Attribute Fusion in a Latent Process Model for Time Series of Graphs

Carey E. Priebe, Nam H. Lee, Youngser Park, Minh Tang

Johns Hopkins University
Department of Applied Mathematics and Statistics
Baltimore, Maryland 21218-2682 USA

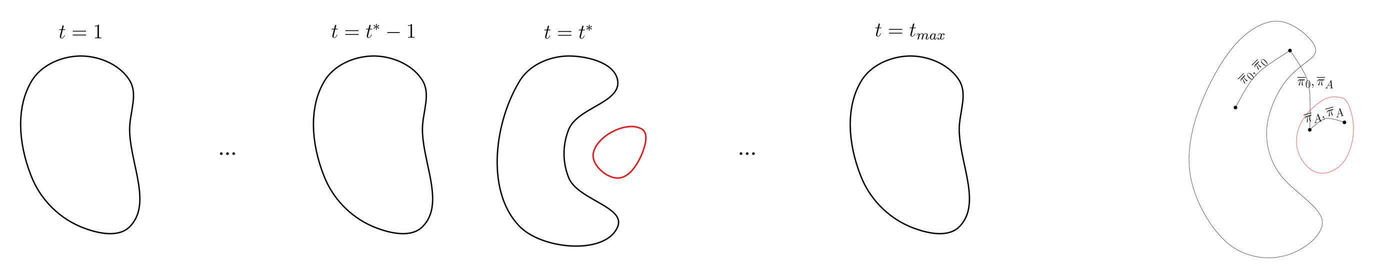
Abstract

We consider anomaly/change point detection given a time series of graphs with categorical attributes on the edges. Various attributed graph invariants are considered, and their power for detection as a function of a linear fusion parameter is presented.

1. Time Series of Attributed Graphs

Given a time series of attributed graphs $G(t) = (V, \varphi(\cdot, t))$, $t = 1, 2, \cdots$, where the vertex set $V = [n] = \{1, \cdots, n\}$ is fixed throughout and edge attribution functions $\varphi(\cdot, t) : \binom{V}{2} \to \{0, \cdots, K\}$ are time-dependent, we wish to detect anomalies and/or change points. Let us consider vertices to represent "actors," and an edge between vertex u and vertex v at time t ($uv \in E(t)$, where the edge set $E(t)$ is given by $E(t) = \{uv \in \binom{V}{2} : \varphi(uv, t) > 0\}$) represents the existence of a communications event between actors u and v during the time period $(t-1, t]$. Categorical edge attributes $\varphi(uv, t)$, when non-zero, represent some mode of the communication event between actors u and v during $(t-1, t]$; for instance, a topic label derived from the content of the communication.

The specific anomaly we will consider is the "chatter" alternative – a small (unspecified) subset of vertices with altered communication behavior during some time period in an otherwise stationary setting, as depicted in Figure 1(a). This figure notionally depicts the entire vertex set V behaving in some null state for $t = 1, \cdots, t^* - 1$; then, at time t^*, a collection of vertices $V_A \subset V$ ($|V_A| = m, 2 \leq m \ll n$) exhibits probabilistically altered behavior. (The remaining vertices remain in their null state throughout.) Our statistical inference task is to determine whether or not there has emerged a chatter group at some time $t = t^*$.



(a) Time series of graphs   (b) "Kidney & Egg"

Figure 1: *Notional depiction of a time series of graphs in which the entire vertex set V behaves in some null state for $t = 1, \cdots, t^* - 1$ and then, at time t^*, a subset of vertices V_A exhibits a change in behavior.*

Our latent process model in [1] produces a dependent time series of attributed graphs $G(t)$, each of which is itself a latent position model with conditionally independent edges given a continuous time, finite state stochastic process $\{X_v(\tau)\}_{v \in V, \tau \leq t}$. The model allows two simplifying approximations; a second-order (central limit theorem) approximation with temporally independent attributed random graphs each of which is itself a random dot product latent position model, and a first-order (law of large numbers) approximation with temporally independent attributed random graphs each of which is itself an independent edge random graph model.

