Dynamic Social Network Group Anomaly Detection Using Hierarchical Bayesian Model

Rose Yu, Xinran He and Yan Liu
University of Southern California
{qiyu, xinranhe, yanliu.cs}@usc.edu

USC UNIVERSITY OF SOUTHERN CALIFORNIA

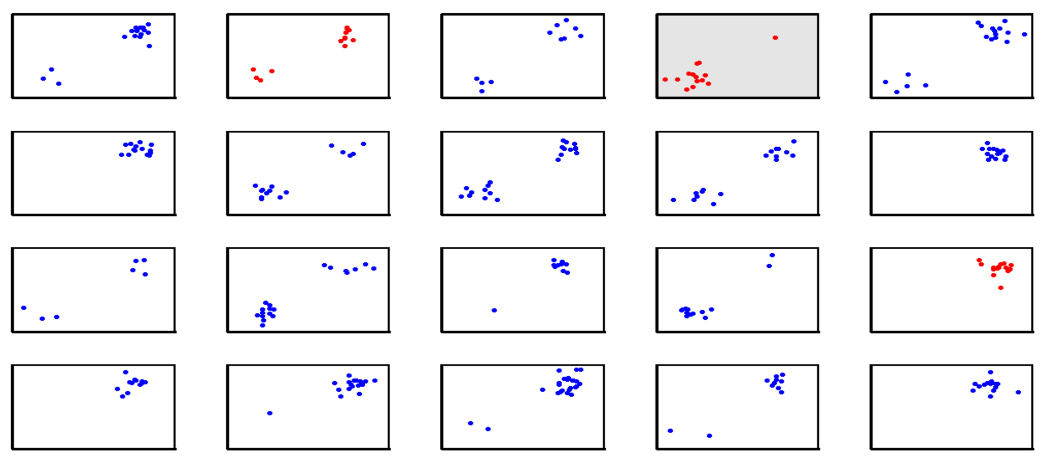
1. Abstract

Detecting anomalies on social network is important and challenging for many settings such as network intrusion, fraud detection and cyber security. Traditionally, anomaly detection algorithms aim at finding anomalous individuals. However, social networks are dynamic in nature and rich in group structures, the incident of anomalies usually take place at group level. In this paper, we study the evolving latent structure of social network groups and formulate the group anomaly detection problem into the change point detection framework. Specifically, we take a generative approach and propose a novel Dynamic Groupwise Latent Anomaly Detection (DGLAD) model. We derive a variational Bayesian inference algorithm of DGLAD and demonstrate its effectiveness on synthetic and real world data sets.

2. Motivation

Detecting anomalous latent structure in groups have been studied in [4], [5] and [6].

However, existing approaches require the group information to be given beforehand and can only deal with groups with fixed size, which is not feasible for dynamic social networks.



An example of latent group anomaly in two components Mixture Gaussian data. Shaded box is the anomalous group.

Two viewpoints of community in the social network:

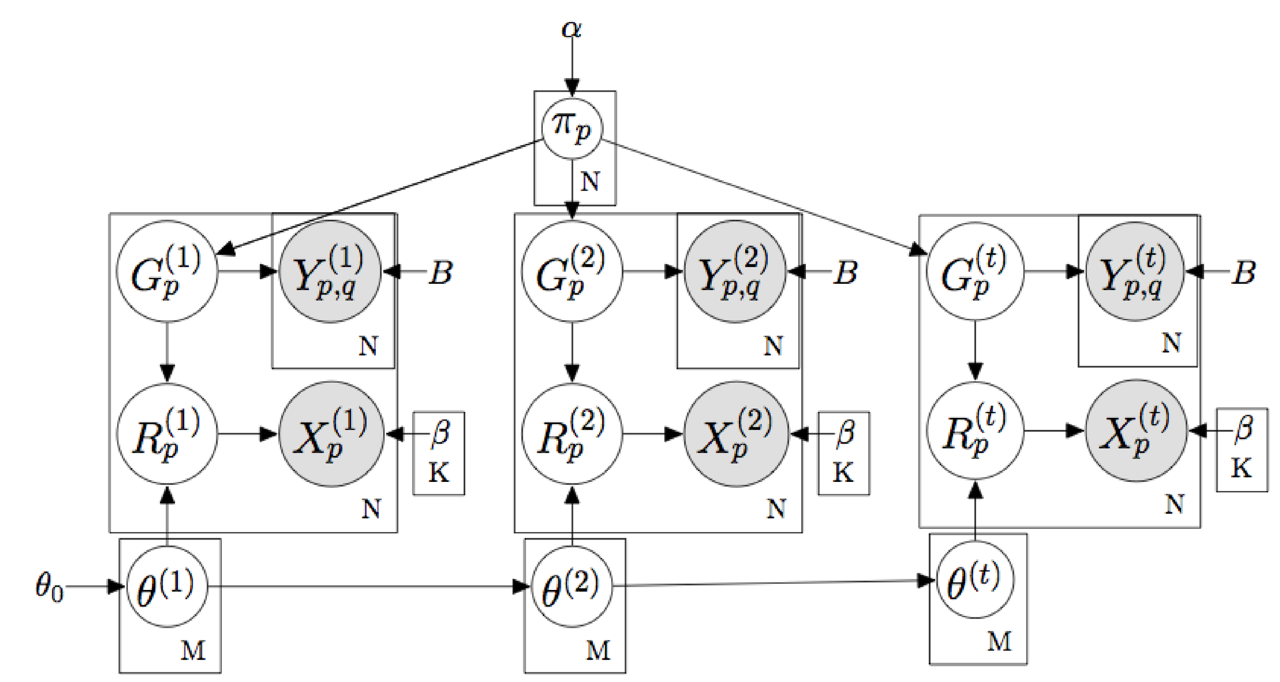
- people in a community have strong ties (pair-wise)
- people in a community have largely overlapped attributes. (point-wise)

