Question of interest

Learning the causal structure of networks of multivariate time series in continuous time
Example 1: Information Diffusion

- Consider a network of users
- We observe a sequence of discrete events in continuous time: tweets, Facebook posts...
Example 1: Information Diffusion

• Consider a network of users

• We observe a sequence of discrete events in continuous time: tweets, Facebook posts...

• Questions of interest:
  Who influences whom?
  How does fake news spread?
Example 1: Information Diffusion

- Consider a **network of users**

- We observe a **sequence of discrete events in continuous time**: tweets, Facebook posts...

- Questions of interest:
  - Who **influences** whom?
  - How does **fake news spread**?
Example 2: Disease Dynamics

• Consider a **network of hospitals**

• We observe a **sequence of discrete events** in **continuous time**: interactions, infections, recoveries...

• Questions of interest:
  
  Who **infected** whom?
  
  How does the **disease spread**?
  
  How to **control** it?
Example 2: Disease Dynamics

• Consider a **network of hospitals**

• We observe a **sequence** of **discrete events** in **continuous time**: interactions, infections, recoveries...

• Questions of interest:
  - Who **infected** whom?
  - How does the **disease spread**?
  - How to **control** it?

---

**The New York Times**

Aug. 8, 2014

**U.N. Agency Calls Ebola Outbreak an International Health Emergency**
How do we usually solve it?
**Method:** Multivariate Hawkes Process (MHP)

- **Temporal Point Process**

- Widely used model to learn **causal structure between time series**

- Captures **mutually exciting** patterns of influence between dimensions

\[
\lambda_i(t|\mathcal{H}_t) = \mu_i + \sum_{j=1}^{d} \sum_{\tau \in \mathcal{H}_i^j} \kappa_{ij}(t - \tau)
\]
Method: Multivariate Hawkes Process (MHP)

• Temporal Point Process

• Widely used model to learn causal structure between time series

• Captures mutually exciting patterns of influence between dimensions

\[
\lambda_i(t|H_t) = \mu_i + \sum_{j=1}^{d} \sum_{\tau \in H_i^j} \kappa_{ij}(t - \tau)
\]

Exogenous intensity: constant, independent of the past

Endogenous intensity: due to excitation from past events, with excitation kernel

\[
\kappa_{ij}(t) = \alpha_{ij} e^{-\beta t} \mathbb{1}_{\{t > 0\}}
\]
**Method:** Multivariate Hawkes Process (MHP)

- Prior work assume **perfect traces without noise**
- What if the observed stream of events is subject to a **random and unknown time shift**?
How to learn MHPs under noisy observations?
Multivariate Hawkes Process under Synchronization Noise

- What events have systematic measurement errors?
Multivariate Hawkes Process under Synchronization Noise

- What it events have **systematic measurement errors**?
Events can enter the observation window... or escape it.
What events have systematic measurement errors?

Edges learnt by maximum likelihood estimation can be significantly affected by even small delays.
New approach DESYNC-MHP

- **Idea:**
  - Consider the noise as parameters
  - Maximize the joint log-likelihood over both MHP parameters and noise
New approach DESYNC-MHP

• **Idea:**
  • Consider the noise as parameters
  • Maximize the joint log-likelihood over both MHP parameters and noise

• **Challenges:** resulting objective is
  • Non-smooth
  • Non-convex
New approach DESYNC-MHP

• **Idea:**
  • Consider the noise as parameters
  • Maximize the joint log-likelihood over both MHP parameters and noise

• **Challenges:** resulting objective is
  • Non-smooth
  • Non-convex

• **Solution:**
  • Approximate the objective with a smooth approximation
  • Use SGD to escape local minima
Experimental Results

![Graph showing experimental results]

- **Average accuracy (± std)** vs **Noise variance $\sigma^2$**
- **Classic MLE**
- **DESYNC-MHP MLE**
- **$1/\beta = 1.0$**
Learning Hawkes Processes Under Synchronization Noise

William Trouleau      Jalal Etesami      Negar Kiyavash
Matthias Grossglauser  Patrick Thiran

EPFL      Georgia Tech      BOSCH

Come check out our poster tonight!