Information and Opinion Dynamics in Online Social Networks

Presented By: **Niloy Ganguly, Department of Computer Science**

iitkgpcnerg

@cnerg

CNeRG IIT Kharagpur

www.cnerg.org

Collaborators

Abir De

Bidisha Samanta

Sourangshu Bhattacharya

Isabel Valera

Manuel Gomez Rodriguez

Activities over Social Network and Social Media

News spread in Twitter

Product reviews and sales in Amazon

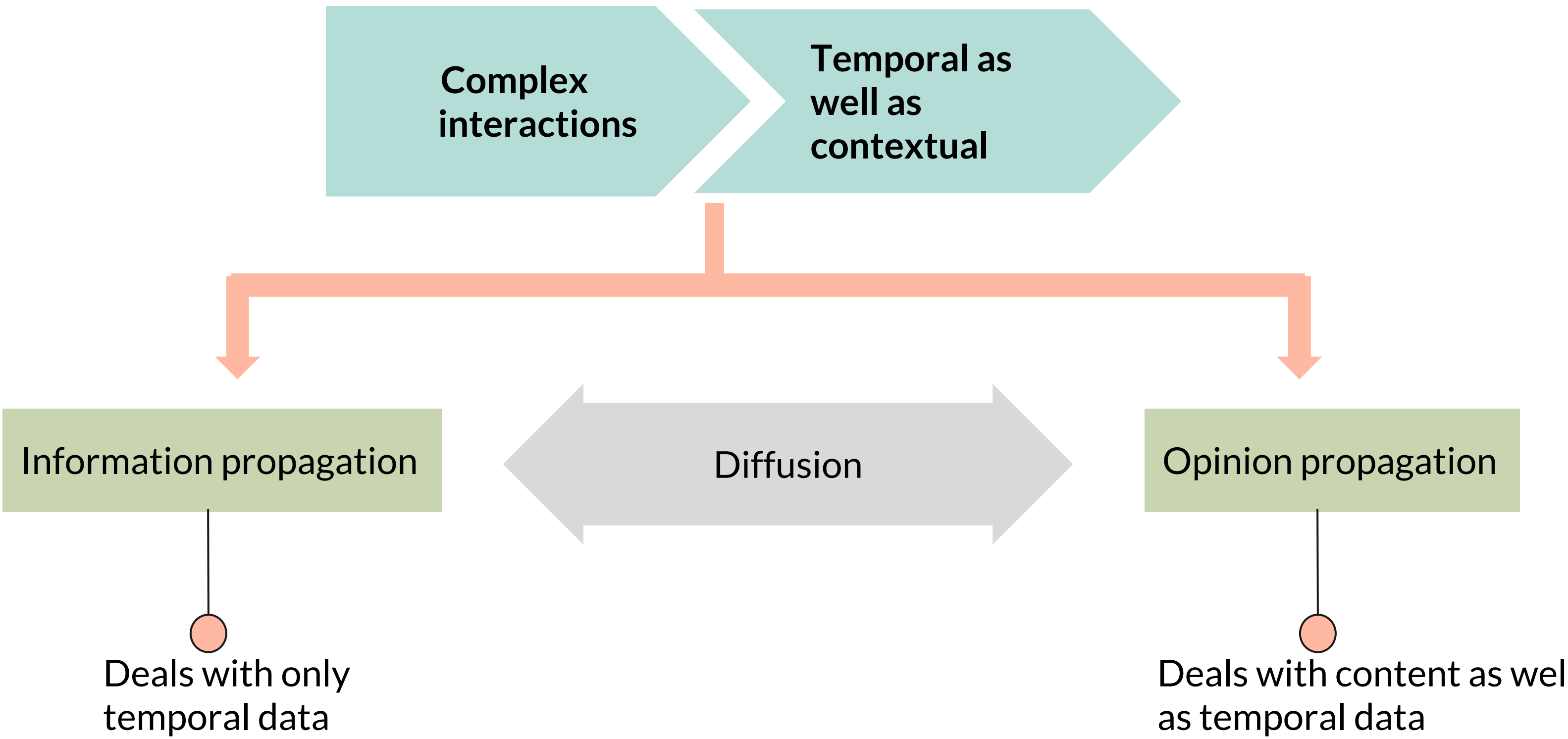
A user gains recognition in Quora

Video becomes viral in Youtube

Posts on several events becomes viral

Events are observations of complex dynamic process

Events are observations of complex dynamic process



Challenges to understand the diffusion process

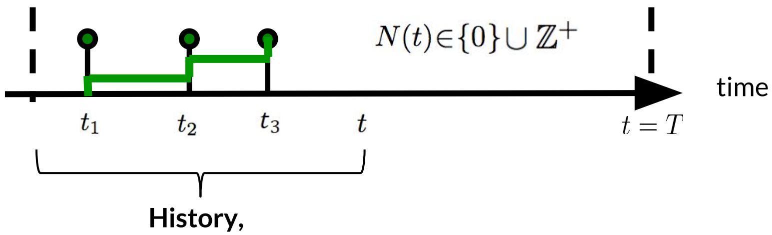
