use of anonymizing tools? The hypothesis is that protests increase the use of anonymizing tools and the null hypothesis is that no significant change in usage patterns exists prior to or during protests. In addition, this research asks if this relationship is more pronounced in authoritarian regimes where public protests may be illegal or subject to military response. The goal was to produce quantitative analytic results that demonstrate that technologies like the Tor Network are used for important societal purposes such as opposing repressive governments or circumventing censorship, rather than solely criminal activities such as terrorist recruiting, drug marketplaces, and child abuse material exchange.

LITERATURE REVIEW

The literature review covers the evolution of crisis event databases and the improved quality to enable the type of research conducted in this study. The review continues with the affordances of technology, such as mobile communications and social media, to civil unrest and protest event, and how computers, the Internet, and particularly, anonymity networks have provided support to protests and social movements.

Crisis Event Data and Analysis

Development of systematic, geographical disaggregated event datasets have helped foster progress in conflict research (Gleditsch, Metternich, & Ruggeri, 2014). Crisis event data has evolved over the years, and specialized databases now exist for social and political science research for a variety of hypothesis testing. These datasets record a *who did what to whom* scheme within the context of political actions. This observational data is a record of individual events between a source actor and target actor and provides a disaggregated view of political events. Modern automated machine coded datasets, such as Integrated Crisis Early Warning System (ICEWS), Global Database of Events, Language, and Tone (GDELT), and Phoenix, are based on news sources and provide insights into what has been reported and collected (Arva et al., 2013). There are several idiosyncratic aspects of these datasets that complicate analysis. A trivial event can attract huge media attention, while major political events may not be reported adequately. Reports may also be false, incomplete, contain duplicates, or have incorrect coding. The Armed Conflict Location and Event Dataset (ACLED) provides a more focused, geographically disaggregated dataset for analyzing protest events and provides a more accurate specification for the locations and dates of the observed events (Raleigh, Linke, Hegre, & Karlsen, 2010).

Most research using event data has focused on monitoring and mining content for emerging trends, and processing trends into forecasts of political events. Quantitative analysis of crisis event data has also focused on predicting event onset. Ramakrishnan et al. (2014) describes how Early Model Based Event Recognition using Surrogates (EMBERS) outperforms existing methods in event forecasting of civil unrest using open source indicators (OSI) from social media, news events, blogs, and economic indicators. Muthiah (2014) provides a more detailed investigation into this EMBERS framework, while Agarwal and Sureka (2015) provide a review of related literature. Similar studies have analyzed the geography of civil unrest. Buhaug and Gates (2002) analyze geographical factors that determine the scope of conflict and the location of the conflict relative to the capital. Sensitivity analysis of empirical results on conflict onset is covered in Hegre and Sambanis (2006), with findings that confirm large population and low per capita income increase risk of civil war. Schutte and Donnay (2014) present a causal analysis that reveals that Iraqi civilians actively supported US military in reaction to indiscriminate insurgent violence.

Affordances of Technology to Collective Behavior

Protests are a specific class of conflict events in which cyberspace has played a significant role. Information sharing and communication technologies are critical to planning and organizing social movements. Different forms of communication, including social media and anonymous communications, can provide unique benefits in facilitating these events. Yuan (2017) discussed the use of mass media and location-based social media data and shows a strong distance decay effect on connections and interactions between Chinese provinces. Bastos, Recuero, and Zago (2014) investigated the relationship between onsite and online protesting activity, concluding that online participants are geographically distant from street protesters. Conover et al. (2013) examined how the goals of a protest movement are reflected in the geographic patterns and information sharing practices of its communication network. Myers (1994) examined the contribution of computer networks to the formation and functioning of social movements and collective behavior.

Another relevant area of research aims to understand the processes of activist computer use and the results for social movements. Bennet (2003) addressed the capacity of digital media to change the political game and how the Internet is implicated in the new global activism. The Internet does not just reduce the costs of communications, but rather transcends the geographical and temporal barriers associated with other communication media. Bennet et al. (2014) focused on crowd peer production and how crowd-enabled networks are activated, structured, and maintained. Garrett (2006) surveyed how new information and communication technologies are changing the ways in which activists communicate, collaborate, and demonstrate. Massa (2016) discussed the role and growing influence of online communities and how they support faster, cheaper, and more flexible organizing. Massa (2016) also identifies the lack of empirical studies of online communities and how they become agents of social change. Lacking from these studies is how participants in social movements can benefit from anonymous communication and how privacy tools can help protect their identity and circumvent censorship.

Anonymity technologies, such as the Tor Network, can provide critical benefits to participants during social movements. According to the Tor Project, dissidents are advised to use Tor to ensure their privacy and safety. However, the use of these tools during periods of civil unrest has not been studied from a quantitative perspective. While there has been extensive research relating to the contribution of computer networks and social media to activism, there has been almost no quantitative research on the role and contribution of anonymity networks to social movements. One example, is Jardine (2016) who asked why people use anonymity-granting technologies when surfing the Internet, hypothesizing that people use online anonymity services to evade repressive regimes infringing on individuals' civil and political rights. The results of this study concluded that Tor usage is driven by political repression (e.g. avoiding surveillance and circumventing censorship) as well as highly liberal contexts (e.g. free and easy access to technology). Another study by Rady (2013) showed that anonymity networks can become terrains for government-population conflict as they enable citizens to overpower governments' conventional control mechanisms over cyber-information exchanges. Sandberg (2017) developed anomaly detection and similarity methods that used Tor client and Twitter usage data to detect country-level anomalous behavior and identify similar patterns across multiple countries. This research provided novel techniques to detect both censorship and protest events and leveraged social media data to help explain the occurrence of events.

