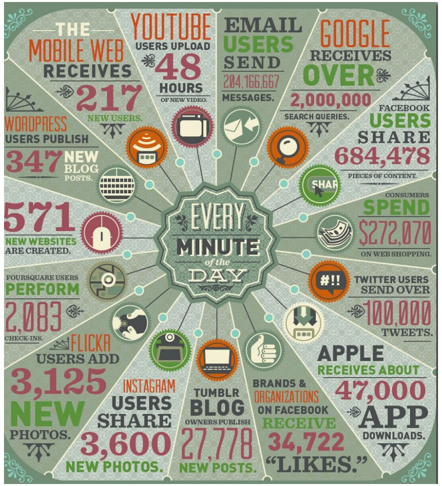
Topics are not Marks : Modeling Text-based Cascades using Multi-network Hawkes Process

Srikanta Bedathur (IIT Delhi), Indrajit Bhattacharya (TCS Research), Jayesh Choudhari, Anirban Dasgupta (IIT Gandhinagar)

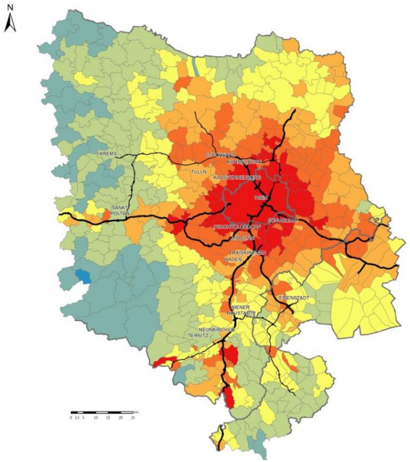
Motivation

Network (hidden) + Time-series of Events

Online Activity

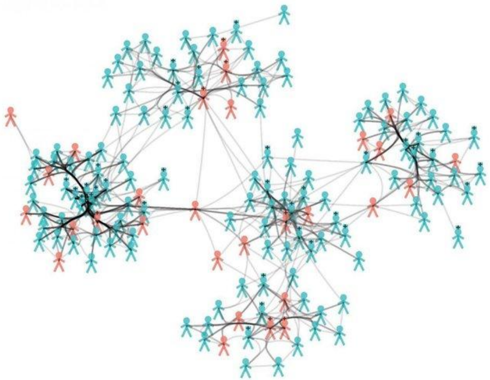


Mobility Dynamics

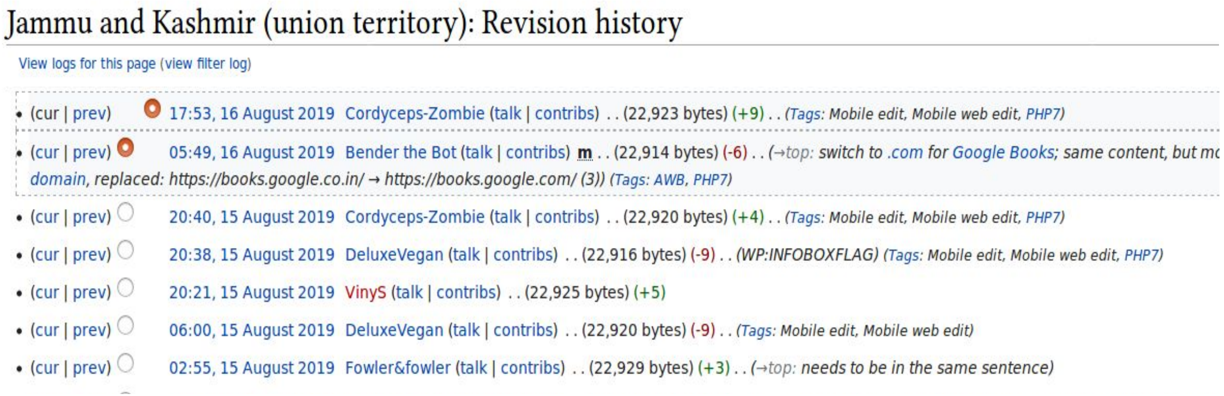


Financial Trading

Disease Dynamics

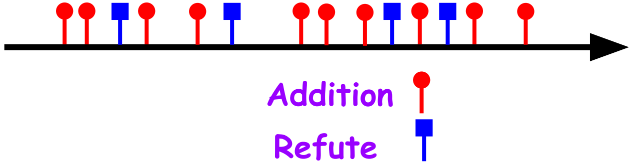


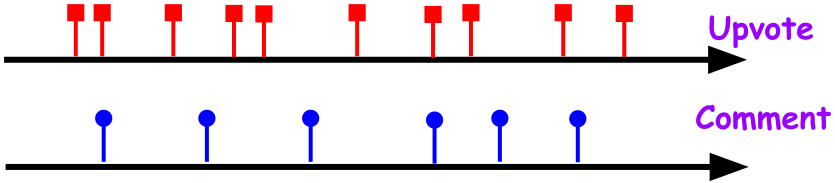
Network (hidden) + Time-series of Events


Jammu and Kashmir (union territory): Revision history

View logs for this page (view filter log)