Understanding which pieces of information are getting through and which are ignored

Accurate Modelling

Accurate Forecasting

Mathematical device to capture the diffusion process

Temporal Point Process



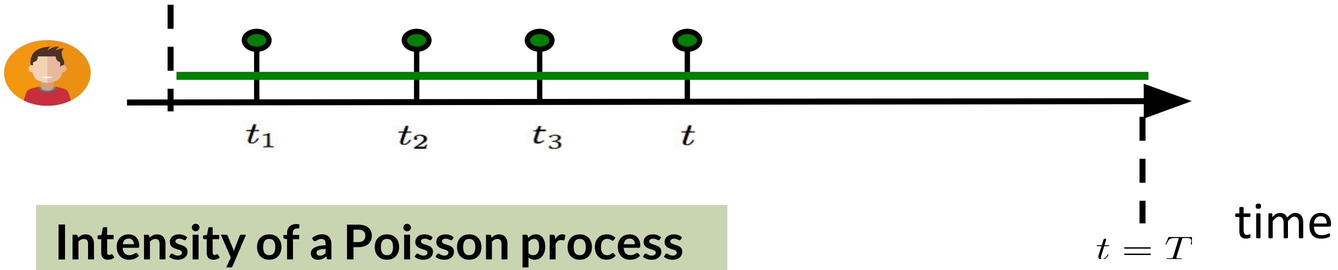
- Models arrival rate of different process

- Captures various phenomenona of interest like text

What is a temporal point process ?

A random process whose realization consists of **discrete events localized in time**

Poisson Process



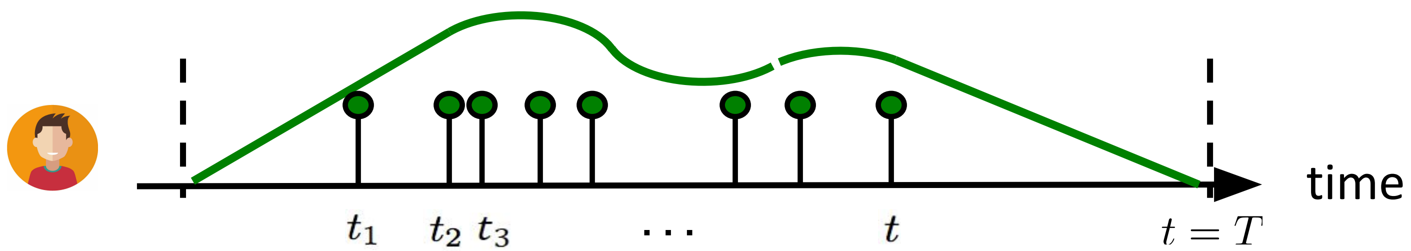
Intensity of a Poisson process

$$\lambda^*(t) = \mu$$

Observations:

1. **Intensity independent of history**
2. **Uniformly random occurrence**
3. **Time interval follows exponential distribution**