Consider a social network with N nodes. For every node p, we observe its attributes at time t : $X_p^{(t)}$ and its link to neighbor q : $Y_{(p,q)}^{(t)}$. We associate each person with a group identity $G_p^{(t)}$ and a role identity $R_p^{(t)}$. The latent structure we are interested in is the role mixture rate θ within the group. We define the anomaly score for group m as

$$\text{Anomaly Score} = \|\theta_m^{(t)} - \theta_m^{(t-1)}\|_2$$

This definition captures the evolution of the mixture rate over time and detects the anomalous groups whose role mixture rates suffer dramatic change.

3. Dynamic Groupwise Latent Anomaly Detection



DGLAD Model

Generative Process of DGLAD

for $t = 1 \to T$ **do**
  for $m = 1 \to M$ **do**
    Draw $\theta_m^{(t)} \sim \text{Gaussian}(\theta_m^{(t-1)}, \sigma)$
  end for
  for individual $p = 1 \to N$ **do**
    Draw membership distribution $\pi_p^{(t)} \sim \text{Dir}(\alpha)$
    Draw $G_p^{(t)} \sim \text{Multinomial}(\pi_p)^{(t)}$
    for individual $q = 1 \to N$ **do**
      Sample $Y_{(p,q)}^{(t)} \sim \text{Binomial}\left((G_p^{(t)})^T B G_q^{(t)}\right)$
    end for
    Draw $R_p^{(t)} \sim \text{Multinomial}(R_p^{(t)} | \mathcal{S}(\theta^{(t)}_{G_p^{(t)}}))$
    Draw $X_p^{(t)} \sim \text{Multinomial}(X_p^{(t)} | \beta_{R_p^{(t)}})$
  end for
end for

4. Variational Inference

We denote model parameters $\Theta = \{\alpha, B, \beta_{1:K}\}$. Visible variables $v = \{X_{1:N}^{(1:T)}, Y_{1:N,1:N}^{(1:T)}\}$, hidden variables $h = \{\pi_{1:N}, G_{1:N}^{(1:T)}, R_{1:N}^{(1:T)}, \theta_{1:M}^{(1:T)}\}$. The complete likelihood of observed data and latent variables can be written as :

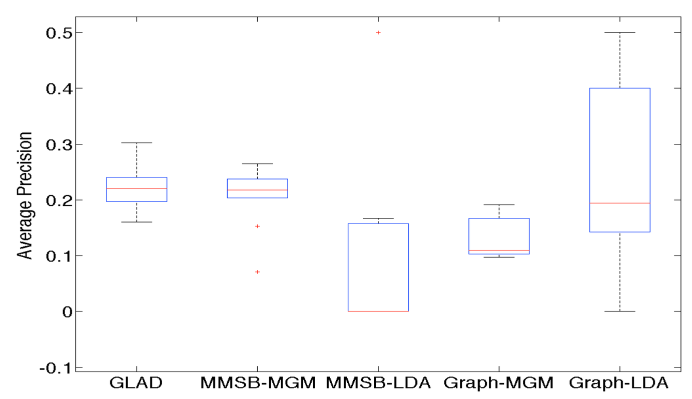
$$\begin{aligned} p(v,h|\Theta) &= \prod_t \prod_p p(\pi_p|\alpha) p(G_p^{(t)}|\pi_p) \times \prod_t \prod_{p,q \neq p} p(Y^{(t)}|G_p^{(t)}, G_q^{(t)}, B) \\ &\times \prod_m p(\theta_m^{1:T}|\theta_0, \sigma) \times \prod_t \prod_p p(X_p^{(t)}|R_p^{(t)}, \beta_{1:K}) p(R_p^{(t)}|G_p^{(t)}, \theta_{1:M}^{(t)}) \end{aligned}$$

