One-Pass Algorithms for MAP Inference of Nonsymmetric Determinantal Point Processes

Aravind Reddy

Northwestern

Joint work with Ryan A. Rossi, Zhao Song, Anup Rao, Tung Mai, Nedim Lipka, Gang Wu, Eunyee Koh (Adobe), Nesreen K. Ahmed (Intel)

Northwestern | McCORMICK SCHOOL OF ENGINEERING

Computer Science

ICML 2022

International Conference On Machine Learning

Adobe

Determinantal Point Processes (DPP)

- Probability distribution on all subsets of a ground set of items [n] = {1, 2, …, n} characterized by kernel matrix $\boldsymbol{L} \in \mathbb{R}^{n \times n}$ such that

$$\Pr(S) \propto \det(L_S)$$

- For example, in E-Commerce applications, the subsets $\boldsymbol{S}$ are baskets (carts) bought by users.

- DPPs have traditionally been used to encourage *diversity* in recommender systems.

Nonsymmetric DPP

- Discrete DPPs were introduced in ML literature in 2010 and work upto 2017 operated under the constraint that $\boldsymbol{L}$ needs to be symmetric.
- Gartrell et.al (2021) *low-rank* NDPP decomposition:

$$\boldsymbol{L} = \boldsymbol{V}\boldsymbol{V}^{\boldsymbol{T}} + \boldsymbol{B}\boldsymbol{C}\boldsymbol{B}^{\boldsymbol{T}}$$

where $\boldsymbol{L} \in \mathbb{R}^{\mathrm{n}\times\mathrm{n}}, \boldsymbol{V} \in \mathbb{R}^{\mathrm{n}\times\mathrm{d}}, \boldsymbol{B} \in \mathbb{R}^{\mathrm{n}\times\mathrm{d}}, \boldsymbol{C} \in \mathbb{R}^{\mathrm{d}\times\mathrm{d}}$ with $d \ll n$ and $\boldsymbol{C}$ is a skew-symmetric matrix.

Our Problem Setup

- Stream: $(\boldsymbol{v}_1, \boldsymbol{b}_1), (\boldsymbol{v}_2, \boldsymbol{b}_2), \ldots, (\boldsymbol{v}_{\mathrm{n}}, \boldsymbol{b}_{\mathrm{n}})$ where $\boldsymbol{v}_{\mathrm{t}}, \boldsymbol{b}_{\mathrm{t}} \in \mathbb{R}^{\mathrm{d}}$ and $d \ll n$.
- At every time step t, want to output subset S to

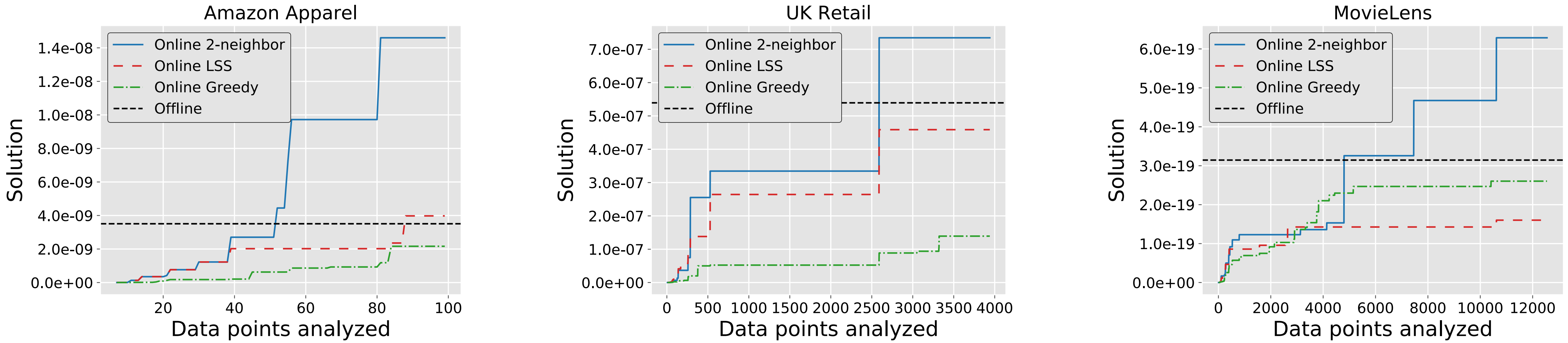
$$\max f(S) = \det(V_S^T V_S + B_S^T C B_S)$$

- $\Pr(S) \propto f(S)$ according to Non-symmetric Determinantal Point Process (NDPP).
- Maximum a Posteriori Inference (MAP Inference).

Main Contributions

- First formulation of the streaming and online MAP Inference problem for Non-symmetric Determinantal Point Processes (NDPPs).
- Design new one-pass algorithms for these problems and show that empirically they perform comparably to (or even better than) the offline greedy algorithm while using substantially lower memory.
- Hard instance for one-pass MAP inference of NDPPs in the online setting.

Experiments



- **Key findings**:
  - Comparable (sometimes even better!!) than the offline greedy algorithm while:
    - Taking a **single** pass over the data.
    - Maintaining a valid solution at each time step.
    - Using a ***fraction*** of the memory (when compared to offline algorithms)
    - Fast update time after seeing any new point
  - Tradeoff between Space and Solution quality.

Future Directions

- Proving approximation factor bounds for our online algorithms under data-assumptions (beyond worst-case analysis).

- Designing new streaming and online algorithms using ideas from algorithms for submodular function maximization.

- Algorithms with constant number of passes (more than 1 but less than k)

Thank you!

Poster today from 6.30 - 8.30 pm
at Hall E #1124