Forecasting International Conflict from Wire News
The Rice Events Data System

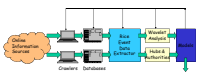
Benedict Lee, Clement Pang, Troy Ruths, Derek Singer, Adam Stepinski, Richard Stoll, Devika Subramanian

Department of Computer Science, Department of Political Science

Problem

Armed conflict between countries remains a significant feature of today's world. Is it possible to forecast international conflict by tracking information on interactions between countries, extracted from publicly available information such as wire news stories?

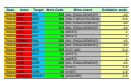
The goal of our research is to develop a system to provide early warning of the onset of serious international conflict. Using on-line media sources, analysis techniques from computer science, statistics, and engineering, and findings from political science, we seek to predict militarized interstate disputes (MIDs) four to eight weeks prior to its outbreak.



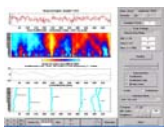
Approach

- Obtain current and archived news stories from AP, AFP, BBC, and Reuters. Index them in large, searchable, text database (currently, 700,000 news articles on events in the Middle East).
- Extract events by analyzing the first line of each relevant wire news story ---determine who did what to whom and when. Events are coded into a discrete conceptual hierarchy called WEIS, developed by political scientists, and used by them for manual coding of news stories.
- Analyze the events data time series to reveal the course of conflict in a region of the world. Such analyses are derived by connecting the dots in "objective" low-level events data.

Singularity Detection in Events Data



- Events data is very noisy; noise arises from errors made by automated coding programs and from events that should not be included in the analysis based on context.
- Events data is highly non-stationary: there are trends, singularities and regime shifts.



Singularity analysis of aggregated events data from the Middle East extracted from Reuters wire stories from 4/1979 to 10/1999.

Time Points

12-25 Sep 1980
102-110 Feb 1988 – Mar 1988
245-248 Jan 1991 –Feb 1991
403-411 Sep 1994 – Oct. 1994
500-end Jul 1998

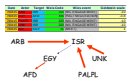
Conflict(s)

Onset and initial phases of Iran-Iraq War
End of Iran-Iraq War
Gulf War
USA-Iraq and Israel-Lebanon conflicts
Operation Desert Fox and the Intifada

- Aggregated events data has information --- singularities detected by the Mallat's modulus maximum technique correspond to major conflicts!.
- Open question: Can we predict the singularities?

Hubs and Authorities Analysis

- Key idea: use tools from social network analysis to understand temporal variations in key actors/targets in events data.



Model interactions between countries in a weighted directed graph.

- A hub node is an important initiator of events; an authority node is an important target of events.
- Hypothesis Identifying hubs and authorities over a particular temporal chunk of events data tells us who the key actors and targets are!
- The HITS algorithm computes hub score (h) and authority score (a) for every node in a weighted directed graph.

$$h^{(t+1)}(p) = \sum_{p \to q} w(p,q) * a^{t}(q)$$

$$a^{(t+1)}(q) = \sum_{p \to q} w(p,q) * h^{(t)}(p)$$

- On a data set of 177,344 events coded from Reuters stories pertaining to the Middle East from April 1979 to October 2003, the top hubs and authorities are the principals in all but 3 of the 41 conflicts recorded in the MID database over this time period. The number and size of the connected components in the event interaction graph predict the outbreak of 38 of 41 of these major conflicts in a 4-8 week horizon.