- (cur | prev) 17:53, 16 August 2019 Cordyceps-Zombie (talk | contribs) . . (22,923 bytes) (+9) . . (*Tags: Mobile edit, Mobile web edit, PHP7*)
- (cur | prev) 05:49, 16 August 2019 Bender the Bot (talk | contribs) **m** . . (22,914 bytes) (-6) . . (*→top: switch to .com for Google Books; same content, but mo domain, replaced: https://books.google.co.in/ → https://books.google.com/ (3)*) (*Tags: AWB, PHP7*)
- (cur | prev) 20:40, 15 August 2019 Cordyceps-Zombie (talk | contribs) . . (22,920 bytes) (+4) . . (*Tags: Mobile edit, Mobile web edit, PHP7*)
- (cur | prev) 20:38, 15 August 2019 DeluxeVegan (talk | contribs) . . (22,916 bytes) (-9) . . (*WP:INFOBOXFLAG*) (*Tags: Mobile edit, Mobile web edit, PHP7*)
- (cur | prev) 20:21, 15 August 2019 VinyS (talk | contribs) . . (22,925 bytes) (+5)
- (cur | prev) 06:00, 15 August 2019 DeluxeVegan (talk | contribs) . . (22,920 bytes) (-9) . . (*Tags: Mobile edit, Mobile web edit*)
- (cur | prev) 02:55, 15 August 2019 Fowler&fowler (talk | contribs) . . (22,929 bytes) (+3) . . (*→top: needs to be in the same sentence*)






What should I do as a computer science undergraduate?

Jeff Erickson, CS professor, University of Illinois at Urbana-Champaign

Answered Tue · Upvoted by Harsh Suryavanshi, M-Tech Computer Science & Information Security, Indraprastha Institute of Information Technology, D...

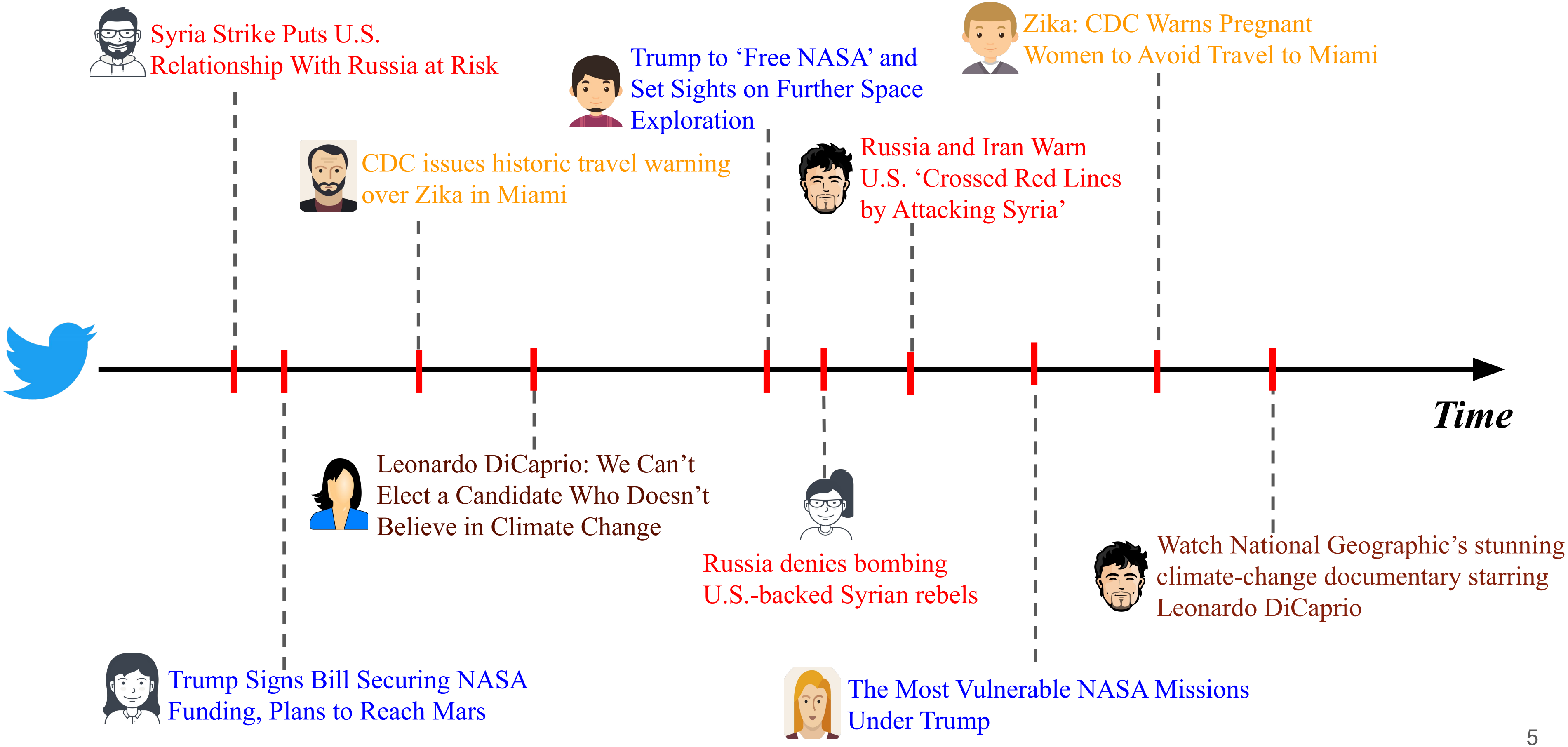
1. Eat. Sleep. Bathe. Go outside. Exercise. Make friends. Have fun.

