SAND2012-6729C

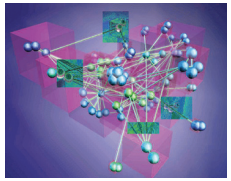
Leveraging Sociological Models for Prediction

Rich Colbaugh **Kristin Glass**

Sandia National Laboratories
New Mexico Institute of Mining and Technology

August 2012

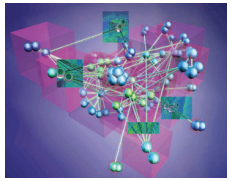
Introduction



Motivation

- Many informatics applications involve making predictions about human behavior, for instance concerning future occurrences or ongoing but "hidden" activities.
- Machine learning (ML) has proved useful for such forecasting and "predicting the present" tasks.
- It is natural to suspect that the predictive performance of ML methods might be improved by incorporating auxiliary information obtained from the social sciences.

Introduction



Objective

Explore extent to which theories and models from the social sciences can be leveraged to enable useful predictions concerning human behavior.

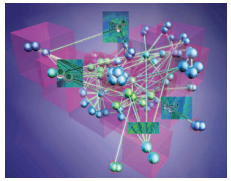
Outline for talk

- Main example: predicting adversarial network structure and dynamics

  semi-supervised classifier for prediction, structural balance theory to specify classifier feature-weight "priors", sample results – edge-sign prediction and network fission forecasting.

- Two additional examples:
  - early warning for "complex" contagions;
  - predictive defense against adaptive adversaries.

Adversarial Social Networks



Background

There is considerable motivation for studying social networks with *signed edges*, in which positive/negative edges reflect friend/foe relationships:

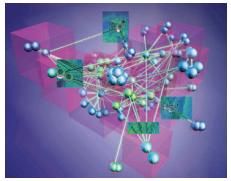
- inferring edge-signs (predicting the present) is useful for a variety of intelligence applications;
- predicting the way an adversarial network will fission when stressed is central to many "influence" campaigns.

See, for example, [Leskovec et al. 2010] for motivation and discussion, and [Colbaugh/Glass 2012] for security applications.

Goal: prediction with limited data

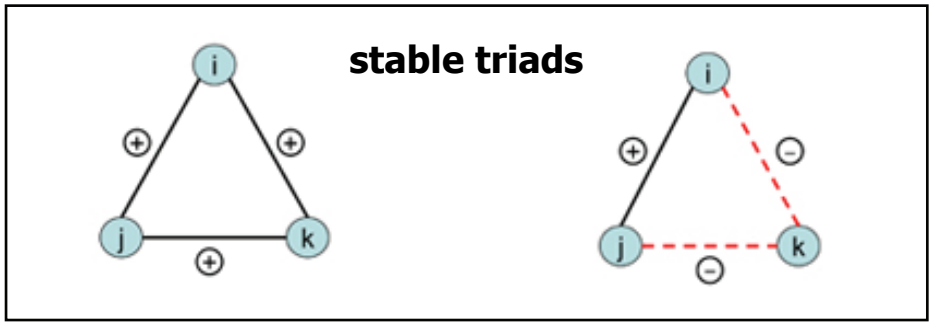
Exploit structural balance theory to 1.) infer edge-signs and 2.) predict group fission outcomes, in situations where only limited data is available concerning the network of interest.

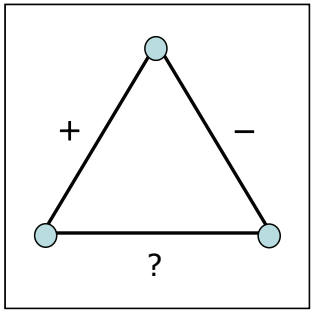
Adversarial Social Networks



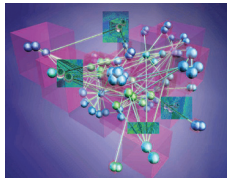
Structural balance theory

Structural balance theory (SBT) [Heider 1946] posits that edge triads in friend/foe networks will be stable if they consist of an odd number of friendly edges (one or three "positive" edges) and will be unstable otherwise.



