Measuring Global Instability Through Big Data

David Mace
Zack Polizzi
California Institute of Technology

Motivation

We believe that if you wish to improve something, you must first understand it. The world needs improving, yet it is difficult to even get a picture of what is going on in the world at any point in time. Traditional sources of global information, such as news sources, are biased towards the views and interests of the publisher, and lead to reporting only on events that improve readership and profits. Therefore, it is difficult or even impossible for an individual, an international aid organization, or even a national government to get an accurate, quantifiable, and real-time picture of world happenings. Without such information, it is difficult to plan operations, route aid efforts and diplomacy, or perform any other operation requiring an accurate idea of the world.

Goals

Political instability is a metric that is very important to quantify and understand, yet traditionally extremely difficult to accurately measure. Therefore, the goal of this project is to create an algorithm that ingests various sources of data and compiles them into a metric of political instability for each country. The algorithm should:

- Take in a variety of big data sources
- Operate on a time scale small enough to be useful day-to-day
- Operate without expert knowledge or manual classification

It will:

- Create sub-metrics for various more-easily measured quantities
- Combine sub-metrics into a holistic system to quantify instability
- Be designed with a flexible, plug-in based system
- Be easily modified, improved, and updated with new metrics

Input Data

News:
About 1.7 million news articles were collected from the Guardian open content API. These articles cover all the news types published by The Guardian, and are tagged with metadata covering topic, date, geographical region of interest, and various other tags.

Foreign Exchange:
Currency exchange rates (relative to USD) for 30 different countries for the past 5-15 years (dependent on country). Data in 1-minute resolution.

Twitter:
1% of all tweets posted on Twitter for a 3-month period. Due to the short time-scale of this sample (and the difficulty/expense of acquiring a larger sample), this data was used for training, testing, and comparison but not direct prediction.

Algorithm Implementation

Designing an algorithm to predict political instability from raw, loosely-structured data has no clear, intuitive solution. Instead of attempting to train a huge neural network or some other opaque, unintuitive system to directly compute instability from the input data, we chose to view the problem at a higher level. While directly computing political instability from our input data is difficult, there are plenty of metrics or classifiers that we can intuitively compute from the inputs. For example, in this paper we explore detecting terrorist attacks from news headlines, measuring economic instability from foreign exchange volatility, and measuring general sentiment through news articles and social media. Creating each of these metrics from the input data is relatively simple and intuitive, and combining each of these sub-metrics into the larger metric of political instability is simple and logical step. By breaking the problem into various sub-problems, we are able to compute a “hard” problem by combining multiple solutions to “easier” problems.