The simplicity of the first-order approximation, depicted in Figure 1(b) for the special case of the *kidney-egg* stochastic blockmodel structure considered herein for the anomaly, provides a useful framework for description. If the vertex processes $\{X_v(t)\}_{v \in V}$ are independent and identical, with stationary probability vector $\pi_0 = [\pi_{0,0}, \pi_{0,1}, \cdots, \pi_{0,K}]'$, then the first-order approximation produces a temporally independent series of homogeneous independent edge attributed random graphs with $P[\varphi(uv, t) > 0] = \langle \overline{\pi}_0, \overline{\pi}_0 \rangle$ and $P[\varphi(uv, t) = k | \varphi(uv, t) > 0] = \pi_{0,k}^2 / \langle \overline{\pi}_0, \overline{\pi}_0 \rangle$, where $\overline{\pi}_0 = [\pi_{0,1}, \cdots, \pi_{0,K}]'$. The vertex processes $\{X_v(t)\}_{v \in V_A}$ change at time $t^* - 1$, taking on stationary probability vector π_A, so that $G(t^*)$ is a *kidney-egg* independent edge random graph with attribute probabilities defined using $\overline{\pi}_0$ and $\overline{\pi}_0$ for $u, v \in V \setminus V_A$, $\overline{\pi}_0$ and $\overline{\pi}_A$ for $u \in V \setminus V_A$, $v \in V_A$, and $\overline{\pi}_A$ and $\overline{\pi}_A$ for $u, v \in V_A$.

2. Invariants

In [2] the scan statistic graph invariants are introduced and applied to the problem of detecting chatter anomalies in time series of Enron graphs. In [3] various graph invariants (size, maximum degree, scan statistic, etc.) are considered for their power as test statistics and it is demonstrated via Monte Carlo that no single invariant is uniformly most powerful, while in [4] it is demonstrated that asymptotics can provide misleading comparative power analysis for size vs. maximum degree except for astronomically large graphs; see also [5] for a summary.

In this paper, we consider the problem of detecting chatter anomalies in time series of graphs using attributed invariants. Specifically, we consider linear attribute fusion with parameter $\lambda \in \mathbb{R}^K$ via

$$\begin{aligned} size_\lambda(G(t)) &= \sum_{k=1}^K \lambda_k \sum_{uv} I\{\varphi(uv, t) = k\}, \\ maxd_\lambda(G(t)) &= \max_v \sum_{k=1}^K \lambda_k \sum_{u \in N[v]} I\{\varphi(uv, t) = k\}, \\ scan_\lambda(G(t)) &= \max_v \sum_{k=1}^K \lambda_k \sum_{u,w \in N[v]} I\{\varphi(uw, t) = k\}, \end{aligned}$$

where $N[v] = \{u : uv \in E(t)\} \cup \{v\}$ is the closed neighborhood of vertex v in graph $G(t)$.

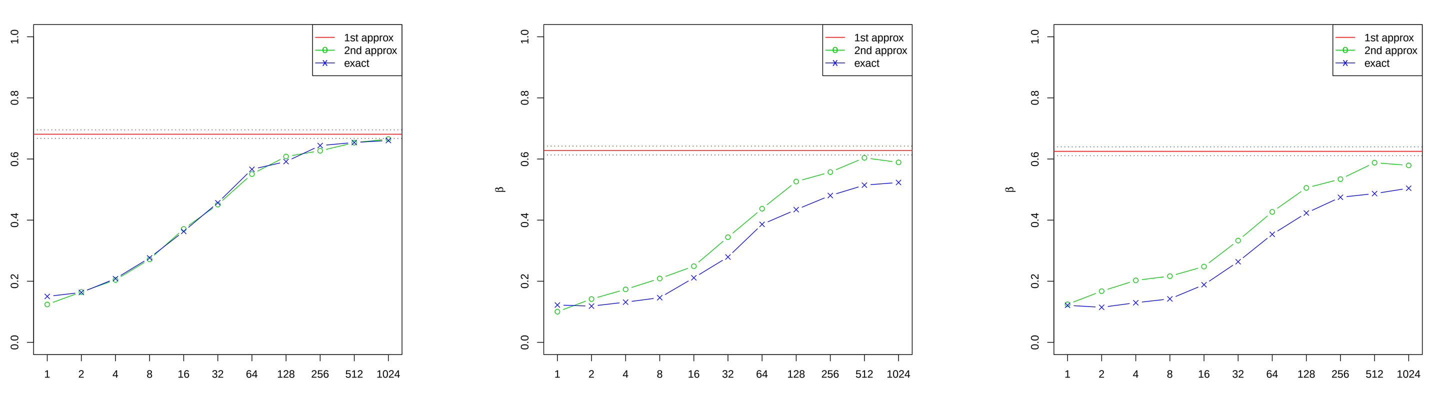
We present experimental results for anomaly detection on time series of simulated data from the model in [1], and demonstrate that optimal attribute fusion depends on invariant.