Computation of the maximizer for $p(v|\Theta)$ requires integration over all the latent variables in above equation, which is intractable here, therefore we apply Variational Bayesian[3] approach to approximate the posterior. We choose the variational distribution $q(h|\Theta)$ in the form of

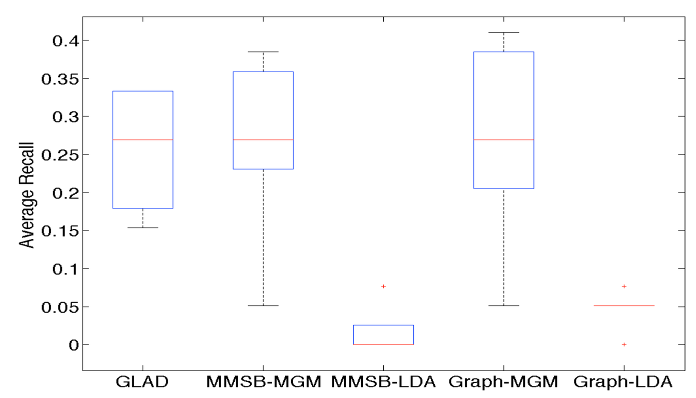
$$q(h|\Theta) = \prod_t \prod_p q(\pi_p|\gamma_p) q(G_p^{(t)}|\lambda_p^{(t)}) q(R_p^{(t)}|\mu_p^{(t)}) \times \prod_m p(\theta_m^{1:T}|\hat{\theta}^{1:T})$$

then maximize over $< \log p(v,h|\Theta) >_q$ by taking the partial derivatives to infer the hidden variables. For the component of sequential structure in the variational distribution $\hat{\theta}^{(1:T)}$, we resort to the variational Kalman Filter technique [2].