DATA PREPARATION

The Armed Conflict Location and Event Dataset (ACLED) is a comprehensive daily collection of political and protest events that focus on the African continent[1]. The data represents a high quality, human vetted source of conflict events that captures the most accurate dates, location information, and related details. At the time of this writing, ACLED only provided data for Africa, but they have plans to provide similar data for the Middle East and Asia. A variety of factors contributed to the selection of the nine African countries used in the analysis. The countries were selected based on large populations, high Internet penetration rates[2], and the degree of freedom based on the Freedom House Index (Freedom House, 2018). An even distribution of *free*, *partly free*, and *not free* countries were used in the study (Freedom House, 2018). Freedom indicators supported the determination of whether there existed a stronger signal of Tor usage in authoritarian regimes. Countries were also vetted for having an active history of protest events and that the events occurred over the time period of the study.

Africa is categorized into five regions including Eastern, Central, Northern, Southern, and Western. All regions except the Central region were represented in the analysis. The Central region lacked countries with appropriate levels of Internet penetration to make them viable for use in this analysis. Countries were only included if they had at least a twenty-five percent Internet penetration rate to ensure that the target variable (i.e. Tor escalation) for the classification task provided an adequate concentration of Internet users. Three countries from each freedom index rating were selected. Table 1 shows the nine countries that met all criteria for inclusion. The time period for the data used in the analysis covered September 1, 2011

[1] Event data was collected from the ACLED website https://www.acleddata.com/
[2] Population and Internet statistics were collected from Internet World Stats website https://www.internetworldstats.com/stats1.htm

through December 31, 2016.

Table 1

Selected Countries Ranked by Internet Penetration

Country (ISO code)	Region	Population (2017)	Internet Users (2017)	Penetration Rate (%)	Freedom Index
Kenya (ke)	Eastern	48,466,928	37,718,650	77.8%	Partly Free
Morocco (ma)	Northern	35,241,418	20,207,154	57.3%	Partly Free
South Africa (za)	Southern	55,436,360	28,580,290	51.6%	Free
Tunisia (tn)	Northern	11,494,760	5,800,000	50.5%	Free
Nigeria (ng)	Western	191,835,936	93,591,174	48.8%	Partly Free
Libya (ly)	Northern	6,408,742	2,800,000	43.7%	Not Free
Algeria (dz)	Northern	41,063,753	15,105,000	36.8%	Not Free
Egypt (eg)	Northern	95,215,102	34,800,000	36.5%	Not Free
Ghana (gh)	Western	28,656,723	7,958,675	27.8%	Free

The ACLED database provides detailed codes for the specific type of event, the source and target actors involved in the event, and records with date and geographic coordinates (Raleigh et al., 2010). The information was highly curated and manually coded providing a reliable, thorough, and usable source of conflict event data; a list of event types and actors included in the ACLED database are listed in Table 2.

Table 2

ACLED Event and Actor Categories

Event Type	Actor Categories
Battle-Government regains territory	No Actor
Battle-No change of territory	Government
Battle-Non-state actor overtakes territory	Rebel force
Headquarters or base established	Political militia
Non-violent transfer of territory	Ethnic militia
Violence against civilians	Rioters
Remote violence	Protesters
Riots/Protests	Civilians
Strategic development	Outside/external force

Table 3 shows the ACLED events and median tor usage by countries in this analysis. This analysis used the nine different ACLED event types and eight actor categories as shown in Tables 2. Actor categories were included for both the source actor (perpetrator) and the target actor (victim).

Table 3

ACLED Events and Median Tor Usage by Country

Country	Total Events	Riot & Protest Events (%)	Tor Usage (Median)
Egypt	6900	3801 (55%)	4581
Nigeria	6766	2438 (36%)	1541
South Africa	6480	5433 (84%)	5688
Libya	4481	789 (18%)	681
Tunisia	2548	1962 (77%)	2229
Kenya	2195	955 (44%)	927
Algeria	1707	1049 (61%)	2833
Morocco	708	512 (72%)	3026
Ghana	368	264 (72%)	759

The number of events recorded in ACLED can vary from zero to dozens on any given day. For analysis purposes, an equally measured time series dataset was created with the temporal unit of a single day. The data was prepared to minimize any loss of information. To accomplish this, the discrete categorical event and actor features were numerically coded and then binary encoded with a new Boolean feature for each event type and actor category. Only one of the features takes on the value of ‘1’ for each sample, and the transformed data becomes a higher dimensional sparse matrix. This method is typically referred to as one-hot encoding and is performed to improve a dataset for machine learning. Because there can be multiple events occurring on a single day, a single record per day was created by summing the event types and actor categories to create a count value for each of these features. Days without any events occurring were recorded with zeros for all features. This approach collapsed the raw ACLED data into a single instance per day with a count of the total number of events, a count for each event type, a count for each source and target category, and a count of fatalities. Together, these event features provide a representation of conflict behavior for each day in each country. The next step was to prepare the target variable based on Tor usage.

The Tor network is the most mature and largest deployed anonymous communication network that conceals the user's location, identity, and activity. The entire public Tor network includes over 7,000 relay nodes with varying levels of bandwidth. Relays and directory authorities publish relay descriptors so that clients can select relays for their paths through the Tor network. Tor relay nodes perform data collection services throughout the Tor network. Tor usage metrics are collected by the Tor relay nodes and country-level aggregate data is provided by the Tor Project for the research community[3]. There are several data layers provided by the Tor network to support client usage metrics, including the relay server descriptors, relay extra-info descriptors, and network status consensus.

