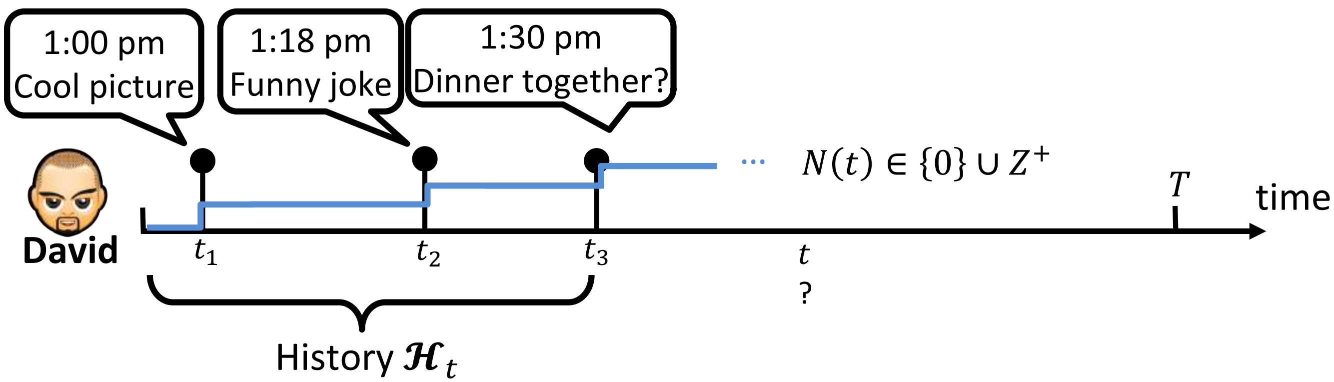
Learning Temporal Point Processes via Reinforcement Learning

Shuang Li[1], Shuai Xiao[2], Shixiang Zhu[1], Nan Du[3], Yao Xie[1], Le Song[1,2]

[1]Georgia Institute of Technology
[2]Ant Financial
[3]Google Brain

Motivation



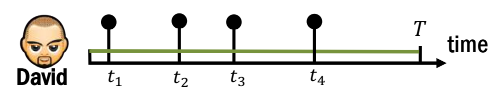
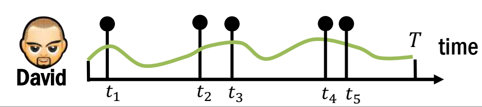
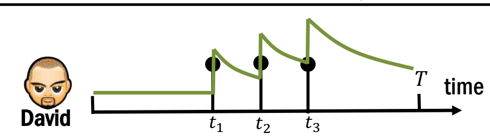
- **Event data** : tweets/retweets, crime events, earthquakes, patient visits to hospital, finance transactions, ….
- Learn **temporal pattern** of event data.
  - Event time is random
  - Complex dependency structure

Point Process Model

- Intensity function

$$\lambda(t|\mathcal{H}_t)dt = \mathbb{E}[N[t, t+dt)]|\mathcal{H}_t]$$

where $N[t, t+dt)$ is the number of events falling in the set $[t, t+dt)$.

Point Process	$\boldsymbol{\lambda_\theta(t|\mathcal{H}_t)}$	**Temporal Pattern**
Poisson	constant	
Inhomogeneous Poisson	$\lambda_\theta(t)$	
Hawkes	$\mu + \alpha \sum_{t_i \in \mathcal{H}_t} \exp(-|t - t_i|)$	

Traditional Maximum-Likelihood Framework

- Model conditional intensity

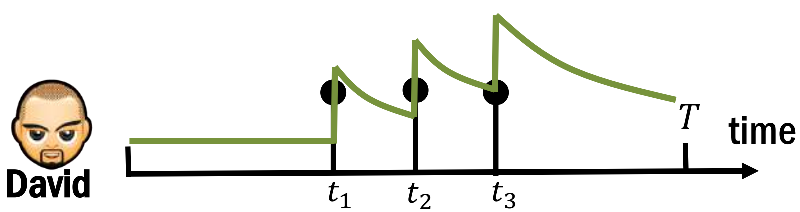
$$\lambda_\theta(t|\mathcal{H}_t)$$

as a parametric/non-parametric form.