2. Read. Write. Ask. Listen. Learn. Practice. Try. Fail. Improve. Repeat.

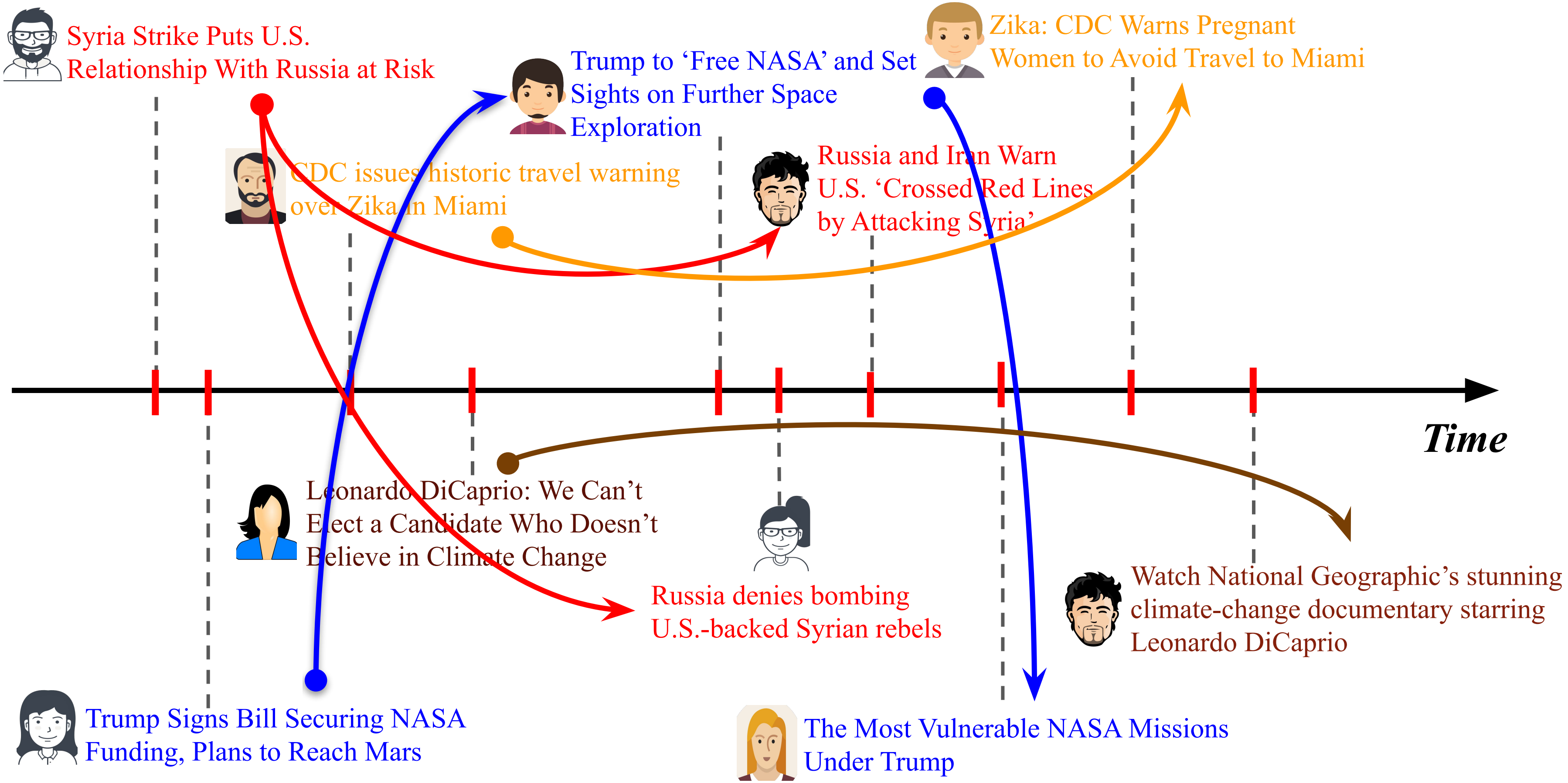
949 views · View Upvoters · View Sharers · Answer requested by Naballa Blesson