Inhomogeneous poisson process



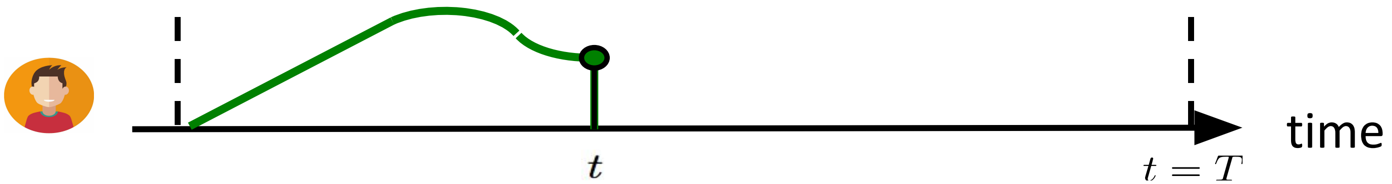
Intensity of an inhomogeneous Poisson process

$$\lambda^*(t) = g(t) \geqslant 0$$

Observations:

1. **Intensity independent of history**

Terminating (or survival) process



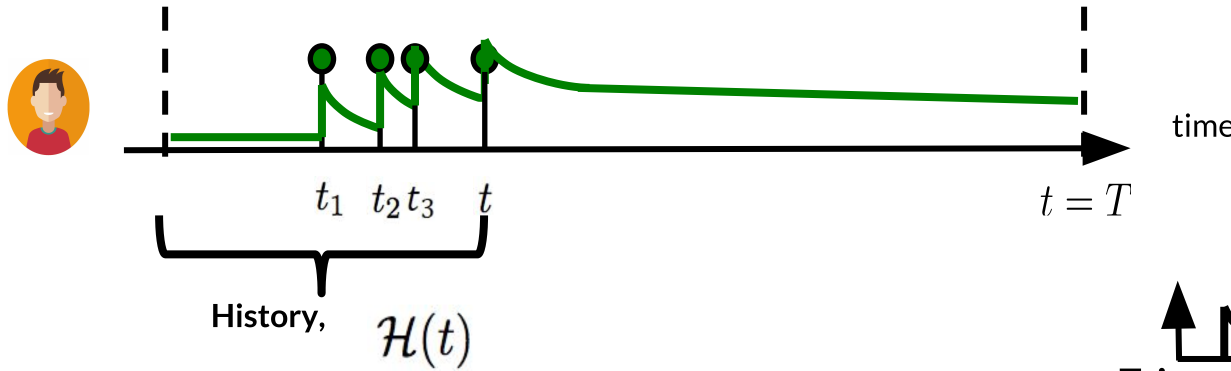
Intensity of a terminating (or survival) process

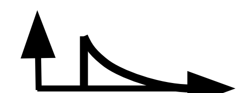
$$\lambda^*(t) = g^*(t)(1 - N(t)) \geqslant 0$$

Observations:

1. **Limited number of occurrences**

Hawkes Process





Intensity of self-exciting (or Hawkes) process:

$$\lambda^*(t) = \mu + \alpha \underbrace{\sum_{t_i \in \mathcal{H}(t)} \overbrace{\kappa_\omega(t - t_i)}^{\text{Triggering kernel}}}_{\text{The effect of the history}}$$

Two important problems in information diffusion

- Hashtag Diffusion (Temporal)
  [INFOCOM'17, IJCAI'17]

- Opinion dynamics (Data driven and Temporal)
  [NIPS'16][ICDM'17]

Hashtag Diffusion in social media

IJCAI’17, Infocom’17

Hashtag diffusion in social media - is it new?