3. Example

We present here a simple yet illustrative Monte Carlo experiment using the following parameters: $K = 2, n = 100, m = 9, \ell = 10$, with stationary probability vectors $\pi_0 = [0.60, 0.10, 0.30]'$, $\pi_A = [0.35, 0.25, 0.40]'$, and transition matrices

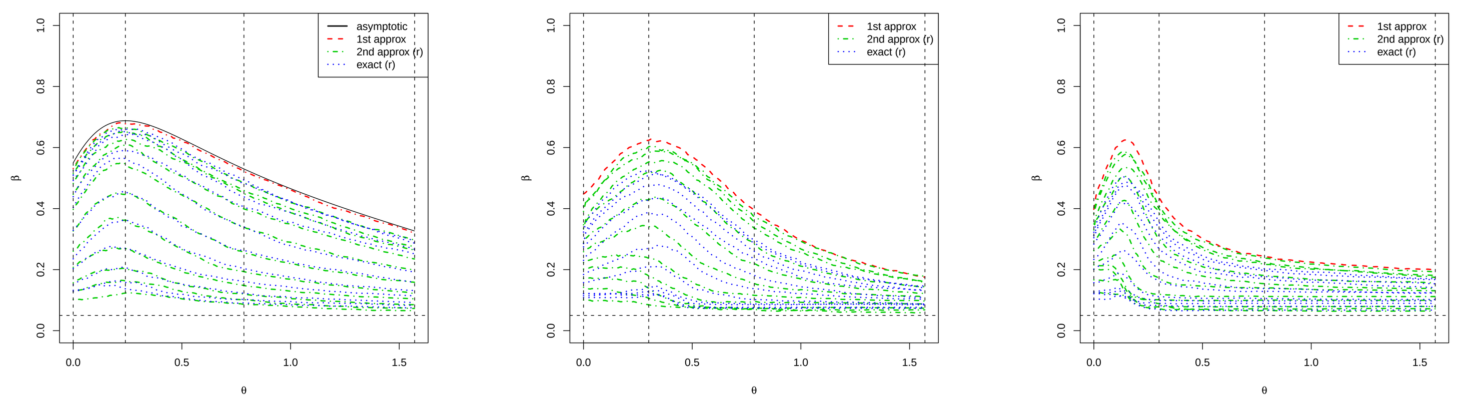
$$Q_0/r = \begin{bmatrix} -2/3 & 1/6 & 1/2 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}, \; Q_A/r = \begin{bmatrix} -13/7 & 5/7 & 8/7 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}.$$

Power estimates for our three attribute fusion statistics for this example are presented in Figures 2 and 3. We consider first approximation, second approximation, and exact model power estimates obtained via Monte Carlo simulation. Figure 2 shows power as a function of the vertex process rate parameter r. Figure 3 shows power as a function of angle θ, where $\lambda = (\cos(\theta), \sin(\theta))$.