Upvote · 95 Share · 1

Twitter: Network + Time-series of Tweets

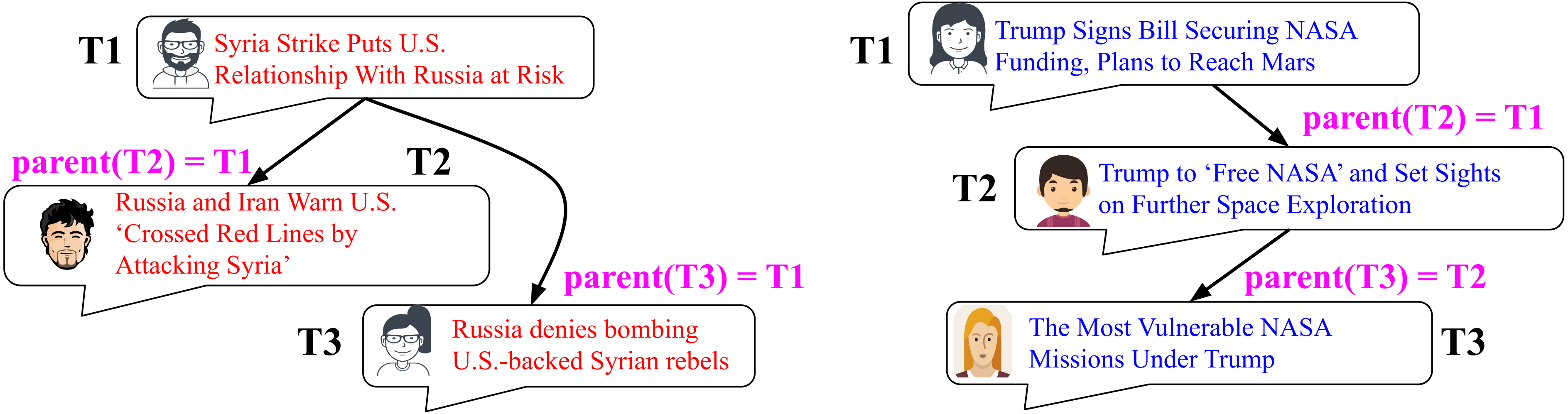


Mixture of Conversations (Twitter) : Topic interactions



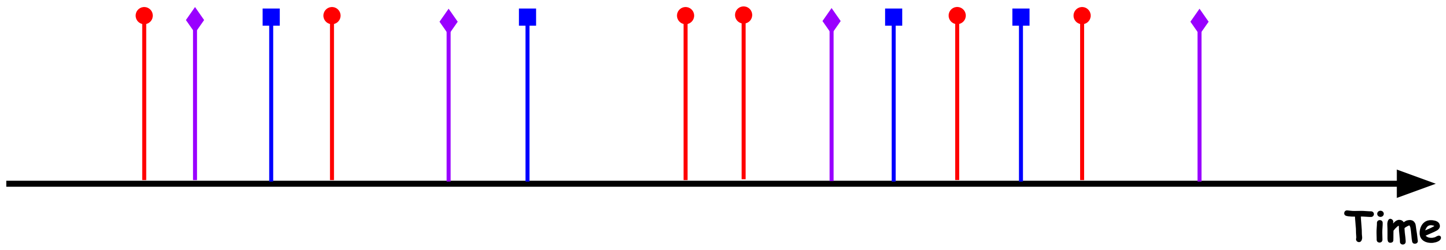
Topics are not Marks: Modeling Cascades

Cascades (Separate Conversations)



Just separate this conversations out!!!

Marked Temporal Point Process (MTPP)

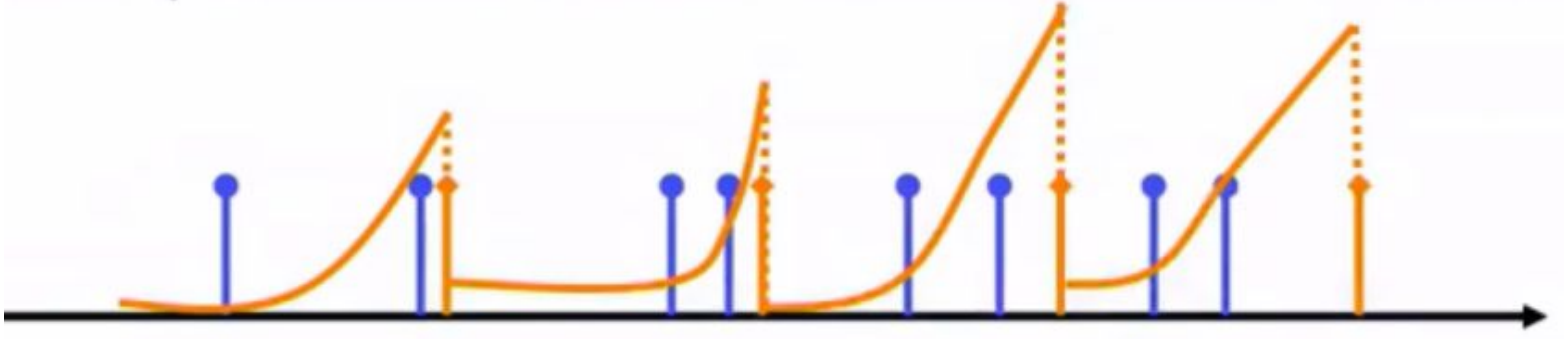


$$\mathcal{H} = \{e_0 = (t_0, \eta_0), e_1 = (t_1, \eta_1), \ldots, e_n = (t_n, \eta_n)\}$$

$$t_i \in \mathbb{R}, \eta_i \in \mathbb{Z}$$

- Sequence of events of type η_i at times t_i
  - Continuous Time
  - Discrete, continuous (or mixed) marks (could be vector of marks)

Marked Temporal Point Process (MTPP)