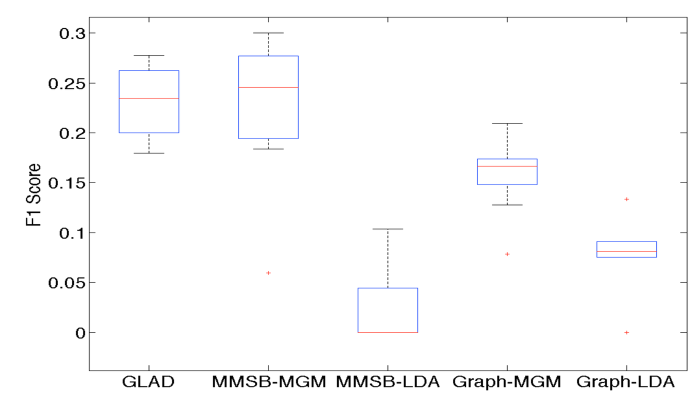
5. Results on Synthetic Dataset



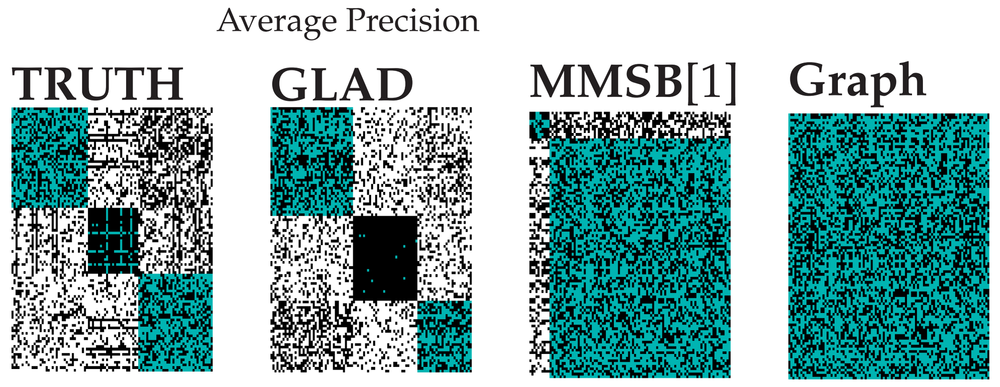
Average Precision

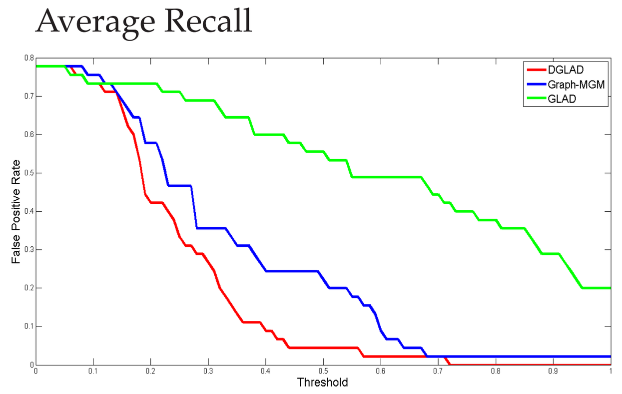


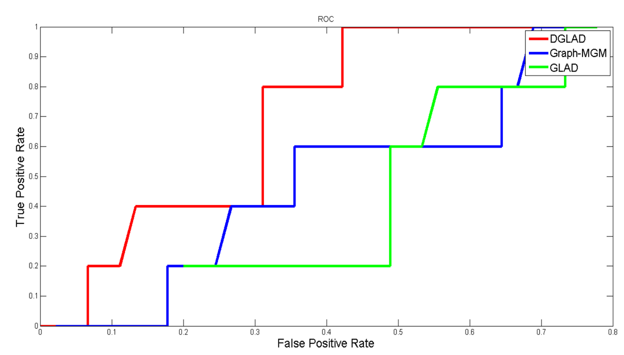
Average Recall



F1 score

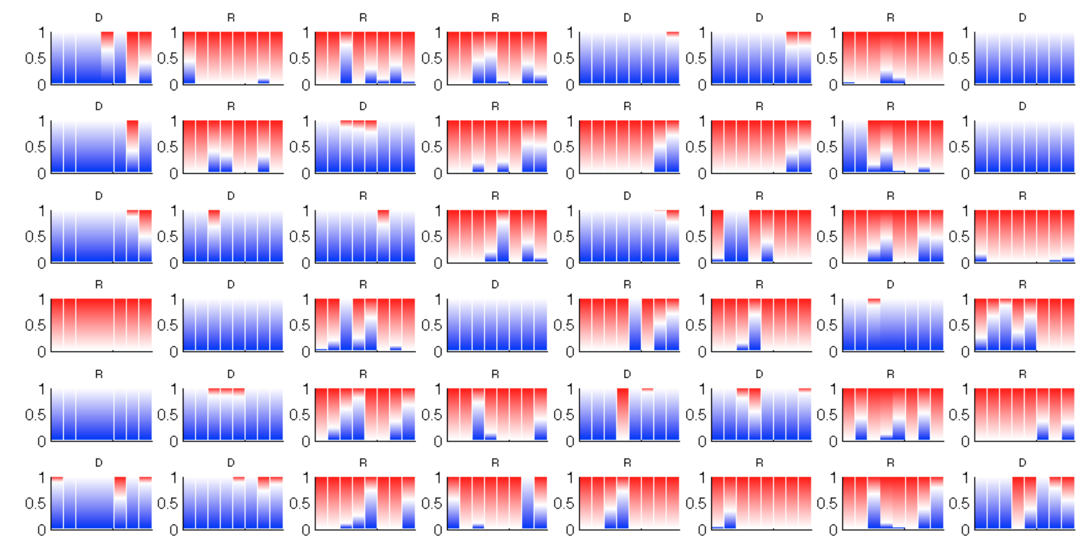




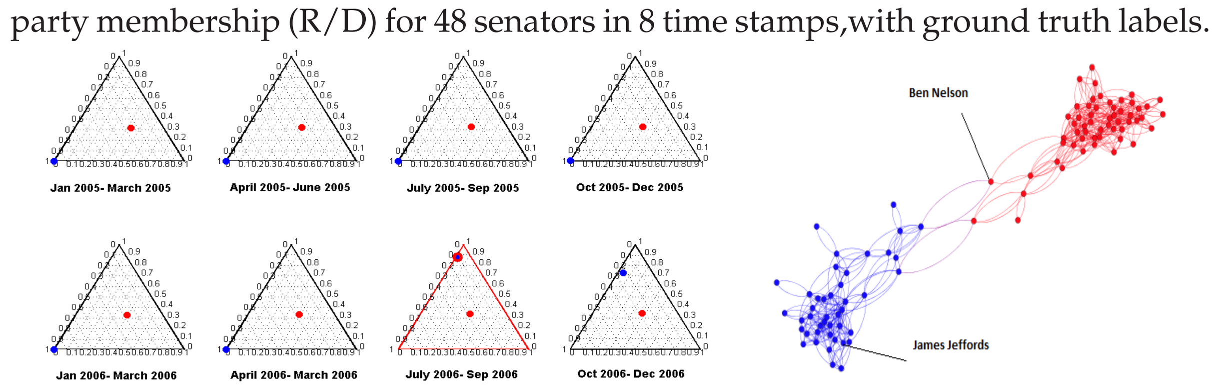


6. Results on Senator votes

100 senators' voting records spanning from Jan 1st 2005 to Dec 31st 2006, with 8 snapshots of senator common votes network and 6 voting features.



Group discovery results on senators' voting using GLAD. Inferred party membership (R/D) for 48 senators in 8 time stamps,with ground truth labels.



In September 2006, Democratic senator Joseph Lieberman lost the primary election and became a independent Democratic. James Jeffords left the Republican in 2001 to become an Independent and began caucusing with the Democrats. Ben Nelson's votes often placed him at odds with his party.

7. Results on ACM publications

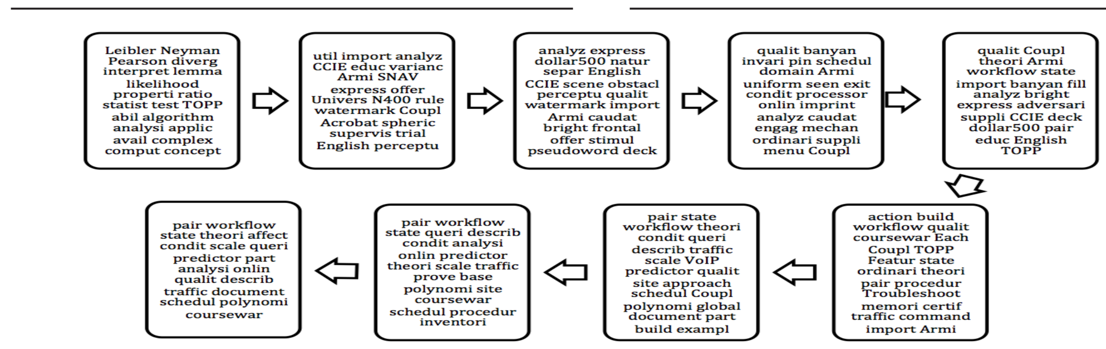
31574 paper abstracts from ACM database, including 4474 authors. We extract 8024 bag of words features for each author and construct 10 years coo-authorship network.

Mixture Rate Change of GLAD and DGLAD

Model	Year 2001	Year 2002	Year 2003
GLAD	0.4950	0.5293	0.3418
