A Trend Pattern Approach to Forecasting Socio-Political Violence

Kurt Rohloff, Rob Battle, Jim Chatigny, Rick Schantz
BBN Technologies
Cambridge, MA 02138
{krohloff, rbattle, jchatigny, schantz}@bbn.com

Victor Asal
University at Albany, SUNY
135 Western Ave.
Albany NY 12203
vasal@email.albany.edu

Abstract

We present an approach to identifying concurrent patterns of behavior in in-sample temporal factor training data that precede Events of Interest (EoIs). We also present how to use discovered patterns to forecast EoIs in out-of-sample test data. The forecasting methodology is based on matching entities' observed behaviors to patterns discovered in retrospective data. This pattern concept is a generalization of previous pattern definitions. The new pattern concept, based around patterns observed in trends of factor data is based on a finite-state model where observed, sustained trends in a factor map to pattern states. Discovered patterns can be used as a diagnostic tool to better understand the dynamic conditions leading up to specific Event of Interest occurrences and hint at underlying causal structures leading to onsets and terminations of socio-political violence. We present a computationally efficient data-mining method to discover trend patterns. We give an example of using our pattern forecasting methodology to correctly forecast the advent and cessation of ethnic-religious violence in nation states with a low false-alarm rate.

Introduction

A major challenge in analyzing complex systems is identifying patterns of behavior which are symptomatic precursors to Events of Interest (EoIs) in entities. In this paper we focus on problems of pattern discovery and forecasting for onsets or terminations of socio-political violence in nation-states. By identifying patterns of behavior that precede onsets or terminations of socio-political violence in nation-states we can begin to both understand the underlying causal structures which drive socio-political violence and forecast these onsets and terminations of socio-political violence.

Due to our motivating context we focus on discovering patterns for EoIs which are considered rare events (rebellions, insurgencies, civil wars, etc...) (King and Zeng 2001; 2003). An event is considered a *rare event* if there are dozens to thousands of times fewer occurrences (such as wars, coups, etc...) than non-occurrences of these events (King and Zeng 2001; 2003). Up to now, the primary literature on analyzing the pre-conditions for rare events has relied on statistical approaches. Unfortunately, statistical approaches can underestimate the probability of rare events (King and Zeng 2001; 2003). The changes in probable likelihood using even sophisticated statistical approaches are very small and do not allow for a smart investment of resources when determining policy based on such small probability bases.

In this paper we present a generalizable sequential pattern concept based on the supposition that the phenomena which cause (or at least are related to) the occurrences of Events of Interest (EoIs) exhibit similar symptomatic behaviors across multiple EoI occurrences. For example, countries experiencing rebellions driven by the desire for freedom by internal ethnic groups commonly exhibit increasing ethnic tension and violence before the occurrence of ethnic rebellions. We formalize our sequential pattern concept using a finite-state machine model of countries' behaviors and use collections of sampled factor data to define the "states" of entities such as countries.

The sampled factor data represents quantifiable measurements of countries at discrete, regular points in time. We use a discrete clock-tick formalism to model the updating of state locations. Example factors from our socio-political domain include GDP, the rates of occurrence of various words in the national press, the average caloric intake, Goldstein measures of conflict/cooperation between governmental entities, etc... These example factors change continuously over time which motivates our use of sampled data. We map the sampled factor data to observed "trends" in this factor data where a factor's sampled measurement can be increasing, decreasing or fluctuating over either the short-term or the long-term.

Our pattern concept is generalizable and we anticipate that it can be applied in other complex system application contexts where patterns of behavior in sampled factor data preceding events needs to be identified and represented. In our motivating context of socio-political violence we are interested in patterns that match the trends of observed behaviors preceding at least two instances of EoI occurrences and which are not present in countries when an EoI does not occur over historical data. We present a generalizable, computationally efficient branch-and-bound back-chaining method to identify the set of factors which define a state space in patterns that match the behavior preceding EoI occurrences in at least two countries from historical data. The backwards chaining methodology permits us to identify which factors change similarly over for multiple countries for several time steps leading up to the socio-political violence onset or termination in selected countries in a computationally efficient manner.