Model-misspecification!

- Learn model by maximizing likelihood

$$P(t_1, t_2, \dots, t_n) = \exp\left(-\int_{(0,T)} \lambda_\theta(t|\mathcal{H}_t)\, dt\right) \prod_i \lambda_\theta(t_i|\mathcal{H}_t)$$



New Reinforcement Learning Framework

- Learn policy

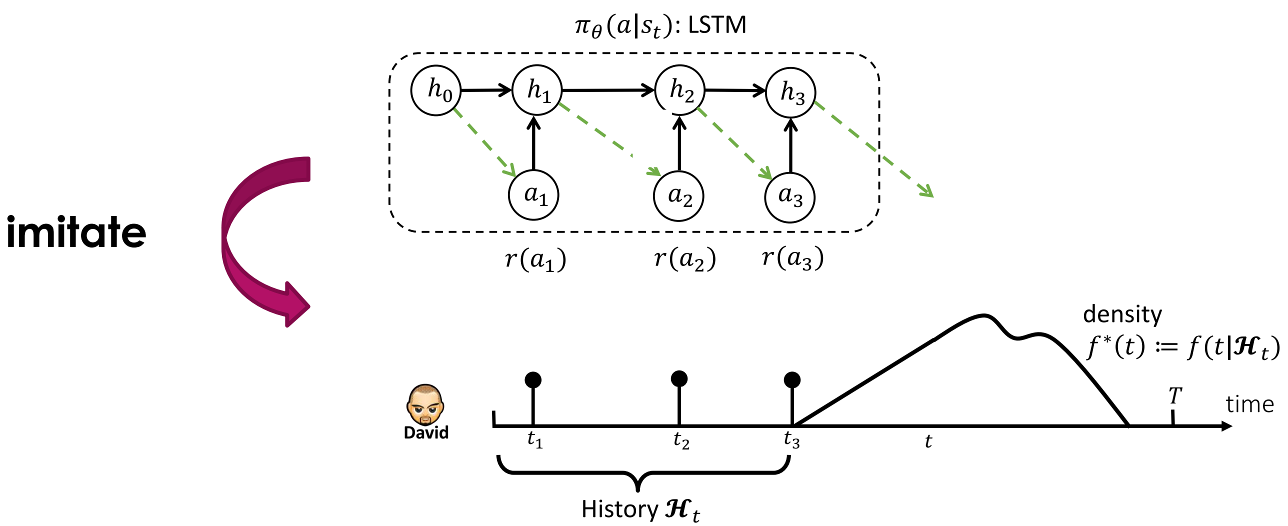
$$\pi_\theta(a|s_t) = p(t_i \mid t_{i-1}, \dots, t_1)$$

where $a \in \mathrm{R}^+$ is the next event time, to maximize cumulative reward .

- Learn reward

$$r(a)$$

to guide policy to imitate observed event data (expert).