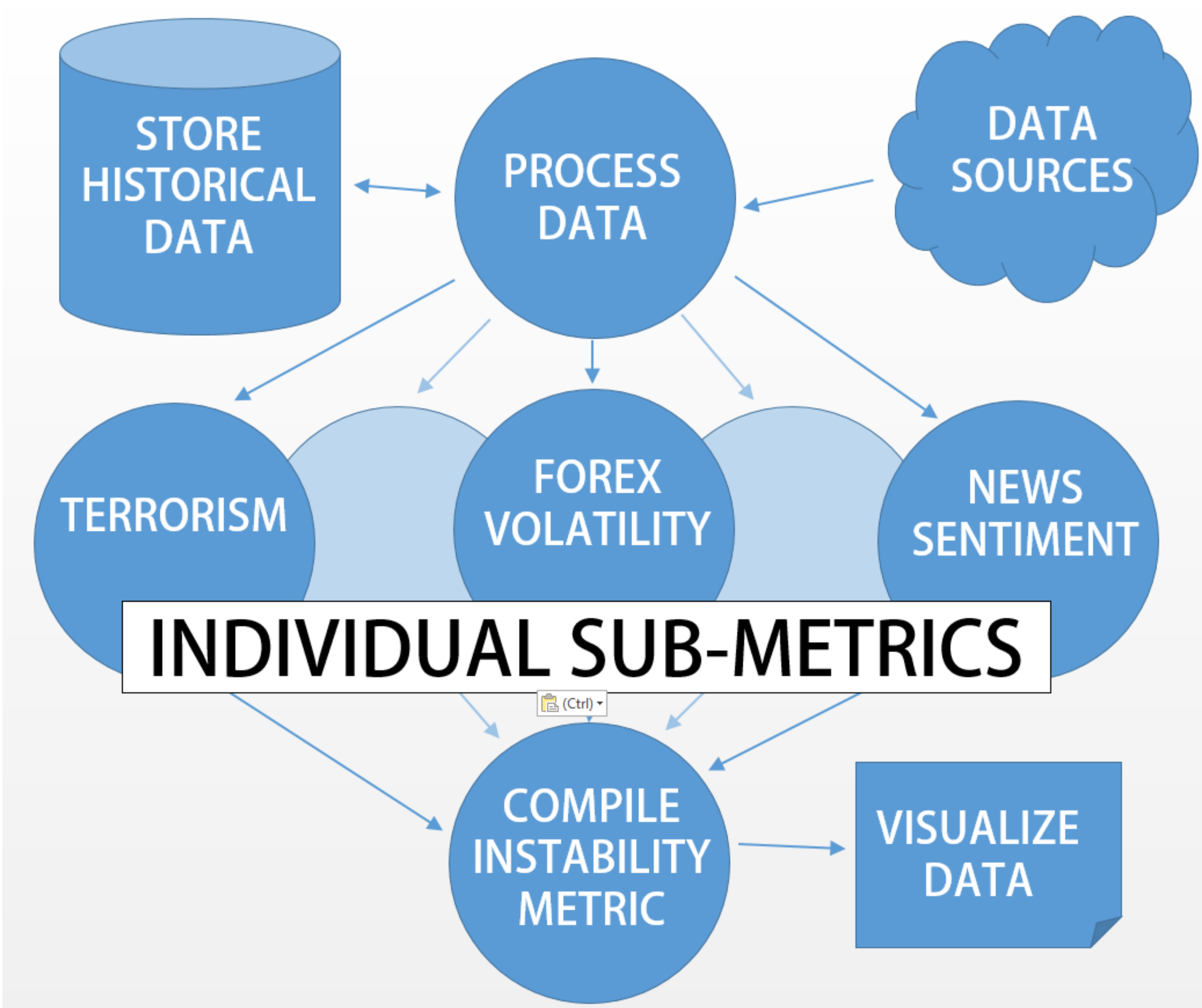


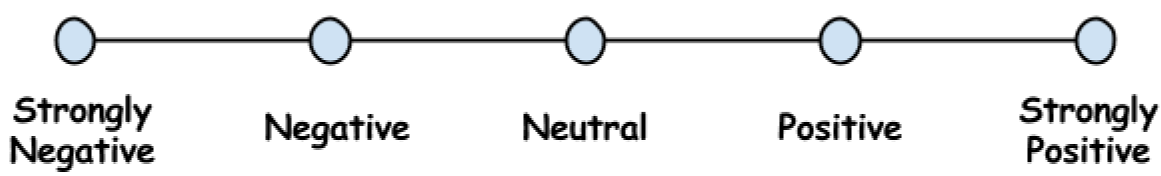
Fig. 2: Algorithm Flowchart

The political-instability prediction system outlined in this paper is extremely flexible and extensible, and is capable of adapting to any set of input data and sub-metric analysis available. Therefore, the following implementation is simply an example of a possible set of inputs and sub-metrics, and is certainly expected to become more robust with additional data and analysis. To build an initial demonstrator of our algorithm, we implemented 3 core sub-metrics.