Counting client users in the Tor network is somewhat challenging since collecting client information could reduce the anonymity and safety of the user. Therefore, the data on Tor usage metrics are only available at the aggregate level. Country-level usage statistics provide client counts representing the number of unique users connecting to the Tor network on a daily basis. Client usage counts for each of the selected countries were extracted from the Tor user metrics. New rows were created for each missing day and the missing values were filled with the previous day's known value. Only four days of Tor usage data was missing over the entire study period. Tor escalation models were developed to measure whether Tor usage was increasing or decreasing on a daily basis. Two escalation methods were developed including a binary escalation model, where a '1' indicates an increase in usage or no more than a 2% decrease from the previous day, and a '0' indicates a decrease of more than 2% from the previous day. The second method used percent change with a split assignment giving all percent changes below the mid-quartile a value of '0' and all values above mid-quartile a value of '1'. Class labels (target variables) were created for each of these escalation models plus additional class labels that captured time lags of one and two days in both directions. This resulted in a total of ten class labels that could be used by the classification methods described in the Methodology section. The set of targets provide a five-day window for analyzing temporal behaviors in Tor usage in relation to protest events. They capture behaviors that may be occurring prior to an event or behaviors in response to an event.

[3] Tor usage metrics were provided by the Tor Project https://www.torproject.org/ and Tor Metrics https://metrics.torproject.org/

Visualizing data helps to understand the relationship between ACLED events and Tor usage. Time series plots are useful for visually displaying this relationship. Figure 1 shows the time series plots for the entire study period for all nine African countries. The plots show the relationship between ACLED events and Tor usage for free countries (Tunisia, Ghana, South Africa), partly-free countries, (Morocco, Nigeria, Kenya), and not free countries (Egypt, Libya, Algeria).

It can be difficult to see clear patterns while visualizing the full time period of the study. It is often helpful to visualize sample periods such as those shown in Figure 2. The plot on the left shows a random sample from the South Africa dataset covering a two-month period of September 1, 2016 through October 31, 2016. The plot on the right shows the behaviors exhibited during the 2013 Egyptian protests, which took place over a three-day period of June 30, 2013 through July 3, 2013. The goal of this protest was to overthrow Mohamed Morsi and suspend the constitution. The sample periods reveal a markedly similar temporal pattern between ACLED events and Tor usage patterns. This study aims to determine if this type of pattern generally exists for the different countries, and if it is possible to accurately predict Tor usage during protest events.

