IBM Research

Proximal Graphical Event Model

Debarun Bhattacharjya, Dharmashankar Subramanian, Tian Gao

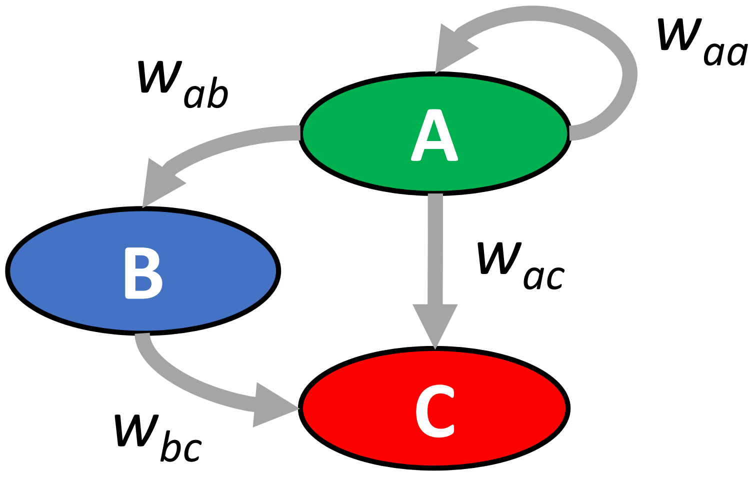
Objective: To learn statistical and causal relationships between event types in the form of graphical models using event datasets

Event datasets: Occurrences of various event types over time

- Examples: web logs; customer transactions; network notifications; political events; financial events; insurance claims; health episodes; other medical events
- Notation: $\boldsymbol{D} = \{l_i, t_i\}, i = 1, \ldots, N; l_i \in L, |L| = M$
  - Assume it is temporally ordered b/w time $t_0 = 0 \leq t_1$ and $t_{N+1} = T \geq t_N$
  - Note that there are M types of event types/labels and N events in the dataset

Proximal Graphical Event Model (PGEM)

- $\mathrm{PGEM} = \{G, W, \Lambda\}$; graph + set of (time) windows on each edge and conditional intensity parameters
- Assumption: The intensity of an event label (node) depends on whether or not its parents have happened at least once in their respective recent histories



Formally, denoting a node X's parents as $\boldsymbol{U}$:

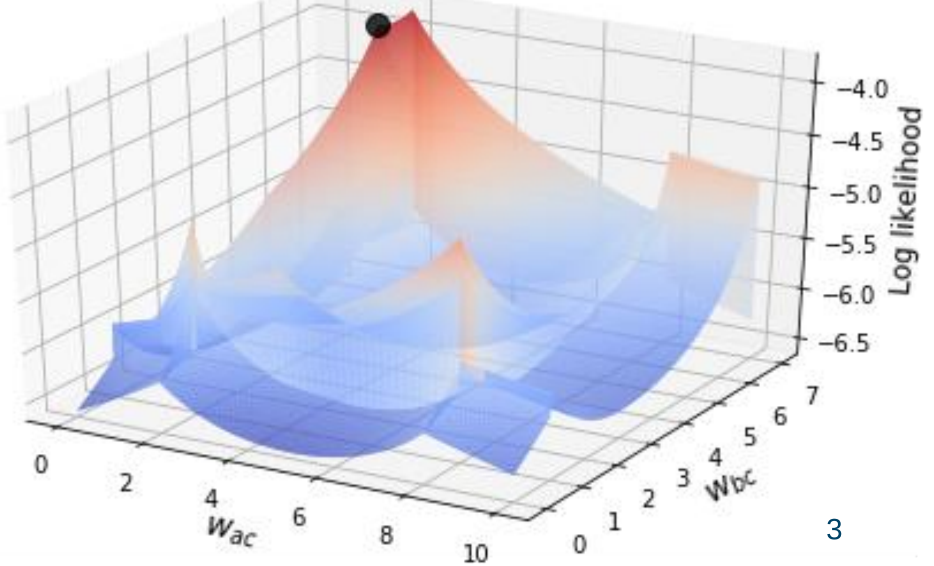
- $G = \{L, E\}$ where L is the event label set
- There is a window for every edge, $W = \{w_x : \forall X \in L\}$, where $w_x = \{w_{zx} : \forall Z \in \boldsymbol{U}\}$
- There is an intensity parameter for every node X and for every instantiation $\boldsymbol{u}$ of its parent occurrences, $\Lambda = \left\{\lambda_{x|\boldsymbol{u}}^{w_x} : \forall X \in L\right\}$