News Sentiment

We developed a bag-of-words model to train a Support Vector Machine to classify news article titles based on their political significance. A training set of about 10,000 titles was manually classified using Amazon Mechanical Turk workers. The workers were asked to rate a set of article titles based on the following question:

Please rate the following news headlines, based on how good or bad they are in relation to political unrest/instability. A headline about war or terrorism should be rated "Strongly Negative" or "Negative", while an article about peace talks should be rated "Positive". NO headline should be rated "Strongly Positive". Headlines unrelated to political unrest should be rated "Neutral", such as headlines about pop culture or sports.



Each headline was rated by 3-5 workers to ensure a reliable rating, and the ratings of each headline were used as the training data for the SVM on the bag-of-words model. Accuracy of the trained model was not exceptional (around 70% correct classification), partially because the Mechanical Turk-classified headlines were somewhat unreliably classified (even with averaging), and partially because the bag-of-words model simply does not capture enough information to accurately classify some headlines. For example, some headlines were difficult to classify even manually, requiring reference to the full article to check the true meaning. Therefore, it is understandable that a simple word-frequency model would not be exceptionally accurate at classification. However, for our purposes, excellent accuracy isn't extremely important, as incorrectly classified articles will be averaged out in the larger model.

Terrorism and Assassinations

A terrorism and assassinations detector was developed by hand-coding language rules to detect these events in news headlines. These language rules were created by manually classifying a number of headlines containing likely keywords, and identifying common language structures among those with the same keyword. These language structures were then hard-coded into a classifier for each keyword. Since headlines that indicate a terrorist attack or assassination and contain a certain keyword have limited structure and variability, this method was very effective (~90% accuracy with minimal tweaking). The addition of another news source would greatly increase this accuracy, as it would allow cross-comparing across the sources to eliminate many false-positives which would likely only be triggered on one news source.

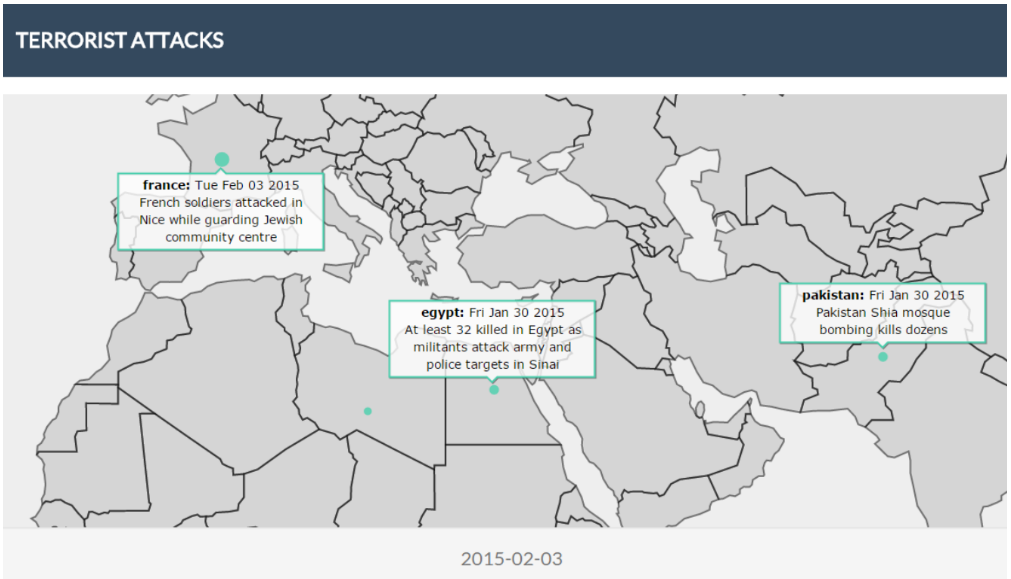


Fig. 1: Sample Output of Terrorism Detector, from February 2015. Terrorism-related events detected in France, Egypt, and Pakistan through language-model detector.