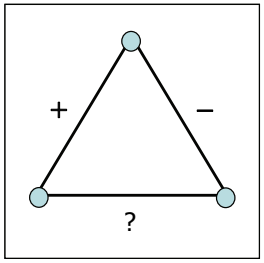


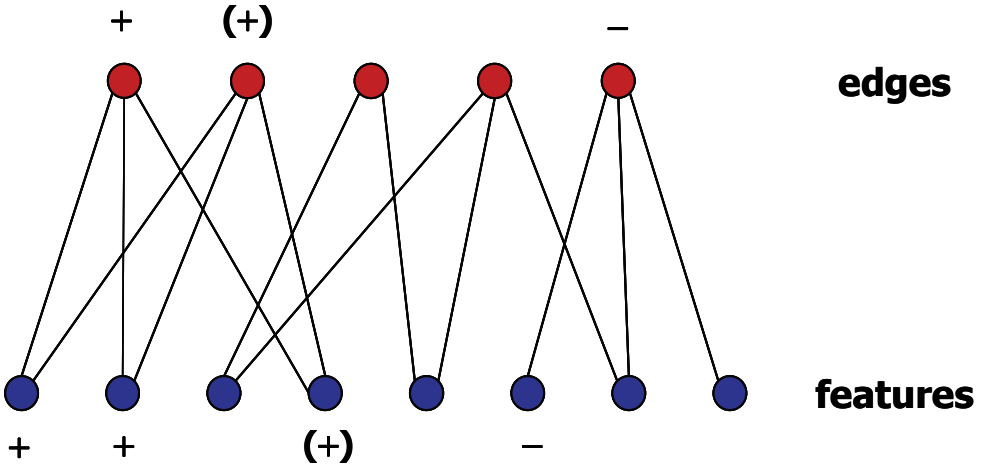
Adversarial Social Networks



Predicting edge-signs via semi-supervised learning

- Data model G_b:



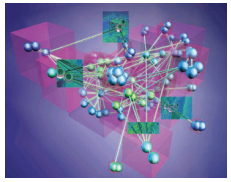


- Quantification:

$$\min_{c_{aug}} \; c_{aug}^{T} L c_{aug} + \beta_1 \sum_{i=1}^{n_1} (d_{est,i} - d_i)^2 + \beta_2 \sum_{i=1}^{|F_1|} (c_i - w_i)^2$$

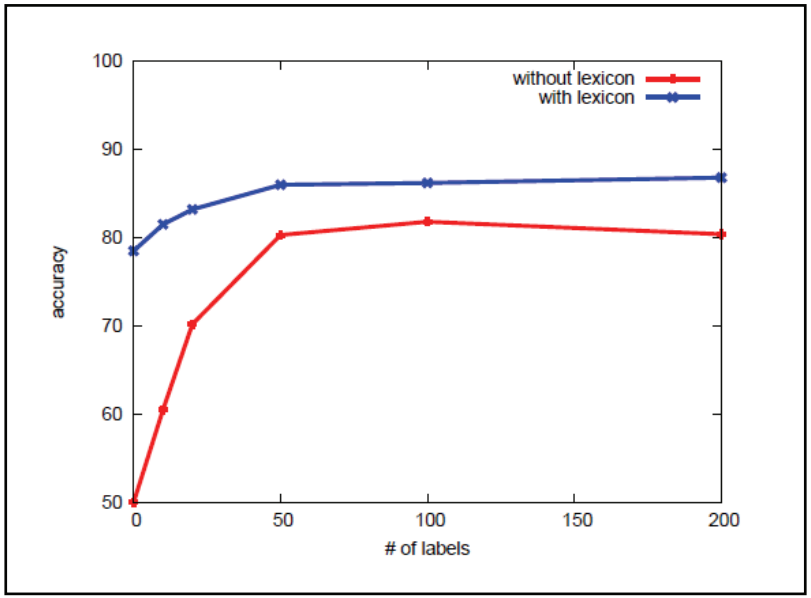
where $c_{aug} = [d_{est}^{T} \quad c^{T}]^{T}$. Sign prediction for new edge $\mathbf{x}$ is $\text{sign}(c^{T}\mathbf{x})$.