Our understanding of patterns of behavior preceding onsets and terminations of socio-political violence is based on the concept of equifinality, "many alternative causal paths to the same outcome (George and Bennett 2005)." A country or other complex system may or may not follow multiple patterns leading to EoI occurrence simultaneously, and there may be multiple means to a same end. For example, a country such as India may contain multiple types of rebellion occurring in different parts of the country and India may exhibit the antecedents for several rebellions simultaneously, each driven by different conditions. As may be expected from our concept of equifinality, not all occurrences of an EoI can be described by a single pattern. Therefore, the major challenge is to find a sufficient number of patterns that describe enough of the possible dynamics preceding the onset or termination.

As a result of our hypothesis that the phenomena which cause (or at least are related to) the onsets and terminations of socio-political violence exhibit similar symptomatic behaviors across multiple onsets and terminations of socio-political violence, we can generate real-time early-warning forecasts of EoIs if early portions of the patterns are observed in a specific country. This forecasting process is based around the notion of *matching* a country's behavior to early parts of historical patterns that end in socio-political violence. If the country's behavior matches the early parts of the pattern then we forecast that onsets or terminations of socio-political violence will occur in the country in the near future. We found that this approach to forecasting using single patterns is inadequate in practice because individual patterns provide a limited representation of the full breadth of all possible behaviors that may precede onsets and terminations of socio-political violence. This motivates our need to generate libraries of patterns that provide a broader representation of the observed preceding dynamics associated with the occurrence of onsets and terminations of socio-political violence.

We search in experimental data for patterns that lead to the onset or termination of socio-political violence driven by (or at least related to and preceding) government policy, expressed conflict or cooperation between political entities and immediate antecedent behavior. These antecedent conditions can create agitation and spark violence amongst a country's population when expectations are let down or there is a spike in repression. Importantly we notice that contextual information is non-trivial: some factors in a pattern state space might not change over time, but they set an important context for the country's state evolution.

We demonstrate our pattern discovery and forecasting methodologies over data of onsets and terminations of ethnic-religious violence in Pacific-region countries from 1998-2006. We show that by discovering patterns for ethnic-religious violence onset and ethnic-religious violence termination over Pacific-region countries from 1998-2004, we can use these patterns to forecast ethnic-religious violence onset and ethnic-religious violence termination over 2005-2006 with a very low false-alarm rate.

The remainder of the paper is organized as follows. In the next section we provide an overview of related work. We then define the notion of trends for our trend pattern concept and then formally defines the trend pattern concept. We subsequently discuss the discovery of patterns in observed factor data and then present our trend pattern forecasting methodology. We next discuss how to tune parameter to improve forecasting performance and then present an example use of our pattern discovery and forecasting methodology to predict advents and cessations of ethnic-religious violence with a low false alarm rate. The paper then closes with a discussion of our results.

Related Work

There has been a lot of work (both recent and historical) on the problem of finding patterns in time-series data. Several recent notable efforts on this topic includes (Das and Schneider 2007; Dzeroski and Struyf 2007; Huang and Chang 2008; Kiernan and Terzi 2008; 2009). Most of this previous work was either generic in nature, or focused on pattern discovery in data generated by engineered systems such as computer networks.

There has been little previous work on the discovery of logical patterns preceding the outbreak of socio-political violence. Notable previous work on pattern discovery in the socio-political violence context includes (Ragin 2000; Chan 2003). This previous work in the social science focused on non-temporal data and examined static conditions assoicated with violence onset or termination. We build on efforts to use Boolean analysis (Ragin 2000) to understand political activity "while allowing for" multiple causal mechanisms (Chan 2003). Most of this work though starts with collecting data on what is trying to be explained rather than collecting data on both the general environment (the dogs that don't bark - and the cats that never would) as well as those occasions where there is an outbreak of an EOI.

We previously published an alternative sequential pattern definition with associated discovery methodologies were presented in (Rohloff and Asal 2008; 2009). This earlier work defined patterns over quantized factor data rather than dynamically defined trends in changes in factor data as is presented in this paper. Also, this paper presents a working methodology with a demonstration example for the forecasting of ethnic-religious violence onset and termination which the previous pattern approaches were not capable of.