Economic Volatility

A measure of the economic instability of a country was developed using the historical foreign exchange data. A simple metric of exchange rate volatility served as an effective indicator of economic instability - in an unstable economy, the opinion of global investors is often changing, resulting in a large amount of volatility in the exchange rate. The sliding-window size of the volatility calculator can be adjusted to the time-scale of interest - to predict large-scale movements of instability, a weekly volatility window might be most effective, while to understand intraday movements, an hour-long window would make much more sense.

Figure 5: US Dollar – EURO exchange rate and volatility chart. High volatility of currency rates is a good indicator of instability within an economy.

Instability Metric

Now that the input data has been processed into a set of sub-metrics, a method of combining the sub-metrics into the overall political instability metric is necessary. Since there has not been significant past research into how "political instability" should be defined, this model was created based on the intuition of the authors into how political instability manifests itself through the the data sources are able to analyze. Later, we will show that our intuition yields a metric that correlates well with expected instability levels in a historical analysis.

The model that we created is based on the intuition that our sub-metrics are able to most effectively measure and detect events that would cause or correlate with an increase in the instability of a country. Therefore, in the absence of such signals, a country is assumed to be relatively stable. Based on this intuition, instability was modelled with an exponential-decay model - events detected by the sub-metrics, such as terrorism, economic instability, or negative sentiment contribute a small, fixed positive increase to the instability of a country, and the instability levels of all countries otherwise slowly decay at a rate proportional to their instability levels. Therefore, high levels of instability, usually caused by short, temporal events in a country, rapidly decay to a lower level, but it takes much longer for the country to decay the same amount from a moderate instability level.