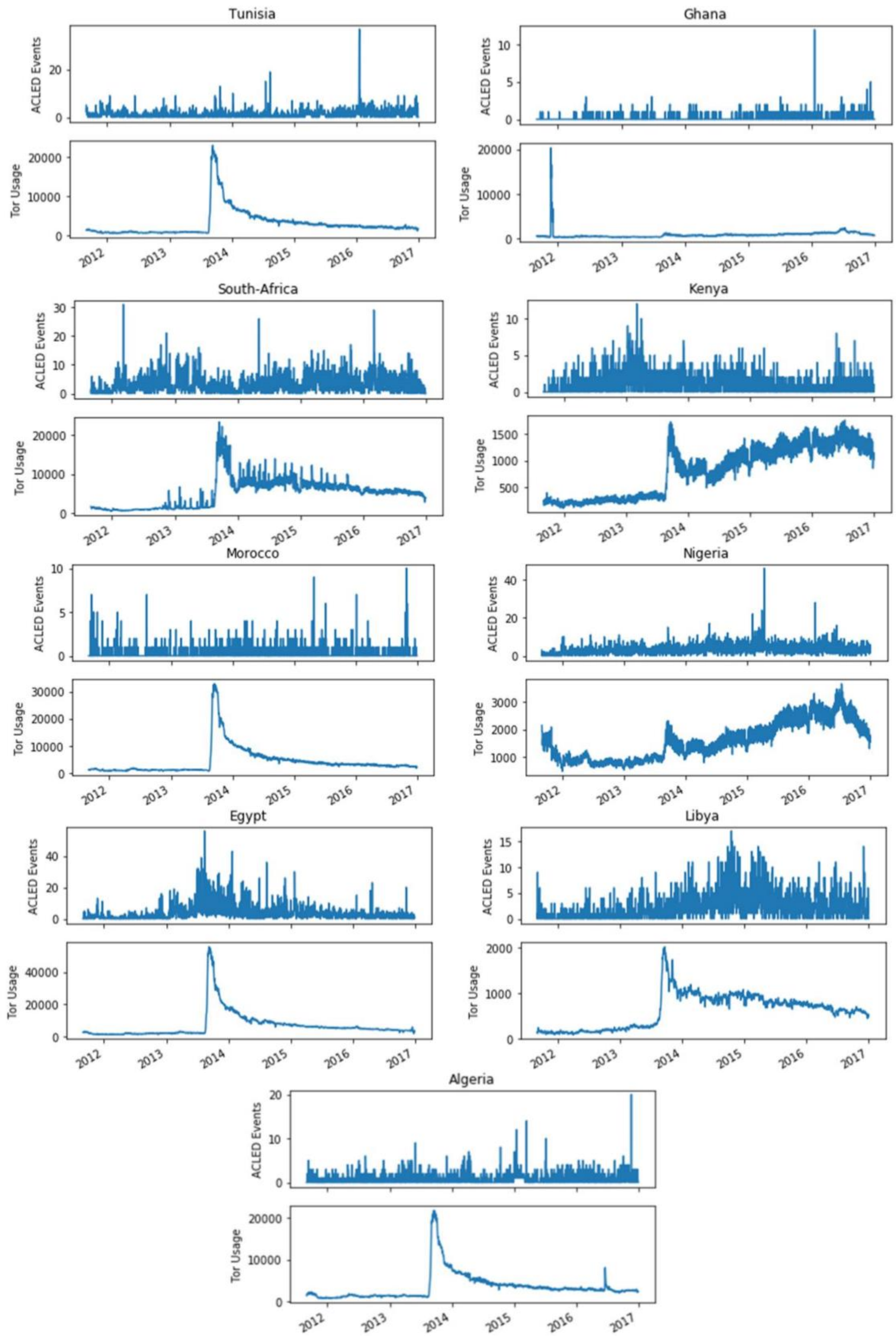

Figure 1. Time series plots for ACLED events and Tor usage.

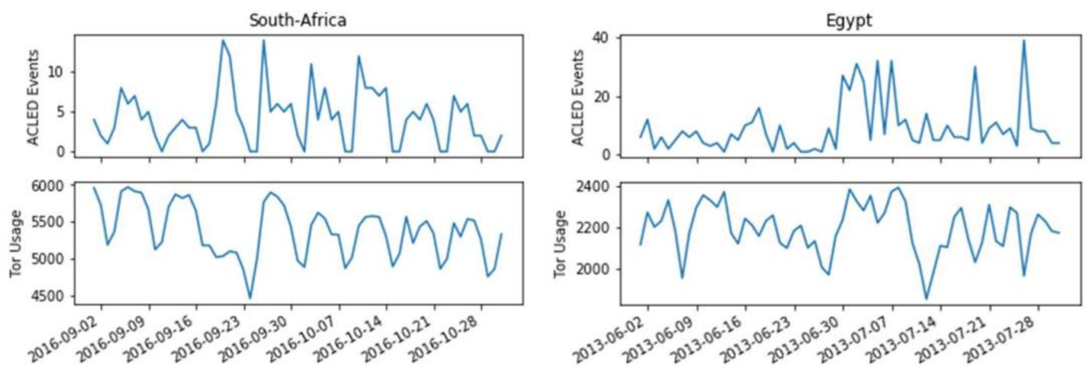

Figure 2. Sample time series plots. The sample plots show patterns of interest when evaluating the relationship between ACLED events and Tor Usage.

The last step in data preparation process included finding spatial clusters of events and using the resulting cluster labels as new features characterizing events. Spatial clustering was used to find spatial groupings for the event data on a per country basis, with the number of clusters represented by the variable K. The K-means algorithm was used to iteratively assign each data point to one of K groups based on the geographic coordinates provided for each event (Arthur & Vassilvitskii, 2007).

The user of the K-means algorithm is required to select the value for K, which is the number of clusters to group the data. This can be difficult to do since there is no easy way to select the best K. One option is to use metrics on the clusters to automatically find the best value for K. The silhouette coefficient measures mean intra-cluster distance (m) and mean nearest-cluster distance (n) for each sample and then calculates a single metric using the formula (n - m) / max (m, n) (Rousseeuw, 1987). The silhouette defines how well clusters are fit by the model and ranges from -1 (poorly matched to neighboring clusters) to 1 (well matched to its own cluster). This metric was used to iteratively find the optimal value for K when fitting the K-means clustering algorithm to each country dataset.

Map visualizations help to understand the resulting clusters, which include the geographic extent and spatial relationships among conflict events. Again, a sample is used to visualize resulting clustering behaviors. In Figure 3, the map shows the spatial patterns and grouping of events into clusters for South Africa. The map displays five spatial clusters with the distribution of events per cluster. An additional overlay shows the distribution of the events that occurred over the sample time period from Figure 2. This sample distribution is spread across all clusters indicating that this particular slice into the data does not have a spatial-temporal pattern consistent with a particular geographic cluster.

New features were derived from the spatial clusters and used to determine if spatial information can improve prediction accuracy of the classifiers. The labels resulting from the spatial clustering were added to each event in the datasets, applying the same one hot encoding process used for event types and actor categories described above. Figure 4 displays the entire set of features as a sparse matrix time series. The features include the total number of events that occurred each day, along with the count of the event types, source actors, target actors, fatalities, and cluster labels.