Parameter and Structure Learning

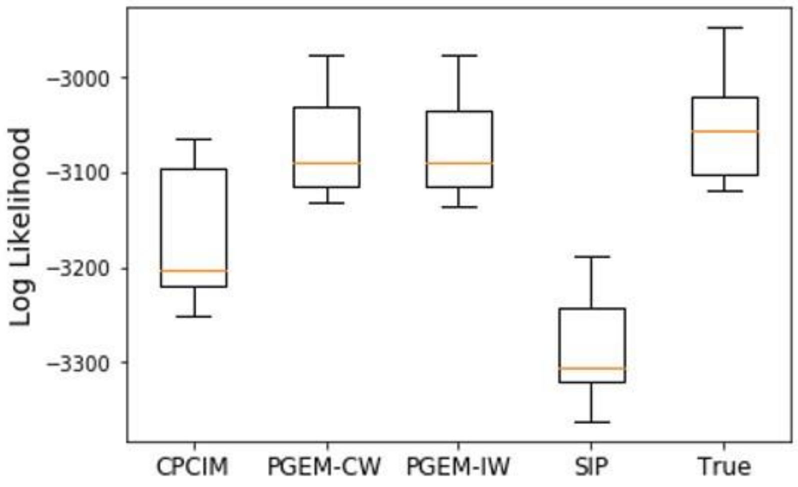
Learning problem: Given an event dataset $\boldsymbol{D}$, learn $\text{PGEM} = \{G, W, \Lambda\}$

- Log-likelihood: $\log\text{L}(D) = \sum_X \sum_{\mathbf{u}} \left(-\lambda_{x|\mathbf{u}} D(\mathbf{u}) + N(x;\mathbf{u}) \ln(\lambda_{x|\mathbf{u}})\right)$
  - $N(x;\boldsymbol{u})$: # of times X is observed and the condition $\boldsymbol{u}$ is true in the relevant windows
  - $D(\boldsymbol{u})$: duration over the entire time period where the condition $\boldsymbol{u}$ is true

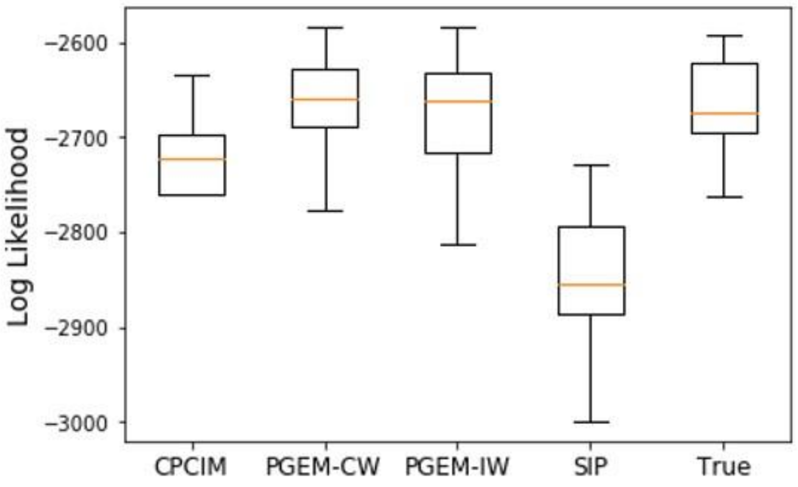
- For a given graph, finding the optimal (MLE) conditional intensities when given the windows is easy, but finding the optimal windows is hard!
- **Contribution 1: Analysis and proof that reduces the window search to a finite set that is algorithmically constructed.**
- **Contribution 2: A method to search over graph structures, with some theoretical results on efficient search and consistency justification**



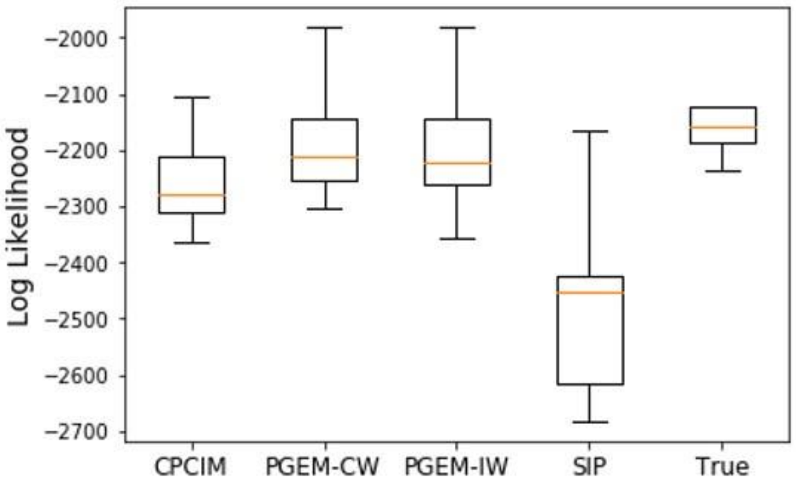
Results: Synthetic Datasets



a) PGEM #1



b) PGEM #2



c) PGEM #3

Wed Dec 5, 5:00 – 7:00 pm, Room 210 & 230 AB #6