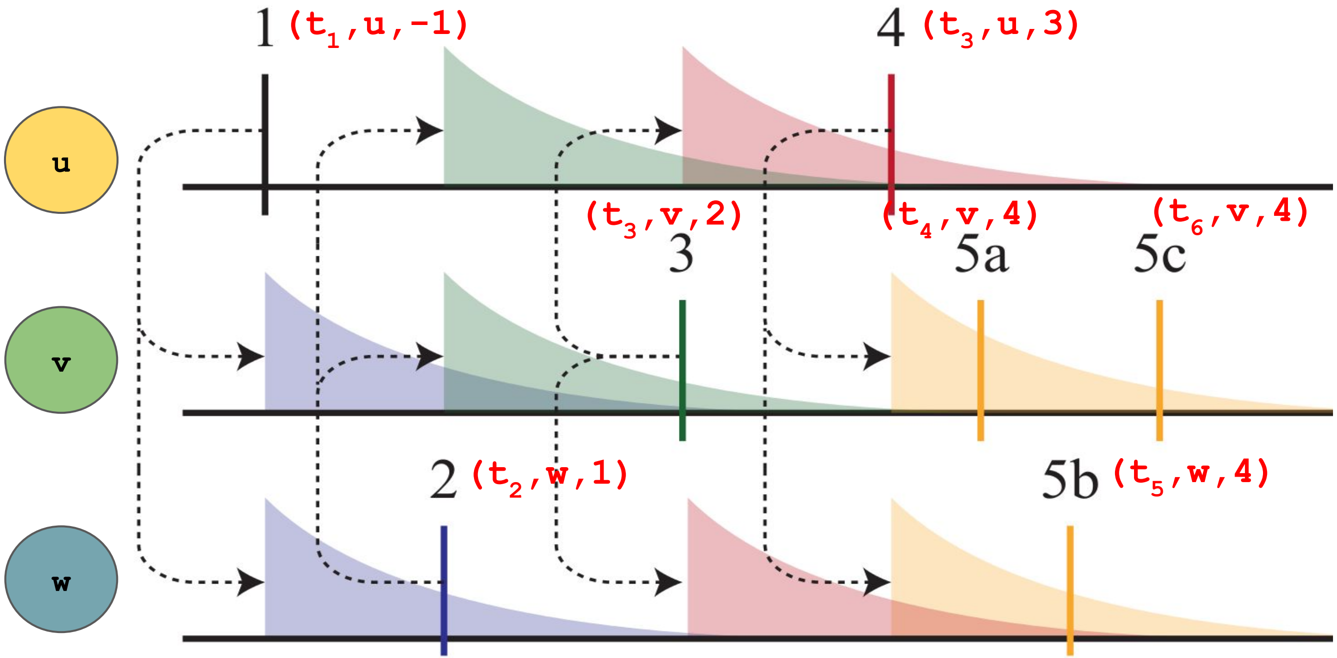
Time-stamps characterized by an *intensity* function:

$$\lambda(t) := \mathbb{P}\{\text{event in } [t, t+dt) | \mathcal{H}_t\}$$

Image Source: Deep Reinforcement Learning of Marked Temporal Point Processes, 2019, Abir et al.

Multivariate Hawkes Process (HMHP, HTM, NetHawkes)

$$\lambda_v(t) = \mu_v(t) + \sum_{n=1}^{|\mathcal{H}_{t-}|} h_{c_n,v}(t - t_n)$$



For each event, topics are sampled later independently of the time-stamps

Likelihood MTPP

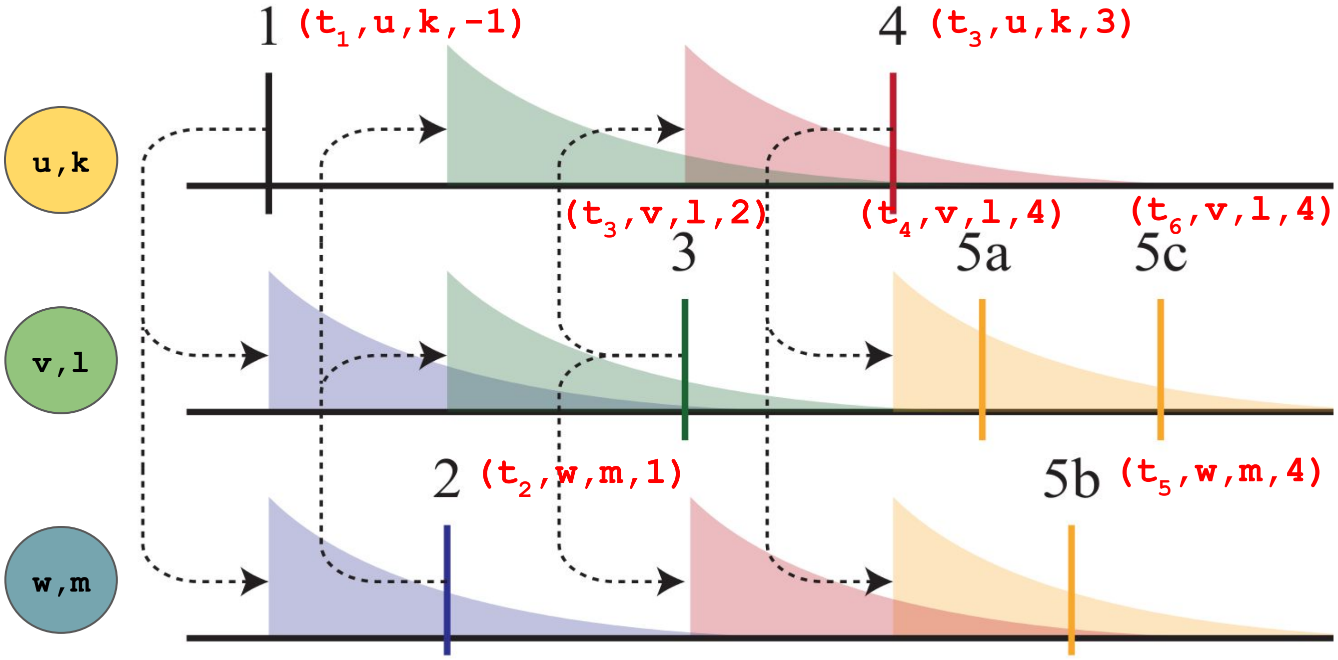
$$\mathbb{P}\left(\mathcal{H}_T\right) := \left(\prod_{e_i \in \mathcal{H}_T} \overbrace{\lambda_{v_i}(t_i)}^{\text{Prob. of an action at } t_i} \underbrace{m^*\left(\eta_i\right)}_{\text{Prob. of mark } \eta_i} \right) \prod_{v \in V} \overbrace{\exp\left(\int_0^T \lambda_v(\tau) d\tau \right)}^{\text{Prob. of no actions at } t \in [0,T] \setminus \{t_i\}}$$

Note: The time-stamps t_i and the marks η_i are modeled independently.