(a) $size_\lambda$ (b) $maxd_\lambda$ (c) $scan_\lambda$

Figure 2: *Power $\beta(\theta_r^*, r)$ at $\theta_r^* = \arg\max_\theta \beta(\theta, r)$ as a function of r for the three invariants at test size $\alpha = 0.05$. (10000 Monte Carlo replicates yields $\sigma_{\widehat{\beta}} \leq 0.005$.) The horizontal line represents first-order approximation $\pm$ three standard deviations, and the two curves represent the second approximation and exact model.*



(a) $size_\lambda$ (b) $maxd_\lambda$ (c) $scan_\lambda$

Figure 3: *Power β as a function of angle θ for $\lambda = (\cos(\theta), \sin(\theta))$. The plot shows analytic asymptotic result (black, available for $size$ only), and first approximation (red), second approximation (green), and exact model (blue) estimates for the various invariants via Monte Carlo. (10000 Monte Carlo replicates yields $\sigma_{\widehat{\beta}} \leq 0.005$.) (Horizontal line is $\alpha = 0.05$.) The four vertical lines correspond to $\theta \in \{0, \pi/2, \pi/4, \theta_{size}^* \approx 0.24\}$, which correspond to attribute weighting schemes of: consider only attribute 1, consider only attribute 2, consider both attributes equally, and the attribute weight ratio of approximately 3.7:1.*

4. Conclusions

One notable implication of this work, inferred from Figure 2, is that inferential performance in the mathematically tractable first- and second-order approximation models does indeed provide guidance for methodological choices applicable to the exact (realistic but intractable) model; furthermore, to the extent that the exact model is realistic, we may tentatively conclude that approximation model investigations have some bearing on real data applications.

Our main result regarding linear attribute fusion is that the optimal linear fusion parameter depends on the invariant considered. In particular, the results depicted in Figure 3 yield $\widehat{\theta}_{size}^* = 0.22$ (compared to the theoretical $\theta_{size}^* \approx 0.24$), $\widehat{\theta}_{maxd}^* = 0.31$, and $\widehat{\theta}_{scan}^* = 0.15$. These optimal fusion parameter differences are statistically significant; combining this result with the "no uniformly most powerful invariant" result, we conclude that optimal linear attribute fusion theory requires significant additional development. Toward this end, the approximation models from [1] promise to be of assistance.

In addition to the social network analysis scenario considered herein for illustration (wherein vertices represent individual actors or organizations), hypothesis testing on time series of attributed graphs has application in areas as diverse as connectome inference (wherein vertices are neurons or brain regions) and text processing (wherein vertices represent authors or documents). These applications, and others, may benefit from generalization to directed, hyper, multi, loopy, weighted graphs, as well as the consideration of inference with *errorful* attributes through an attribute confusion matrix.

References

[1] N. H. Lee and C. E. Priebe, "A Latent Process Model for Time Series of Attributed Random Graphs," *Statistical Inference for Stochastic Processes*, forthcoming.

[2] C. E. Priebe, J. M. Conroy, D. J. Marchette, and Y. Park, "Scan Statistics on Enron Graphs," *Computational and Mathematical Organization Theory*, vol. 11, no. 3, pp. 229–247, 2005.

[3] H. Pao, G. A. Coppersmith, and C. E. Priebe, "Statistical Inference on Random Graphs: Comparative Power Analyses via Monte Carlo," *Journal of Computational and Graphical Statistics*, vol. 20, no. 2, pp. 395–416, 2011.

[4] A. Rukhin and C. E. Priebe, "A Comparative Power Analysis of the Maximum Degree and Size Invariants for Random Graph Inference," *Journal of Statistical Planning and Inference*, vol. 141, pp. 1041–1046, 2011.

[5] C. E. Priebe, G. A. Coppersmith, and A. Rukhin, "You Say Graph Invariant, I Say Test Statistic," *ASA Sections on Statistical Computing Statistical Graphics SCGN Newsletter*, vol. 21, no. 2, December 2010.

This work is supported in part by the Johns Hopkins University Human Language Technology Center of Excellence and a National Security Science and Engineering Faculty Fellowship (NSSEFF) award. Relevant manuscripts available at `http://www.cis.jhu.edu/~parky/SSP/`.