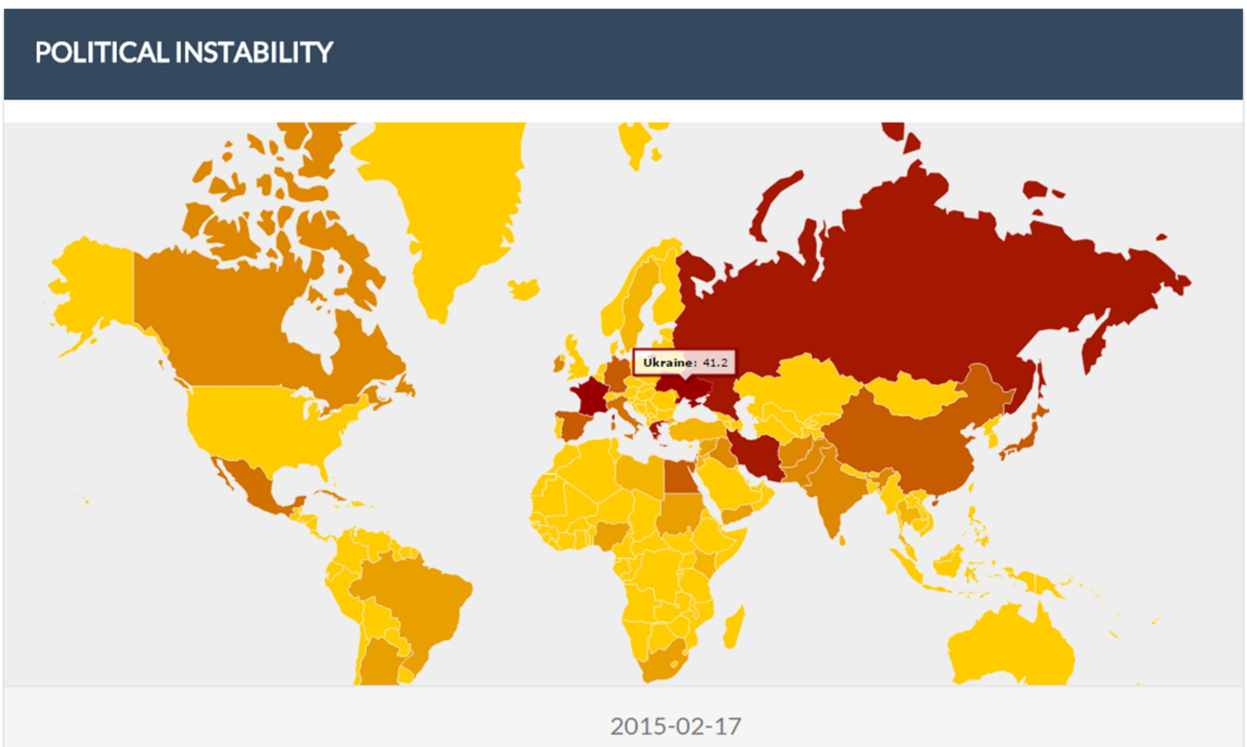


Fig. 4: Output Metric of Political Instability on February 17, 2015. Color value scale included in bottom left, with darker colors indicating greater instability. Ukraine highlighted as an example, with high instability values due to the Ukraine Crisis.

Analysis of Instability Metric

Analysis of the quality of the metric is difficult due to the lack of similar metrics to compare to. There exist some very coarse attempts at measuring instability, but these are usually updated once a year or so. Our metric is aimed at having resolution of a day or so, so comparing these metrics would be difficult at best. Instead, we will look at the large-scale historical results of our metric on a well-known country.

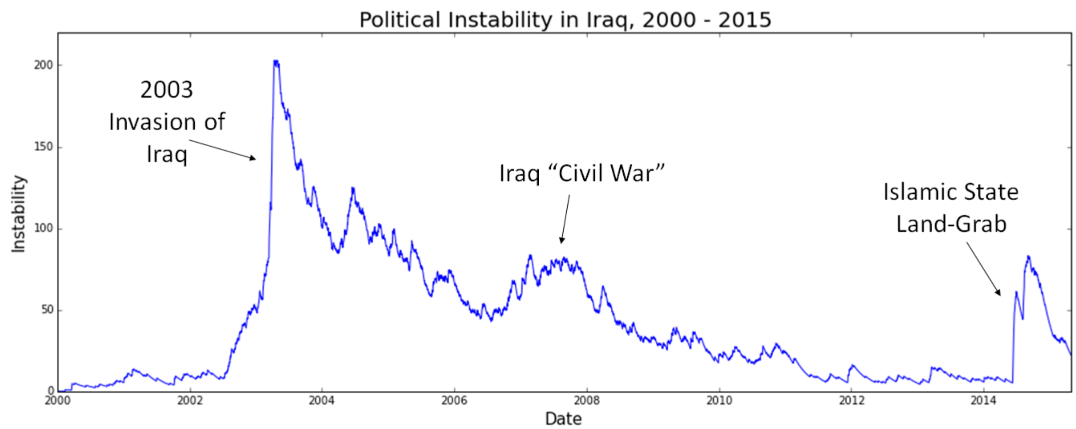


Fig. 3: Political Instability metric for Iraq from 2000 to 2015. Significant events in Iraq's history can be clearly observed in this chart, such as the spikes in instability during the 2003 Invasion of Iraq, the "Civil War" around 2006, and the recent land-grab by the Islamic State.

As we can see from the calculated historical instability of Iraq, our metric correlates accurately with significant events in Iraq's recent history, and the levels of instability we would roughly expect from such events. A similar analysis of a number of other countries shows that our metric is relatively accurately correlated with estimated historical instability levels.

Prediction

Of course, while being able to accurately measure global political instability is extremely important, being able to predict this metric into the future with some statistically significant level of accuracy would be even more groundbreaking. We believe the sub-metrics and overall metric, when combined with other coarse data (ex. CIA World Fact Book), contain enough information that relationships can be determined among instability movements in each country – allowing predictions to the future movement of the instability metric to be made. Initial testing with simple clustering and cross-correlation has been promising.

Future Work

The work done so far is only a small sample of the possible data and sub-metrics that could be computed. We are currently working on numerous additional sub-metrics, covering employment, education, criminal violence, and level of social activism, among many others. Each additional sub-metric will improve the overall metric accuracy, while lowering the bias towards any one source. Further work on predicting the metric should also prove exciting.