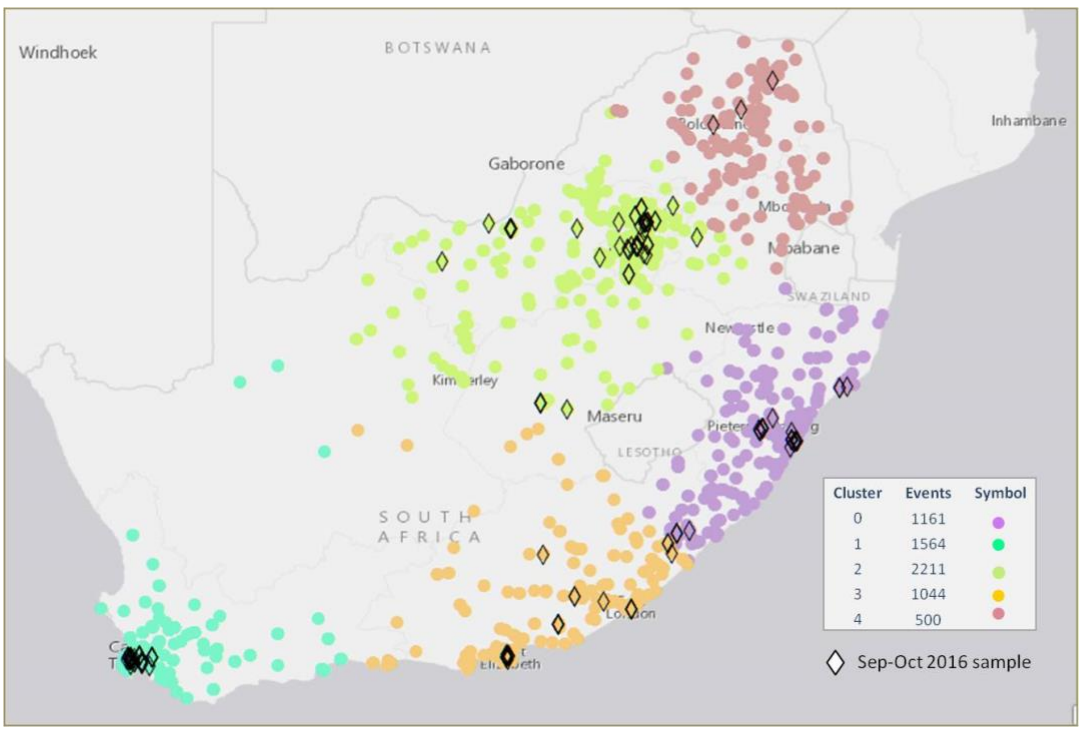

Figure 3. Sample event-based spatial clusters. The events tend to cluster around the population of the nine provinces of South Africa. The largest cluster covers the provinces of Northern Cape, North West, Free State, and Gauteng (2211 events). The other clusters align with the remaining provinces of Limpop and Mpumalanga (500 events), KwaZulu-Natal (1161 events), Eastern Cape (1044 events), and Western Cape (1564 events).

	EVENT COUNT	EVENT TYPE									SOURCE ACTOR								TARGET ACTOR									FATALITY COUNT	CLUSTER LABELS					
Date	E	ET1	ET2	ET3	ET4	ET5	ET6	ET7	ET8	ET9	S1	S2	S3	S4	S5	S6	S7	S8	T0	T1	T2	T3	T4	T5	T6	T7	T8	F	CL0	CL1	CL2	CL3	CL4	Class Label
9/1/2011	3	0	0	0	0	0	0	2	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	1
.																																		
.																																		
.																																		
12/31/2016	6	1	2	0	2	0	1	0	0	0	0	0	4	2	0	0	0	0	0	0	0	0	3	0	0	0	1	0	1	0	0	0	0	0

Figure 4. Input features for the classification tasks. Each feature value is represented as a sum, as multiple events can occur on any given day. The class label indicates whether Tor escalated that day (1) or did not escalate (0).

METHODOLOGY

The goal of supervised machine learning is to make predictions as accurately as possible. A variety of classification methods were selected, implemented, and evaluated to predict the escalation in Tor usage during protests. If models produce accurate scores, then it is reasonable to assume that a relationship exists between events and Tor usage. The approach discovered models that work best for the problem domain and the different datasets. The analytic workflow involved manually exploring, configuring, and evaluating different classification models and then choosing the best performing model.

Twelve different classification methods were evaluated to predict the escalation in Tor usage during political events. This approach resulted in the selection of a classification method that produced the maximum predictive power for each of the nine African country datasets. Each classification method was fit to each of the nine event datasets and iteratively processed against the labeled test data. The algorithms were configured by tuning the hyperparameters or using the recommended default settings. The classifiers evaluated included linear, non-linear, ensembles, and deep leaning estimators. Each classification method uses a specific strategy to determine the decision boundaries given the variances in the datasets. It was assumed that some datasets would be more easily separated linearly, while other datasets would require a non-linear or an ensemble approach to produce an optimal solution.

The twenty-eight data features described in the Data Preparation section captured the unique characteristics of the events for each country. A separate experiment was run that incorporated the cluster labels to determine if this spatial information would improve performance. The input features remained constant across all model iterations. A summary of the methodology is shown in Figure 5.