DGLAD	**0.3593**	**0.2417**	**0.2299**
Model	Year 2004	Year 2005	Year 2006
GLAD	0.3536	0.2591	0.3186
DGLAD	0.5515	**0.1928**	**0.2693**
Model	Year 2007	Year 2008	Year 2009
GLAD	0.4212	0.3628	0.4724
DGLAD	**0.3440**	**0.2201**	**0.3697**

Prediction loglikelihood of GLAD and DGLAD

Model	Year 2001	Year 2002	Year 2003
GLAD	0.4950	0.5293	0.3418
DGLAD	**0.3593**	**0.2417**	**0.2299**
Model	Year 2004	Year 2005	Year 2006
GLAD	0.3536	0.2591	0.3186
DGLAD	0.5515	**0.1928**	**0.2693**
Model	Year 2007	Year 2008	Year 2009
GLAD	0.4212	0.3628	0.4724
DGLAD	**0.3440**	**0.2201**	**0.3697**



8. References

[1] E. M. Airoldi, D. M. Blei, S. E. Fienberg, and E. P. Xing. Mixed membership stochastic blockmodels. 9, 2008.

[2] D. M. Blei and J. D. Lafferty. Dynamic topic models. 2006.

[3] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul. An introduction to variational methods for graphical models. 37(2), 1999.

[4] K. Muandet and B. SchÃűlkopf. One-class support measure machines for group anomaly detection. March 2013.

[5] L. Xiong, B. Poczos, J. Schneider, A. Connolly, and J. Vanderplas. Hierarchical probabilistic models for group anomaly detection. 15:789âĂŞ797, 2011.

[6] L. Xiong, B. PÃşczos, and J. Schneider. Group anomaly detection using flexible genre models. 2011.