Optimal Reward

- Inverse Reinforcement Learning

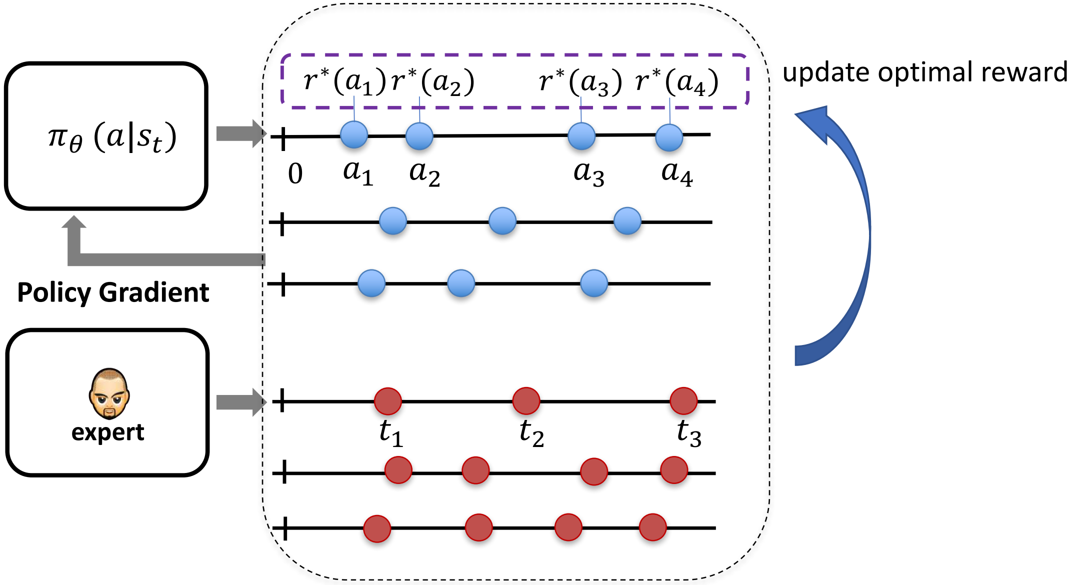
$$r^* = \max_{r \in \mathcal{F}} \left(\mathbb{E}_{\text{expert}} \left[\sum_i r(t_i) \right] - \max_{\pi_\theta} \mathbb{E}_{\pi_\theta} \left[\sum_i r(a_i) \right] \right)$$

- Choose $r \in \mathcal{F}$ be unit ball in Reproducing Kernel Hilbert Space (RKHS). We obtain analytical optimal reward
- Given L expert trajectories, M generated trajectories by policy

$$\hat{r}^*(a) \propto \underbrace{\frac{1}{L} \sum_{l=1}^{L} \sum_{i=1} k\left(t_i^{(l)}, a\right)}_{\text{mean embedding of expert intensity function}} - \underbrace{\frac{1}{M} \sum_{m=1}^{M} \sum_{i=1} k\left(a_i^{(m)}, a\right)}_{\text{mean embedding of policy intensity function}}$$

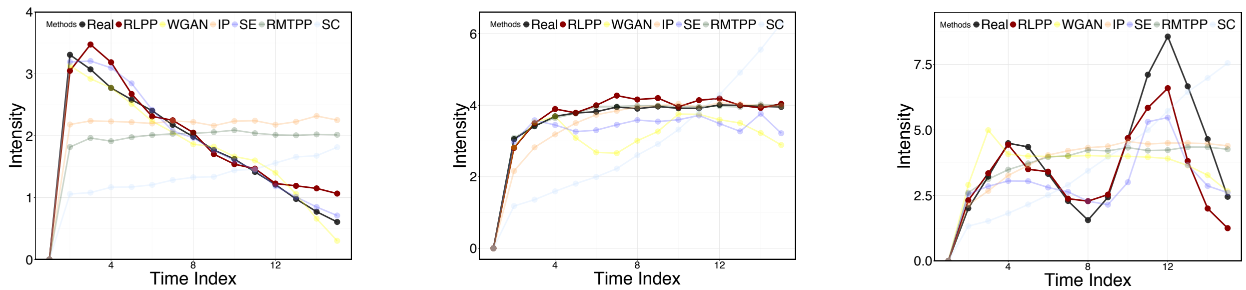
where $k(t, t')$ is a universal RKHS kernel.

Modeling Framework



Numerical Results

- Our method: **RLPP**
- Baselines:
  - State-of-the-art methods: RMTPP(Du et al. 2016 KDD) WGANTPP (Xiao et al. 2017 NIPS)
  - Parametric baselines: Inhomogeneous Poisson (IP), Hawkes (SE), Self-correcting (SC)
- Comparison of learned empirical intensity



- Comparison of runtime

Method	**RLPP**	WGANTPP	RMTPP	SE	SC	IP
Time	80 m	1560m	60m	2m	2m	2m
Ratio	40x	780x	30x	1x	1x	1x

Poster

- Tue Dec 4th 05:00 -- 07:00 PM
- @ Room 210 & 230 AB #**124**