Trends Observed in Factor Data

We start our investigation into behaviors preceding socio-political violence with the belief that the outbreak of violence will be characterized by equifinality, "many alternative causal paths to the same outcome (George and Bennett 2005)." In other words there is not one cause of socio-political onset or termination. There may be a set of factors that make those outbreaks more or less likely but we believe there are a set of potential causal pathways that are likely to lead to outbreak of non-state actor violence towards the state. We are interested in exploring the combinational power of various factors as they lead to the onset or termination of socio-political violence. To do this we build on efforts to use Boolean analysis (Ragin 2000) to understand political activity "while allowing for" multiple causal mechanisms (Chan 2003). Most of this work though starts with collecting data on what is trying to be explained rather than collecting data on both the general environment (the dogs that don't bark - and the cats that never would) as well as those occasions where there is an onset or termination of socio-political violence.

We define our patterns to be sequences of trends of behaviors observed in factors before the onset or termination of socio-political violence. We look for trends in sampled factor data where the sampled factors either increase, decrease or fluctuate over either the short-term or the long-term. For instance, in India in the quarters preceding the onset of ethnic-religious violence in early 2002, we see that the level of cooperating expressed by the government towards opposition parties (as measured by a Goldstein metric) holds fluctuating for several quarters before increasing over a short term and then decreasing shortly before the onset of violence.

With this in mind, we formally define the 6 possible types of trends that can be observed in factors as:

- Long-Term Increasing
- Short-Term Increasing
- Long-Term Fluctuating
- Short-Term Fluctuating
- Long-Term Decreasing
- Short-Term Decreasing

For our application context of socio-political violence we define *short-term* trends as those occurring over 3 time samples or less, and *long-term* trends occur over 3 time samples or more.

We allow the definitions of *increasing*, *decreasing* and *fluctuating* to be country- and factor-specific. Our motivation for this is that "normal" observed factor behaviors change differently not only from factor to factor (as may be intuitive because different factors measure different phenomena), but that "normal" observed factor behavior varies from from country to country for the same factor. As an example, any small change in the level of cooperation expressed by the Chinese government towards potential opposition parties is unusual and significant, but relatively dramatic observed changes in the level of cooperation expressed by the Indian government towards opposition parties is fairly routine.

To map observed changes in factor data to increasing, decreasing or fluctuating trends, we use a weighted threshold test based on the standard deviation of the changes in the factor over a set of training data. For a given set of training data D_{train}, a country c and a factor f, we denote the standard deviation of the changes in the factor f over the country c in the training data D_{train} as $\sigma(D_{train}, c, \delta f)$, or σ if context allows. If the factor f for a country c is observed at times t_i and t_j ($i < j$), we define the following for our application context:

- f is increasing from t_i to t_j if $w\sigma < f(t_j) - f(t_i)$.
- f is fluctuating from t_i to t_j if $-w\sigma \leq f(t_j) - f(t_i) \leq w\sigma$.
- f is decreasing from t_i to t_j if $f(t_j) - f(t_i) < -w\sigma$.

The variable w is a user-defined threshold weighting factor that can be used to tune the trend definitions. A large w permits factor data to be associated with increasing or decreasing trends only if there are relatively drastic changes in a factors' values. Similarly, a small w permits factor data to be associated with increasing or decreasing trends when there are minor changes in a factors' values. We found that the variable w is an important tuning parameter that can be adjusted based on the type of violence that a pattern is defined for.

Trend Patterns

We build on our notion of observed trends in factor data to define patterns based on sequences of observed trends over multiple factors. Formally, we define a pattern P as

$$P = (F, \tau) \tag{1}$$

where F is a set of factors and $\tau : F \rightarrow T^*$ maps factors to sequences of trends. If we use T to represent the set of all possible trends (as define in Section), we define $\tau(f)$ as follows:

$$\tau(f) = \begin{cases} t^f_{-n} \cdots t^f_{-1} & \text{if } f \in F \\ \emptyset & \text{otherwise} \end{cases} .$$

The sequence $t^f_{-n} \cdots t^f_{-1} \in T^*$ represents the sequence of trends that precede the onset/termination of socio-political violence described by the pattern where t^f_{-i} is the trend exhibited by factor f i steps before the onset/termination. (Note that T^* represents the Kleene-closure of T: the set of all sequences constructed from elements of T (Cassandras and Lafortune 1999).)

Note that our pattern definition is independent of specific countries and specific thresholds for trends to be increasing, decreasing or fluctuating. We made this design decision so that patterns can be adaptively applying depending on the local context of what observable factor behavior is "normal" for each country.

An example single-factor pattern that we found in India in the quarters preceding the onset of ethnic-religious violence in early 2002, is for a factor representing the Goldstein measure of the cooperation expressed by the government towards opposition party is the sequence of trends $t^f_{-3}t^f_{-2}t^f_{-1}$ = Short-term Fluctuating, Short-term Decreasing, Short-Term Increasing.