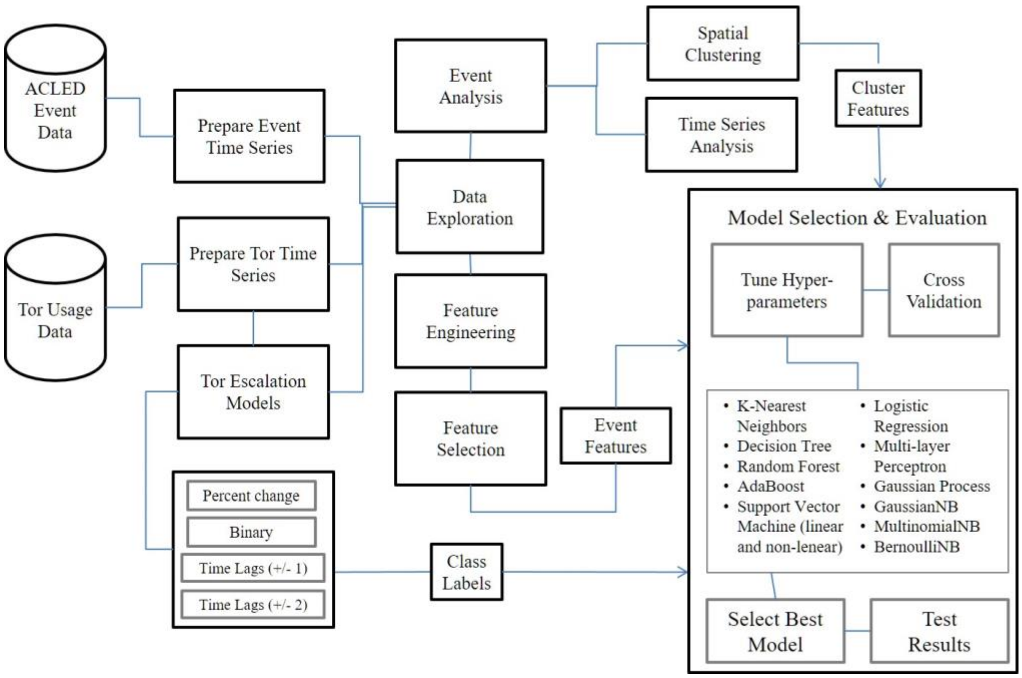

Figure 5. Predictive modeling workflow. The approach allows for a variety of experiments to be conducted easily. New classifiers can be added and different combination of input features can be evaluated.

While certain algorithms are better suited for some problems, particular assumptions and constraints must be considered when selecting an algorithm. For instance, some algorithms can work with categorical features or will automatically convert them to numerical values, while other algorithms fail with this data type. For this reason, all categorical features were one-hot encoded to produce an all numeric dataset. Additionally, some algorithms require missing data imputation to occur prior to learning, while others may be resilient to missing data or automatically impute values. Therefore, the datasets were prepared and held constant across all evaluations to ensure all methods would execute successfully.

Algorithm selection for a problem depends on multiple factors including number of samples and features in the given dataset, the data types and distributions of those features, and the independence or correlations between features and the target variable. The process involved development and exploration of the different classification models using a variety of configurations for each model to optimize its performance. The deeper the understanding of each algorithm can help reduce the amount of time to configure each one properly. Understanding the underlying behavior of the different learning algorithms on each dataset is a challenging undertaking. For each model, parameters were manually tuned via trial and error to achieve optimal performance on each dataset. Table 4 summarizes the models and model hyperparameter settings that were implemented.

Table 4

Summary of Classification Models and Model Hyperparameters.

Model	Model Hyperparameters
K-Nearest Neighbors	n_neighbors=3; weights=uniform; algorithm=brute; p=1
Decision Tree	criterion=entrop; max_depth=3; min_samples_leaf=2
Random Forest	n_estimators=10; criterion=entropy; max_features=3; max_depth=5
AdaBoost	DecisionTreeClassifier (criterion=entropy, max_depth=3, min_samples_leaf=2); n_estimators=10; learning_rate=.01; algorithm=SAMME.R
Support Vector Machine (nonlinear)	C=0.25; kernel=rbf; gamma=0.2
Support Vector Machine (linear)	C=0.1; kernel=linear;
Logistic Regression	penalty=l2; tol=1.0; C=0.01; solver=liblinear; multi_class=ovr
Multi-Layer Perceptron	hidden_layer_sizes=50; activation=logistic; solver=adam; alpha=1.0; max_iter=100; learning_rate=adaptive
Gaussian Process	default parameters
GaussianNB	default parameters
MultinomialNB	default parameters
BernoulliNB	default parameters

For each evaluation of the nine country-level datasets, the data was split into training and testing data using a 90/10 split and run across all models identified in Table 4. Results were compared using prediction test accuracy on held out test data (mean accuracy on the given test data and class labels). Test accuracy was also measured using the F1 score metric, which is commonly used for binary classification problems and considers both precision and recall to compute the score:

$$F1\ Score = 2 * \frac{(Precision * Recall)}{(Precision + Recall)},$$

$$Precision = \frac{True\ Positives}{(True\ Positives + False\ Positives)},$$

$$Recall = \frac{True\ Positives}{(True\ Postives + False\ Negatives)}.$$

To assess how well a model will generalize to unseen data, cross validation was used and the mean cross validation score was reported in the results. A five-fold cross validation technique was used for each evaluation. Finally, a measure to quantify uncertainty was reported to help determine how much confidence a stakeholder should have in the performance of the trained models. Uncertainty of model accuracy was computed using the 95% confidence interval.

RESULTS

The results of the applied machine learning approach are shown in Table 5 with the classifier and target variable shown along with the best overall accuracy scores for each country. The models produced reasonably accurate scores, so it can be assumed that a relationship does exist between conflict events and Tor escalation. Quantitative evidence now exists for the use of Tor escalating prior to a conflict event or responding to an event as demonstrated in the different time lags. Countries that are rated as not free all have the highest accuracy scores and are highlighted in red. This indicates that these countries have the highest signal in the use of Tor in relation to conflict events as compared to those rated as free or partly free. Participants in protests taking place in countries with an authoritarian regime may feel the need to protect their identity and location or need to circumvent censorship during these conflict periods.