Unified Model

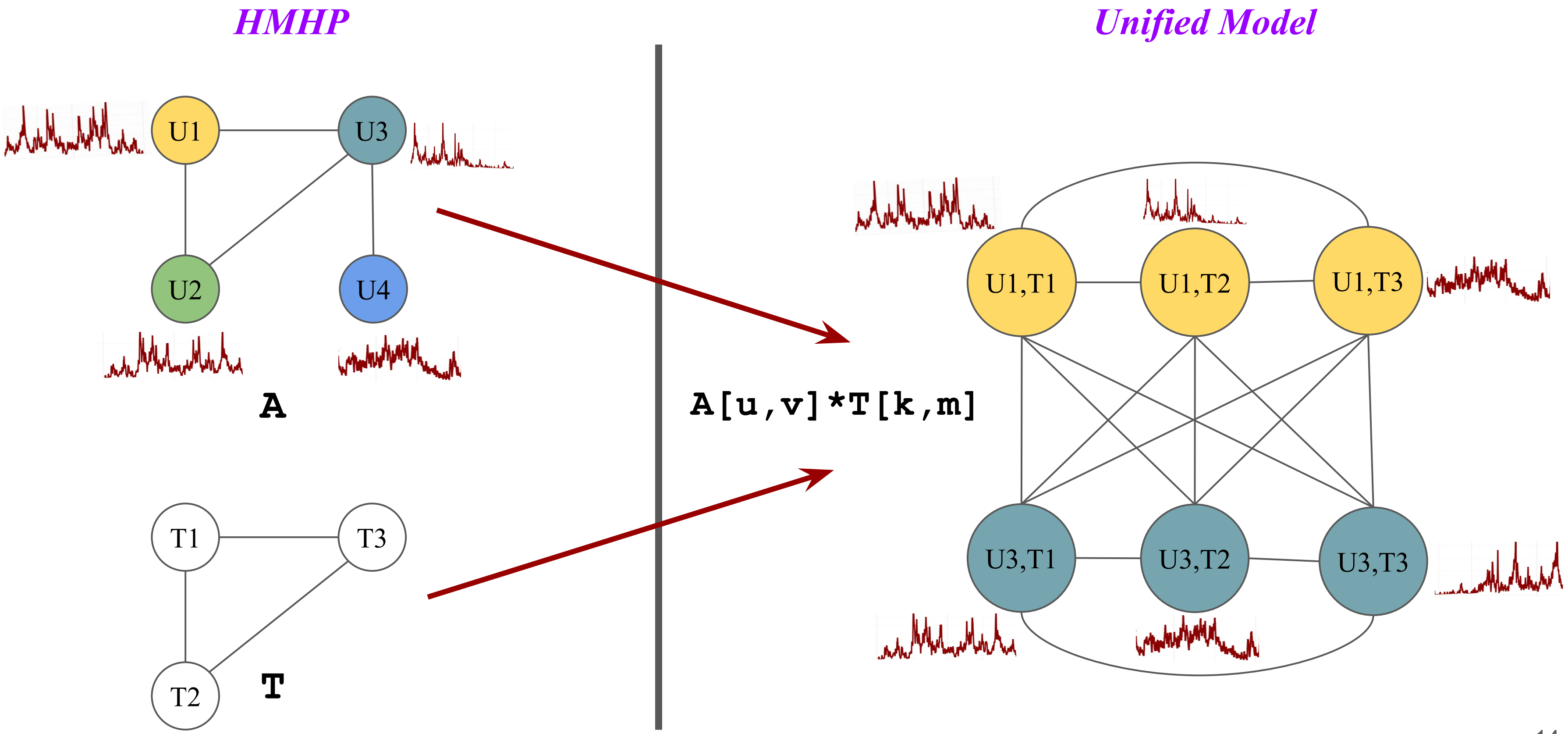
Unified Marked Multivariate Hawkes Process

$$\lambda_{\mathbb{v}}(t) = \mu_{\mathbb{v}}(t) + \sum_{n=1}^{|\mathcal{H}_{t-}|} h_{\mathbb{c}_{\mathbb{n}},\mathbb{v}}(t - t_n)$$