Trend Pattern Discovery

We implemented Algorithm 1 to discover single-factor patterns from a set of factors F, a set of onset/termination events E and a trend threshold weight w for the set of trends T as defined above. The algorithm returns a set of single-factor patterns that we later merge to form patterns with multiple factors.

Data:
A set of factors $F = f_1, \ldots, f_m$.
A set of onset/termination events $E = e_1, \ldots, e_l$.
A scalar weight w.
Result:
A set of single-factor patterns P that correspond to behavior observed in at least two events in E.
T is the set of possible trends;
$Return\_List := \emptyset$;
$Test\_List := \emptyset$;
for $f \in F$ **do**
  for $t \in T$ **do**
    $test := newSequence(t)$;
    $matches := 0$;
    for $e \in E$ **do**
      $data$ = data preceding e;
      if *validSequence(test,f(e),σ,w)* **then**
        $matches++$;
      end
    end
    if $matches > 2$ **then**
      $Test\_List.add(test)$;
    end
  end
  while $Test\_List \neq \emptyset$ **do**
    $current := Test\_List.pop()$;
    for $t \in T$ **do**
      $test :=$ $current.addTrendToSequence(t)$;
      $matches := 0$;
      for $e \in E$ **do**
        $data$ = data preceding e;
        if *validSequence(test,f(e),σ,w)* **then**
          $matches++$;
        end
      end
      if $matches > 2$ **then**
        $Test\_List.add(test)$;
        $Return\_List.add(newPattern(f, test))$;
      end
    end
  end
end
return $Return\_List$;
Algorithm 1: The Pattern Discovery Algorithm

Algorithm 1 is an iterative branch-and-bound algorithm that iteratively expands $Test\_List$, a running list of possible sequences that would form a valid pattern with respect to a set of factor data, a threshold weight and a set of onset/termination events. Algorithm 1 has an outer loop that tests all possible factors $f \in F$ for possible sequences that could form a single-factor pattern for the given data. In the first half of outer for-loop in Algorithm 1, initial sequences for $Test\_List$ are found from the set of all sequences formed by single Trends. The second half of the outer for-loop in Algorithm 1 iteratively attempts to add new trends to each sequence in the $Test\_List$. If this new sequence would form a valid single-factor pattern with respect to the factor f for at least two sequences, then this new sequence is re-added to the test list, and a new pattern is formed from this sequence and factor and added to the return list.

The set of all possible sequences that could be returned by Algorithm 1 can be thought of as a tree. Algorithm 1 searches over the tree using a branch-and-bound approach to find the sequences which define valid single-factor patterns. Although the size of the search tree is exponential in the maximum search depth, we found in practice that this algorithm scales well and runs very efficiently over rea-world data. In Algorithm 1, the algorithm continually and recursively attempts to find larger prefixes of the test data such that the prefix would match the trailing trend, and the suffix of the data matches the remaining sequences of trends. This recursive algorithm returns true in the default case if there are no more sequences to attempt to match. This default case will occur after several recursions if successive prefixes of the original data can be mapped to various trends with respect to the given w and σ.

Data:
A sequence of trends $t^f_{-n} \cdots t^f_{-1}$.
A sequence of factor data $data$.
A standard deviation measure σ.
A scalar weight w.
Result:
$true$ or $false$ to represent whether the trends match the observed factor measurements preceding at least two events.
if $\|t^f_{-n} \cdots t^f_{-1}\| == 0$ **then**
  return true;
end
$i := n$;
$trend := t^f_{-i}$;
for $sub\_data \in prefixes(data)$ **do**
  if *$trend$ fits sub_data with respect to $w\sigma$* **then**
    $post\_data :=$ suffix of $data$ corresponding to sub_data;
    if *valid(post_data, $t^f_{-(n-1)} \cdots t^f_{-1}$),$\sigma$,$w$* **then**
      return true;
    end
  end
end
return false;
Algorithm 2: The Sequence Validity Algorithm

After using Algorithm 1 to exhaustively construct all valid single-factor patterns for a given set of training onset/termination events, we can construct valid multi-factor patterns by merging the single-factor patterns if the source single-factors have at least two training onset/termination events that define the pattern in common. Although we do not discuss this here, we can fuse the single-factor patterns return by Algorithm 1 to generate multiple factor patterns if the source single-factor patterns have at least two such onset/termination events in common. A branch-and-bound version of the validity algorithm from Algorithm 1 is shown in Algorithm 2. Algorithm 2 is linear in the size of the data and the number of trends in the sequence tested. Besides being used during pattern construction, Algorithm 2 can also be used to test if a pattern matches a country's behavior during forecasting. We discuss this further in Section .