Table 5

Predictive Modeling Results

Country	Best Model	Best Target	Accuracy	CV	CI	F1 Score
Egypt*	MultinomialNB	Time Lag 0	0.774	0.624	0.076	0.867
Algeria*	Random Forest	Time Lag 0	0.749	0.695	0.002	0.856
Libya*	Decision Tree	Time Lag -1	0.744	0.691	0.003	0.853
Tunisia	BernoulliNB	Time Lag -2	0.738	0.664	0.019	0.849
Morocco	Gaussian Process	Time Lag +2	0.718	0.655	0.006	0.835
South-Africa	Random Forest	Time Lag -2	0.672	0.626	0.014	0.785
Kenya	Random Forest	Time Lag -2	0.651	0.572	0.005	0.778
Nigeria	Random Forest	Time Lag +1	0.641	0.567	0.044	0.713
Ghana	Random Forest	Time Lag -2	0.621	0.583	0.003	0.761

*Note: * Indicates the most accurate models which are also countries with authoritarian regimes. Random Forest was one of the best performing classifiers and was selected for five of the nine experiments. CV – Cross Validation; CI – 95% Confidence Interval.*

DISCUSSION

It is challenging to quantitatively account for the reason and degree of usage of the Tor Network. The practical uses of anonymity networks are diverse and include identity and location protection, surveillance avoidance, censorship circumvention, and criminal activities. While it is not possible to accurately measure how Tor clients are being used at any particular time, this research successfully detected signals that consistently demonstrated utilization of Tor during periods of civil unrest.

Tor usage patterns for the different countries can vary widely. Therefore, it was necessary to analyze a 5-day window in relation to conflict events so that the different temporal usage patterns would be adequately captured. This approach takes into account whether Tor usage escalates prior to, during, or after onset of a political event. Different Tor escalation models were evaluated and the binary escalation model was used, creating a total of five class labels. The class labels represent a time lag of one or two days prior to or after an event. Experiments were run using each class label against twelve classification methods and the one that produced the best prediction accuracy was selected for that country. If a classification method produces reasonably accurate scores, then it can be assumed that a relationship exists between political events and Tor usage escalation. Different escalation models can be developed that tighten or relax the definition of escalation.

One criterion for selection of countries to include in the study was based on Internet penetration rates. Proxy variable could also be explored to better understand the use of online technologies in target countries, such as quantity of user-generated content on social media platforms, edits on Wikipedia, or the number of active Tor relays or bridges in a country. These proxy variables can help characterize the propensity for users in specific countries to leverage online tools. It should be noted that the absence of Tor relays or bridges in specific countries, such as in Libya, is likely due to the inherent risk to individuals hosting Tor infrastructure and is not a good measure of Tor usage. There is no requirement to have a Tor relay in the host country for users of that country to access the Tor network. Finally, analysis of content on specific Tor Hidden Services, can estimate usage of Tor for specific purposes. Hidden forums, blogs, chat rooms, and social networks that discuss civil unrest and protest events in target countries would provide an estimate on the number of accounts using Tor for specific purposes. These proxy variables and related analyses offer potentially interesting areas to extend this research.

Each country in the study has its own unique geography, demographic, government, and other factors that can shape political events or public discontent. Qualities of governance, degree of freedom, and access to Internet technologies have a direct impact on the patterns of collective action as well as the resulting behaviors in both physical and virtual spaces. These factors are intrinsically represented in the data and therefore create unique problem sets for each country. While spatial information can be used to explore the different temporal and spatial patterns of conflict events, the addition of spatial information provided using the spatial clustering labels as new features did not produce better performing models. Spatial analysis focusing on specific regions in countries and specific event types or actor categories also offer potentially interesting areas to extend this research.

The methodology used in this study aimed to be consistent across all the datasets and to avoid overfitting through extensive feature selection or using specific slices of the input data. All evaluations used the same exact set of input features and used the full dataset for the entire study period for each country prior to fitting the models. Model improvements may be accomplished by removing anomalous periods in the Tor usage data. Additional approaches to consider include using multi-class classification rather than binary classification or conduct regression analysis on continuous changes in Tor usage versus the binary escalation model. Additional countries from the ACLED dataset could be added to the analysis including Sudan, Rwanda, Zimbabwe, and Uganda. Each of these countries have Internet penetration rates above 25%. This study ensured an even distribution of countries across the Freedom House index of free, partly-free, and not-free. These four countries are not-free, except Uganda, which is partly-free. Expanding the study beyond Africa would also be of interest, assuming similar quality crisis event datasets are available. ACLED has recently expanded its data collection activities in the Middle East and Asia.

CONCLUSION

Generally, people use anonymity-granting tools because they are concerned about their privacy and want to avoid governments and companies infringing on their civil rights. This research set out to measure the escalation in Tor usage during periods of civil unrest. To better understand this relationship, a five-day window on Tor usage was captured in the modeling process. These time lags accounted for the different temporal usage patterns that may exist prior to or during protests. The analysis incorporated nine African countries covering a diverse set of geographic regions, event types, and actor categories. The results demonstrated that a reasonably strong relationship exists between social movements and Tor usage. This was the first study to produce quantitative analytic results that demonstrate that anonymizing tools are essential for protecting human rights and used for important societal purposes.

REFERENCES

Agarwal, S., & Sureka, A. (2015). Applying social media intelligence for predicting and identifying on-line radicalization and civil unrest oriented threats. *arXiv preprint:1511.06858 [Cs]*. Retrieved from http://arxiv.org/abs/1511.06858

Arthur, D., & Vassilvitskii, S. (2007). K-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (pp. 1027–1035). Philadelphia, PA, USA: Society for Industrial and Applied Mathematics. Retrieved from http://dl.acm.org/citation.cfm?id=1283383.1283494

Arva, B., Beieler, J., Fisher, B., Lara, G., Schrodt, P. A., Song, W., … Stehle, S. (2013, July). Improving forecasts of international events of interest). *EPSA 2013 Annual General Conference Paper 78.* Retrieved from https://papers.ssrn.com/abstract=2225130