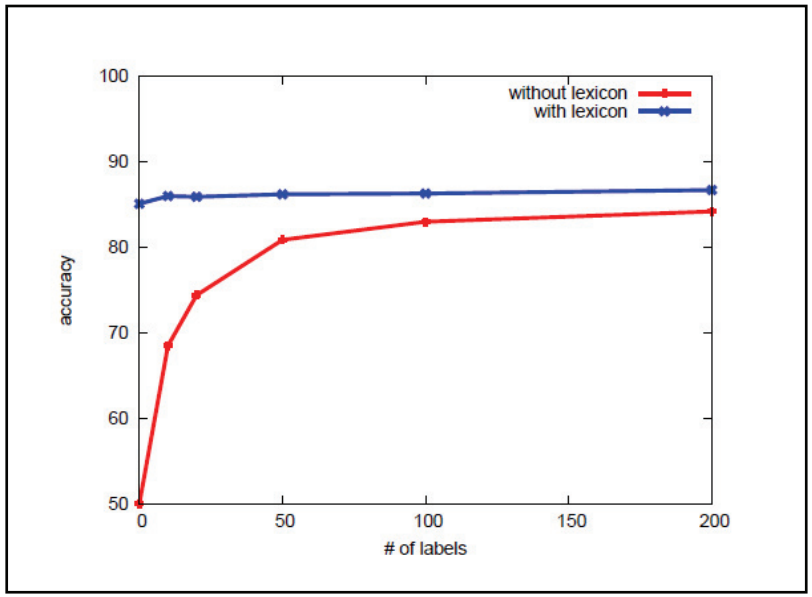
Adversarial Social Networks



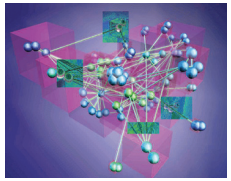
Predicting edge signs: sample results

Edge-sign prediction accuracy for Wikipedia "voting network" (left, 100K edges) and "editing network" (right, 750K edges) (balanced datasets, two-fold cross-validation).





Adversarial Social Networks

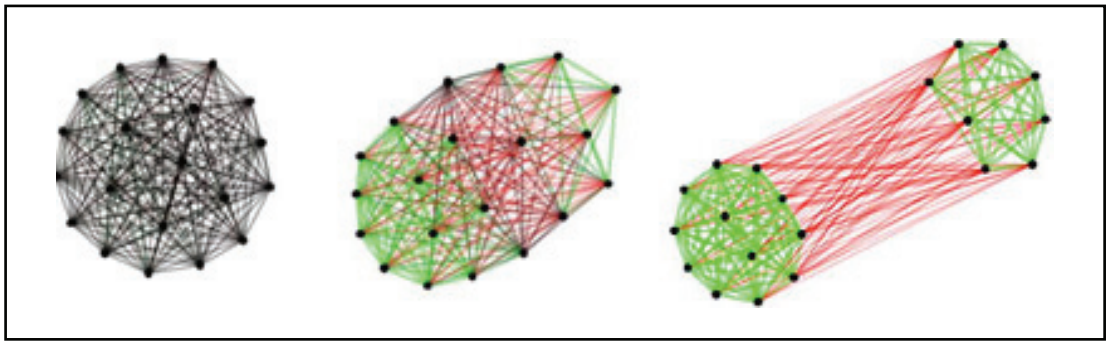


Predicting group fission

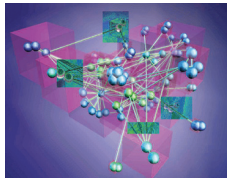
- Problem: given $G_s = (V, E)$ characterizing relationships between entities (individuals, countries), only some of which are known, predict the way G_s is likely to fission if subjected to stress.
- Interesting result: given fully-specified signed adjacency matrix Z_0 for G_s, [Marvel et al. 2011] evolve simple matrix dynamics below until singularity Z_s, and interpret Z_s as defining the fission – partition has all positive intra-group edges and all negative inter-group edges.