We found in practice that Algorithms 1 and 2 are both highly scalable. In our example use of our Algorithms (discussed below), we implemented the Algorithms in Java and ran the implementations on a COTS laptop over datasets with on the order of a thousand factors in less than a minute.

Forecasting with Concurrent Trend Patterns

After discovering a set of patterns that precede the onset/termination of socio-political violence over some training data, we can use these patterns to make out-of-sample forecasts for the onset/termination of socio-political violence over some test data. We found that a relatively simplistic approach to forecasting is generally very effective - we used a weighting voting mechanism based on where the discovered patterns matched out-of-sample observations in the test data to generate forecasts.

To implement our weighted voting mechanism for a given country at a given time, we used Algorithm 2 over multiple patterns and factors to determine which patterns match the observed factor data leading up to that time. If the number of patterns matching the data exceeds a voting threshold v, then we forecast the onset/termination of socio-political violence in that country at that time. In the next section we discuss a tuning operation to find an appropriate trend weight threshold w and a voting threshold v using training data to generate reasonable out-of-sample forecasting performance.

Tuning Parameters to Improve Performance

Although there are multiple mechanisms to find an appropriate trend weight threshold w and a voting threshold v using training data to generate reasonable out-of-sample forecasting performance, we experimented with a simple approach to tuning the parameters w and v based on the maximization of forecasting performance for the onset/termination of socio-political violence over training data. The setting of these thresholds is important in practice - we found that if we set a high threshold for a zero false positive rate then the forecaster would generate *no* forecasts. Conversely, if we set a low threshold for a zero false negative rate then the forecaster would generate incorrect false positive forecasts at all times.

In particular, in addition to a given set of factor data for Pacific-region countries from 1998-2006, we had data on the occurrences of coups in those regions. We used a 2-dimensional search algorithm to find a trend weight threshold w and a voting threshold v that when used to discover patterns over 1998-2004 and forecast coups over 2005-2006, the f-Measure of the resulting forecasts would be maximized. In particular, we found that by setting $w = 0.20$ and v to be the threshold that generates a precision measure of 0.3 over the training data, forecasts over the corresponding test data will result in no false positive forecasts for coups and only two false negatives (out of three possible coups.)

Examples of Forecasting the Advent and Cessation of Ethnic-Religious Violence using Concurrent Trend Patterns

Using the trend weight threshold w and the voting threshold v that maximized the F-measure of forecasts for the onset of coups, we applied our approach to forecast the onset and termination of ethnic-religious violence in Pacific-region countries. We ran this experiment to forecast the onset and termination of ethnic-religious violence using a set of quarterly sampled Goldstein metric factor data that expressed the relative levels of conflict/coopeation between political groups operating in the countries (such as the government, opposition parties, international organizations, etc...).

We split our data into training and test sets. The training data ran from 1998-2004 and the test data ran from 2005-2006. Over the training data there were onsets of ethnic-religious violence in the following countries at the following times:

- China Q1-2004
- India Q1-2002
- Indonesia Q1-1999
- Solomon Islands Q1-2000
- Solomon Islands Q1-2003
- Sri Lanka Q1-2003

Similarly, there were terminations of ethnic-religious violence in the following countries at the following times:

- India Q1-2004
- Solomon Islands Q1-2001
- Solomon Islands Q1-2004

Using these two sets of events and the $w = 0.2$ threshold, we discovered 55 single-factor patterns for the onset of ethnic-religious violence and 19 single-factor patterns for the cessation of ethnic religious violence. When then used these patterns to forecast the onset/termination of ethnic-religious violence using our threshold voting mechanism.

For the onset of ethnic-religious violence, we generated the following forecasts:

- India Q1-2005
- Nepal Q2-2005
- Taiwan Q4-2006

Over the 2005-2006 test data, the only onset of ethnic-religious violence is in India in the beginning of 2005. There are no true ethnic-religious violence outbreaks in Nepal or Taiwan so we generated two false-positive forecasts. This results in a 100% recall and a 33% recall. It is interesting to note however, that in early to mid-2005 in Nepal there was an uptick in the level of violence associated with the smoldering Maoist insurgency in that country.