Bastos, M. T., Recuero, R. da C., & Zago, G. da S. (2014). Taking tweets to the streets: A spatial analysis of the Vinegar Protests in Brazil. *First Monday*, *19*(3). doi: 10.5210/fm.v19i3.5227

Bennett, W. (2003). Communicating global activism. *Information, Communication & Society*, *6*, 143–168. doi: 10.1080/1369118032000093860a

Bennett, W. L., Segerberg, A., & Walker, S. (2014). Organization in the crowd: Peer production in large-scale networked protests. *Information, Communication & Society*, *17*, 232–260. doi: 10.1080/1369118X.2013.870379

Braha, D. (2012). Global civil unrest: Contagion, self-Organization, and prediction. *PLOS ONE*, *7*(10), e48596. https://doi.org/10.1371/journal.pone.0048596

Buhaug, H., & Gates, S. (2002). The Geography of civil war. *Journal of Peace Research*, *39*, 417–433. doi: 10.1177/0022343302039004003

Conover, M. D., Davis, C., Ferrara, E., McKelvey, K., Menczer, F., & Flammini, A. (2013). The Geospatial Characteristics of a Social Movement Communication Network. *PLOS ONE*, *8*(3), e55957. https://doi.org/10.1371/journal.pone.0055957

Freedom House (2018) *Freedom in the World.* Washington, D.C.: Author

Garrett, R. K. (2006). Protest in an Information Society: A review of literature on social movements and new ICTs. *Information, Communication & Society*, *9*, 202–224. doi:10.1080/13691180600630773

Gleditsch, K. S., Metternich, N. W., & Ruggeri, A. (2014). Data and progress in peace and conflict research. *Journal of Peace Research*, *51*, 301–314. doi: 10.1177/0022343313496803

Hegre, H., & Sambanis, N. (2006). Sensitivity Analysis of Empirical Results on Civil War Onset. *Journal of Conflict Resolution*, *50*, 508–535. doi: 10.1177/0022002706289303

Janelle, D. G. (1969). Spatial Reorganization: A Model and Concept. *Annals of the Association of American Geographers*, *59*, 348–364.

Jardine, E. (2016). Tor, what is it good for? Political repression and the use of online anonymity-granting technologies. *New Media & Society*, *20*, 435-452. doi: 10.1177/1461444816639976

Klausen, J. (2015). Tweeting the Jihad: Social media networks of western foreign fighters in Syria and Iraq. *Studies in Conflict & Terrorism*, *38*, 1–22. doi: 10.1080/1057610X.2014.974948

Lee, F. L. F., & Chan, J. M. (2016). Digital media activities and mode of participation in a protest campaign: A study of the Umbrella Movement. *Information, Communication &*

Society, 19, 4–22. doi: 10.1080/1369118X.2015.1093530

Leukfeldt, E. R., Kleemans, E. R., & Stol, W. P. (2017). Origin, growth and criminal capabilities of cybercriminal networks. An international empirical analysis. *Crime, Law and Social Change, 67*, 39–53. doi: 10.1007/s10611-016-9663-1

Massa, F. G. (2016). Guardians of the internet: Building and sustaining the anonymous online community. *Organization Studies, 38*, 959-988 0170840616670436. doi: 10.1177/0170840616670436

Muthiah, S. (2014, July 11). *Forecasting protests by detecting future time mentions in news and social media* (Mastser's thesis). Virginia Tech University, Blacksburg, VA. Retrieved from https://vtechworks.lib.vt.edu/handle/10919/49535

Myers, D. J. (1994). Communication technology and social movements: Contributions of computer networks to activism. *Social Science Computer Review, 12*, 250–260. doi: 10.1177/089443939401200209

Nicholls Walter. (2008). Place, networks, space: theorising the geographies of social movements. *Transactions of the Institute of British Geographers*, *34*, 78–93. doi: 10.1111/j.1475-5661.2009.00331.x

Rady, M. (2013). Anonymity networks: New platforms for conflict and contention *MIT Political Science Department Research Paper No. 2013-5*. Retrieved from SSRN: https://ssrn.com/abstract=2241536

Raleigh, C., Linke, A., Hegre, H., & Karlsen, J. (2010). Introducing ACLED: An armed conflict location and event dataset: Special Data Feature. *Journal of Peace Research, 47*, 651–660. doi: 10.1177/0022343310378914

Ramakrishnan, N., Butler, P., Muthiah, S., Self, N., Khandpur, R., Saraf, P., … Mares, D. (2014). "Beating the news" with EMBERS: Forecasting civil unrest using open source indicators. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1799–1808). New York, NY, USA: ACM. doi: 10.1145/2623330.2623373

Rika Butler. (2007). A framework of anti-phishing measures aimed at protecting the online consumer's identity. *The Electronic Library*, *25*, 517–533. doi: 10.1108/02640470710829514

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. doi: 10.1016/0377-0427(87)90125-7

Sandberg, B. (2017). Geography of Anonymous Communications: Detecting Online Censorship. *Journal of Mason Graduate Research*, *5*(1), 17–36. doi: 10.13021/G8jmgr.v5i1.1782

Schutte, S., & Donnay, K. (2014). Matched wake analysis: Finding causal relationships in spatiotemporal event data. *Political Geography*, *41*, 1–10. doi: 10.1016/j.polgeo.2014.03.001

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, *46*, 234–240. doi: 10.2307/143141

Yuan, Y. (2017). Exploring the Spatial Decay Effect in Mass Media and Location-Based Social Media: A Case Study of China. In *Advances in Geocomputation* (pp. 133–142). Springer, Cham. doi: 10.1007/978-3-319-22786-3_13