$$dz_{ij}/dt = \sum_k z_{ik} z_{kj}.$$

$$dZ/dt = Z^2, \quad Z(0)=Z_0.$$



Adversarial Social Networks

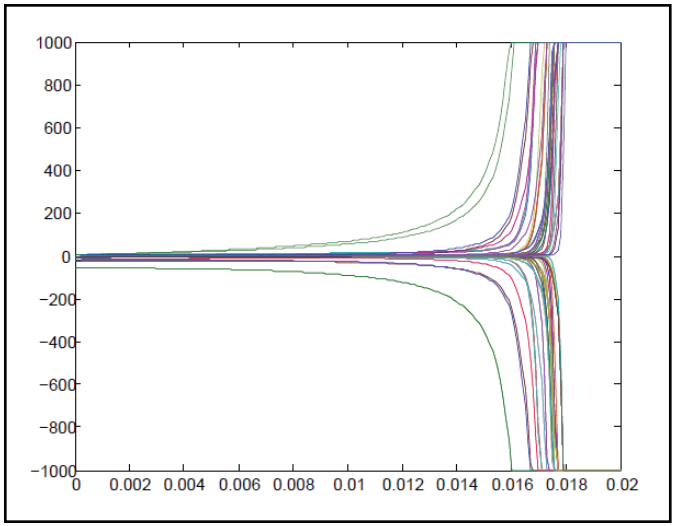


Predicting group fission with limited data

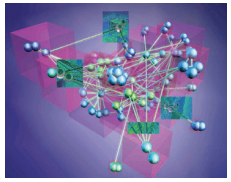
1. Given network of interest G_s and estimates for some of the edge-signs, construct initial estimate Z_0 for signed adjacency matrix via SS-ML.
2. Evolve $dZ/dt = Z^2$, $Z(0) = Z_0$, until it reaches singularity Z_s.
3. Interpret Z_s as defining predicted network fission: factions have positive intra-group edges and negative inter-group edges.

Sample result

Method correctly predicts ultimate alignment of Allied/Axis countries in WWII, even if only 15% of edge-signs are known in advance (data from [Axelrod/Bennett 1993]).

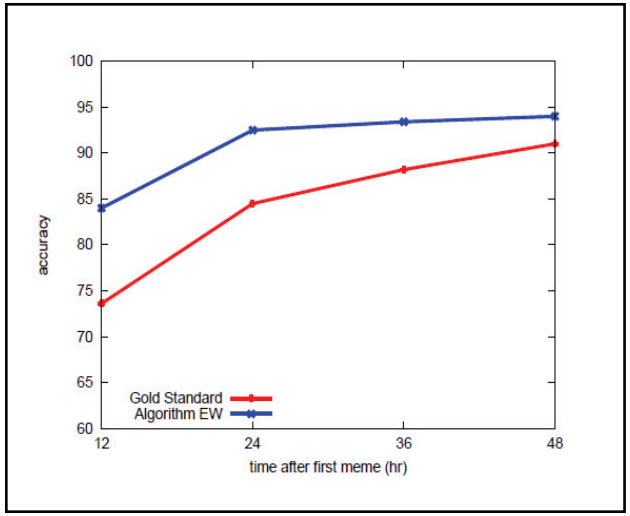


Additional Examples



Early warning for complex contagions

- Goal: distinguish complex contagions that will "go viral" from those which won't early in contagion lifecycle.
- Social science: complex contagion model [Centola 2007, Romero et al. 2011].



Predictive defense against adaptive adversaries

- Goal: develop ML algorithms which enable effective defense against intelligent, adaptive adversaries.
- Social science: combine ML with Stackelberg game theory [Peters 2008].