For the onset of ethnic-religious violence, we generated the following forecasts:

- China Q1-2005
- Sri Lanka Q1-2006
- Sri Lanka Q3-2006

Over the 2005-2006 test data, the only termination of ethnic-religious violence is in China in the beginning of 2005. There are no true ethnic-religious violence terminations in Sri Lanka so we generated two false-positive forecasts. This results in a 100% recall and a 33% recall. It is similarly interesting to note however there was a dip in the ongoing ethnic Tamil insurgency in the beginning of 2006 in Sri Lanka, but this violence picked up again several months later.

For both forecasting both the onset and termination of ethnic-religious violence, we were able to forecast both occurrences of onset/termination correctly along with a reasonably low false-positive rate. Because we were forecasting over 29 countries and two years, the false-positive rate is approximately one false alarm every 100 country-quarters.

Discussion

In this paper we presented a general approach to identifying patterns of behavior preceding the onset/termination of socio-political violence in sampled temporal factor data. We present a computationally efficient method to discover patterns using a branch-and-bound back-chaining algorithm and a method to forecast from these patterns along with to examples of using these patterns to forecast the onset/termination of ethnic-religious violence.

A particular benefit of our pattern formalism is the ability to test if an entity is following a pattern in a computationally efficient manner. Although our motivating problem of testing if countries are following patterns leading to socio-political violence would generally not require real-time computation due to the very long (potentially monthly) sampling rates, we have found algorithms that can be scalably used in real-time to test if an entity is following a pattern. This broadens the applicability of our sequential pattern methodology to other, faster response time application contexts such as for change detection for part general techno-social systems. In order to demonstrate the wider applicability of our approach we would need to apply our sequential pattern methodology over data sets with more occurrence of events of interest. Although we are not necessarily trying to move away from the small-n pedigree of our intended application context, it is difficult to evaluate the effectiveness of our past pattern discovery efforts without a robust set of test data.

References

Cassandras, C., and Lafortune, S. 1999. *Introduction to Discrete Event Systems*. Boston, MA: Kluwer Academic Publishers.

Chan, S. 2003. Explaining war termination: A Boolean analysis of causes. *Journal of Peace Research* 40(1):49–66.

Das, K., and Schneider, J. 2007. Detecting anomalous records in categorical datasets. In *KDD '07: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, 220–229. New York, NY, USA: ACM.

Dzeroski, S., and Struyf, J., eds. 2007. *Knowledge Discovery in Inductive Databases, 5th International Workshop, KDID 2006, Berlin, Germany, September 18, 2006, Revised Selected and Invited Papers*, volume 4747 of *Lecture Notes in Computer Science*. Springer.

George, A. L., and Bennett, A. 2005. *Case Studies and Theory Development in the Social Sciences*. Boston: The MIT Press.

Huang, K.-Y., and Chang, C.-H. 2008. Efficient mining of frequent episodes from complex sequences. *Inf. Syst.* 33(1):96–114.

Kiernan, J., and Terzi, E. 2008. Constructing comprehensive summaries of large event sequences. In *KDD '08: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, 417–425. New York, NY, USA: ACM.

Kiernan, J., and Terzi, E. 2009. Eventsummarizer: a tool for summarizing large event sequences. In *EDBT '09: Proceedings of the 12th International Conference on Extending Database Technology*, 1136–1139. New York, NY, USA: ACM.

King, G., and Zeng, L. 2001. Logistic regression in rare events data. *Political Analysis* 9(2):137–63.

King, G., and Zeng, L. 2003. Explaining rare events in international relations. *International Organization* 55(3):693–715.

Ragin, C. C. 2000. *Fuzzy-Set Social Science*. University of Chicago Press.

Rohloff, K., and Asal, V. 2008. The identification of sequential patterns preceding the occurrence of political events of interest. In *Proceedings of the 2nd International Conference on Computational Cultural Dynamics (ICCCD)*.

Rohloff, K., and Asal, V. 2009. Computational methods to discover sets of patterns of behaviors that precede political events of interest. In *Proceedings of the AAAI Spring Symposium on Technosocial Predictive Analytics*.