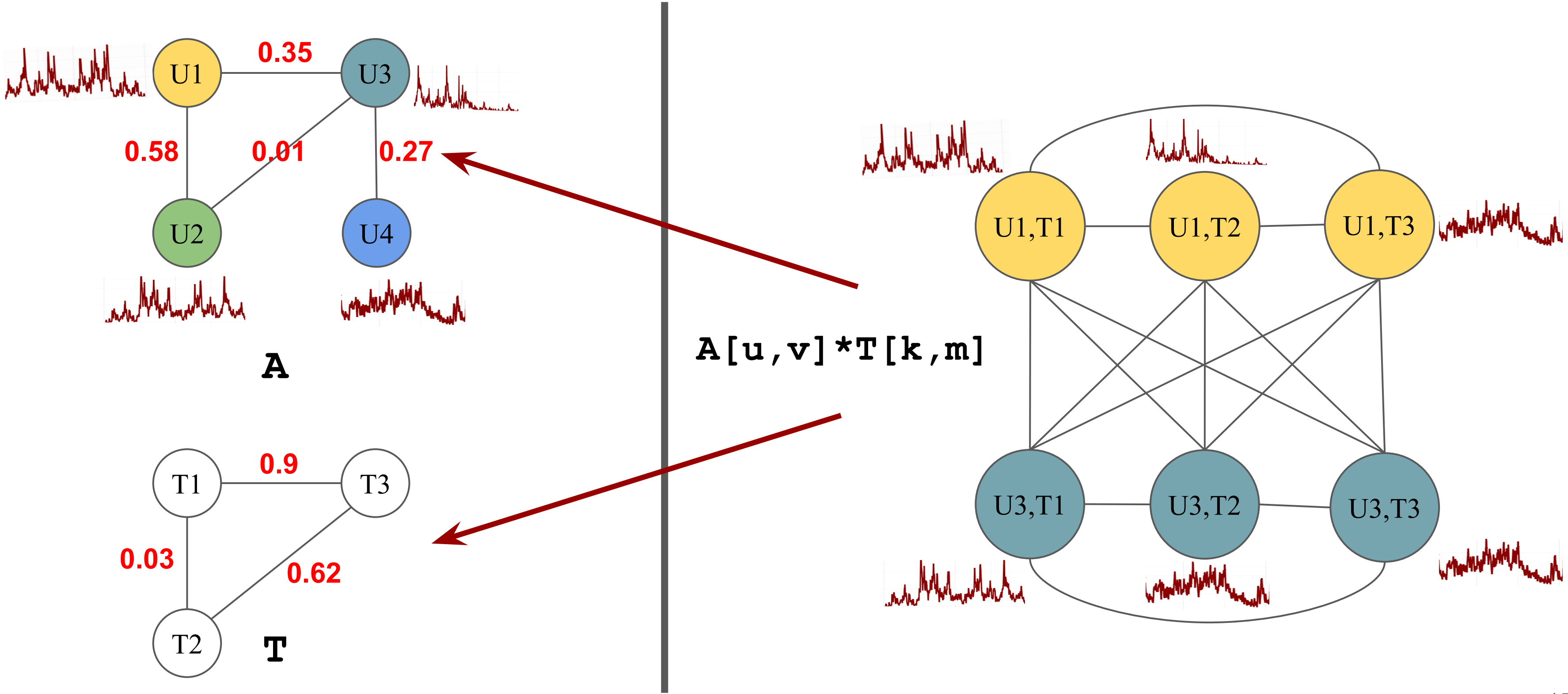
For each event, topics come along with the time-stamps

HMHP Model v/s Unified Model



HMHP Model v/s Unified Model

HMHP ***Unified Model***



Likelihood Unified MTPP

$$\mathbb{P}\left(\mathcal{H}_T\right) := \left(\prod_{e_i \in \mathcal{H}_T} \overbrace{\lambda_{\mathbb{v}_i}(t_i)}^{\text{Prob. of an action at } t_i \text{ with mark } \eta_i} \right) \prod_{\mathbb{v} \in \mathcal{V}} \overbrace{\exp\left(-\int_0^T \lambda_{\mathbb{v}_i}(\tau) d\tau \right)}^{\text{Prob. of no actions at } t \in [0,T] \setminus \{t_i\}}$$

$$\mathbb{v}_i = (v_i, \eta_i)$$

Recall:

$$\mathbb{P}\left(\mathcal{H}_T\right) := \left(\prod_{e_i \in \mathcal{H}_T} \overbrace{\lambda_{v_i}(t_i)}^{\text{Prob. of an action at } t_i} \underbrace{m^*(\eta_i)}_{\text{Prob. of mark } \eta_i} \right) \prod_{v \in V} \overbrace{\exp\left(-\int_0^T \lambda_v(\tau) d\tau \right)}^{\text{Prob. of no actions at } t \in [0,T] \setminus \{t_i\}}$$

HMHP Model (HTM, NetHawkes) v/s Unified Model

HMHP (HTM, NetHawkes)		**Unified Model**
(each user node)	Intensity Function	(a user-topic pair)

Intensity Function

HMHP (each user node): v

$$\lambda_v(t) = \mu_v(t) + \sum_{n=1}^{|\mathcal{H}_{t-}|} h_{c_n,v}(t - t_n)$$

Unified Model (a user-topic pair): $\mathbb{v} =$ v,k

$$\lambda_{\mathbb{v}}(t) = \mu_{\mathbb{v}}(t) + \sum_{n=1}^{|\mathcal{H}_{t-}|} h_{\mathbb{c}_n,\mathbb{v}}(t - t_n)$$

Base Rate

v

$$\mu_v(t) = \mu_v(t)$$

$\mathbb{v} =$ v,k

$$\mu_{\mathbb{v}}(t) = \mu_v(t)\mu_k(t)$$

Impulse Response

$u \rightarrow v$

$$h_{u,v}(\Delta t) = W_{u,v} f(\Delta t)$$

$u,k \rightarrow v,k'$

$$h_{\mathbb{u},\mathbb{v}}(\Delta t) = W_{u,v} \mathcal{T}_{k,k'} f(\Delta t)$$

Unified Model Inference

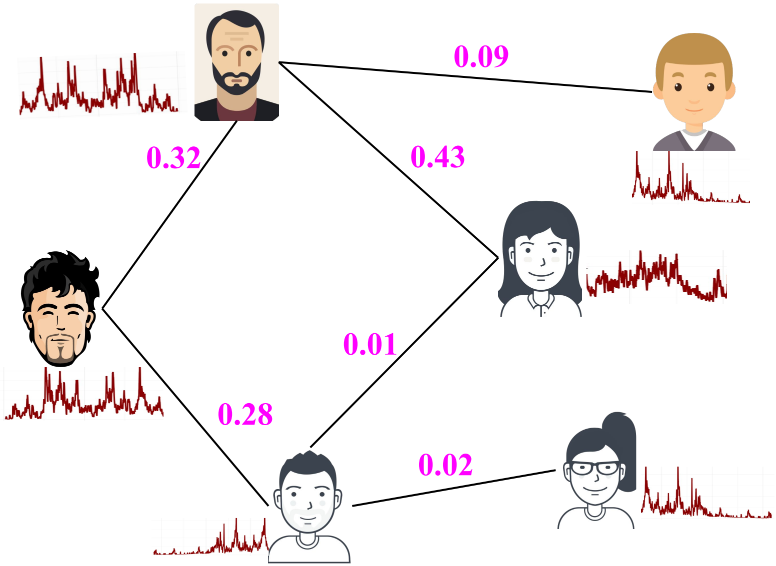
- User-user weights and topic-topic weights are now coupled → lack of conjugacy property, implying that both the set of weights need to be actually sampled

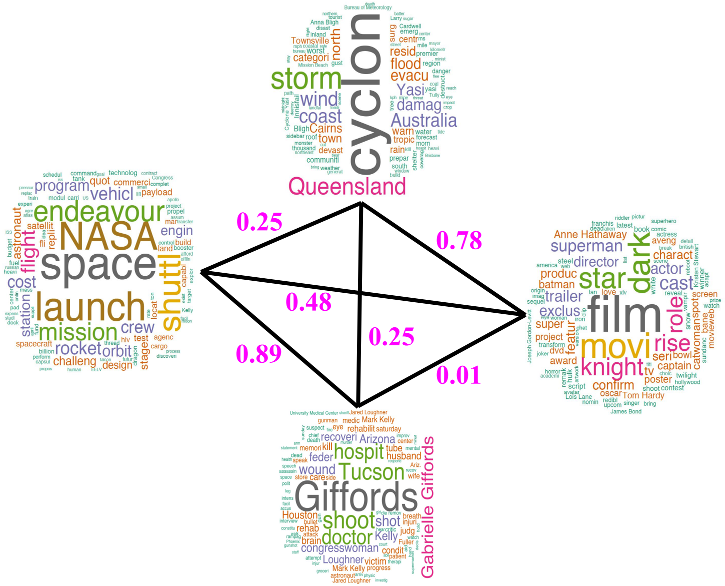
- Interestingly, Gibbs sampling is still efficient as they are both follow conditional Gamma distributions

$$P(W_{u,v}|\boldsymbol{\mathcal{T}}, \boldsymbol{z}) \propto P(W_{u,v}|\alpha_1, \beta_1) \left(W_{u,v}^{N_{u,v}} exp\left(-\beta W_{u,v}\right) C\right)$$

- This coupling in fact, allows the flow of evidence between two user-topic tuple pairs, this is expected to be useful for user pairs or topic pairs between which the data is scarcer

Inference Tasks





- User Temporal Dynamics
- Preferred topics of each user
- Network Strengths (user-user influence)

- Topics
- Topical Interactions

HMHP Model (HTM, NetHawkes) v/s Unified Model

HMHP (HTM, NetHawkes) | **Unified Model**

Base Rate Inference

$$\mu_v = \frac{N_v^{(spon)}}{T}$$

Note: *There is no base rate associated with the topics*

$$\mu_v = \frac{N_v^{(spon)}}{T \sum_{k \in K} \mu_k} \qquad \mu_k = \frac{N_k^{(spon)}}{T \sum_{v \in V} \mu_v}$$

Influence Inference

$$W_{u,v} \sim Gamma\left(N_{u,v} + \alpha_1, N_u + \beta_1\right)$$

Note: *Topic-Topic interaction is integrated out in HMHP because of conjugacy*

$$W_{u,v} \sim Gamma\left(N_{u,v} + \alpha_1, \sum_k \left(N_{u,k} \sum_{k'} \boldsymbol{\mathcal{T}}_{k,k'}\right) + \beta_1\right)$$

$$\boldsymbol{\mathcal{T}}_{k,k'} \sim Gamma\left(N_{k,k'} + \alpha_1, \sum_u \left(N_{u,k} \sum_v W_{u,v}\right) + \beta_1\right)$$

Experiments & Results: Simulated data

User graph:
- Top 50 authors from High energy Physics
- Edge weights generated using a Gamma

Topic graph
- Erdos-Renyi graph with 10 nodes and edge prob = 0.5
- Edge weights generated using Gamma (scale parameter proportional to distance in original graph

Topic-word distributions: Dirichlet

Data generated using the Unified model with above parameters

Tasks:
- Reconstructing the parameters
- Generalization, i.e. likelihood on test data

Experiments & Results: Simulated data

Topic evaluation results

	MAE	Med. AE	Std. Dev.
HMHP	0.009	0.0088	0.0131
Uni-1G	0.009	0.0088	0.0123
Uni-2G	0.009	0.0088	0.0122
Uni-DG	0.009	0.0088	0.0124

Parent identification results

	Acc.	R @1	R @3	R @5
HMHP	0.375	0.417	0.651	0.754
Uni-1G	0.391	0.430	0.668	0.769
Uni-2G	0.391	0.431	0.668	0.770
Uni-DG	0.392	0.431	0.668	0.769

Summary and Future Work

Proposed a new model for modeling information cascades on networks, general enough to incorporate number of previous models

Critical step is joint modeling of topical and temporal information-- topics treated at par with users

Interesting use cases on the way-- when we decide what is observable and what is not !

Joint modeling allows information to flow across user-topic pairs

Initial experiments on simulated data show better performance when compared to related models

Full version under review

Thank You