Temporal Model

Feature engineering
(Rosenfeld et al WSDM ’16, Bourigault et al WSDM’14)

Drawbacks of existing models

Reinforced Poisson Process (RPP) AAAI'14

- Model single tweet propagation dynamics
- Model gives a non-convex objective function

Hawkes Process based models (Hawkes, Seismic) WWW'15, KDD'15

- Model single tweet propagation dynamics
- Some variation of this model is limited to long term forecasting

Pattern based models (spikeM) WWW'15

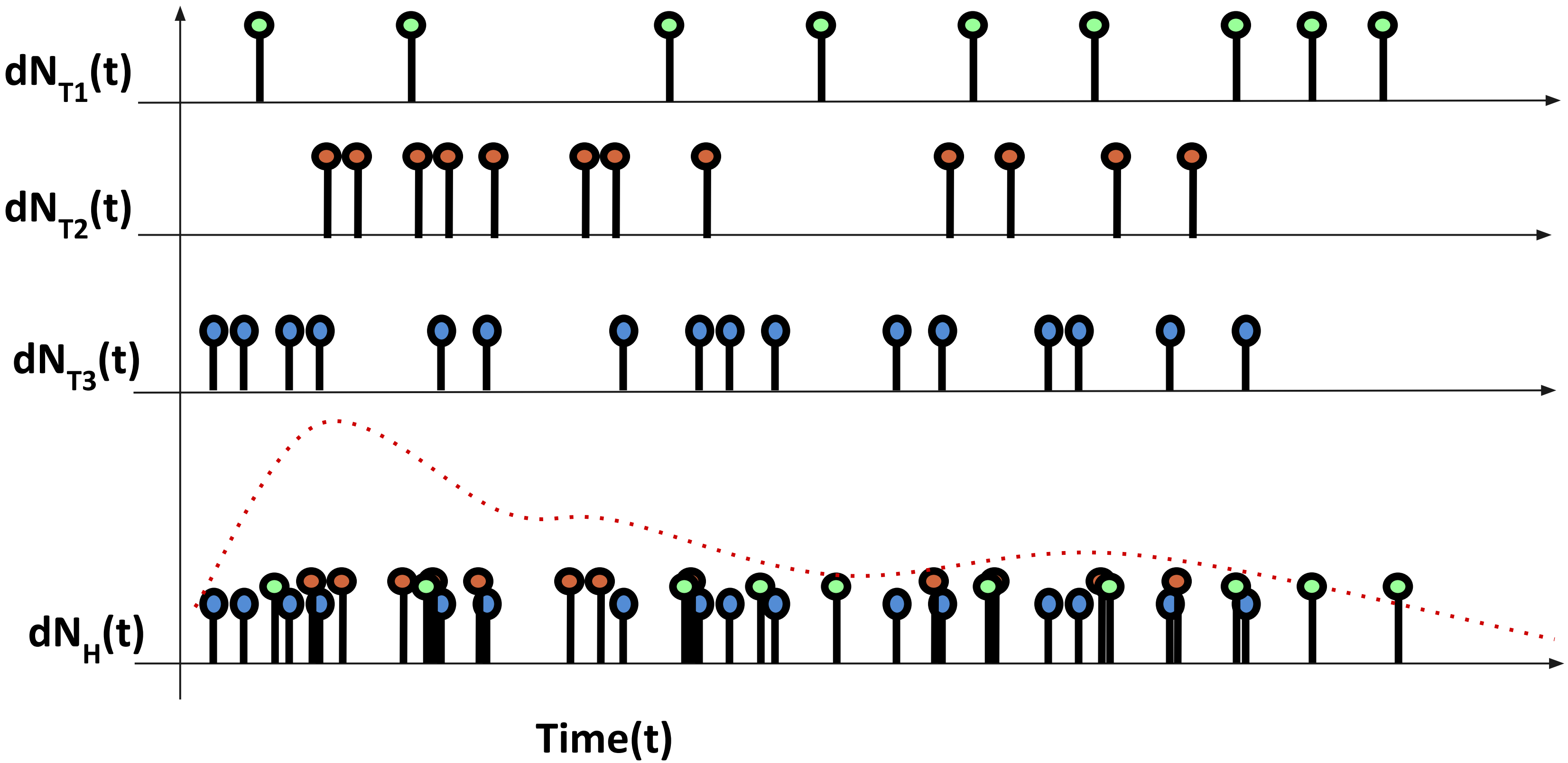
- Heavily dependent on users' past activities (6 pattern of spikes)
- Model fails if user activity changes frequently

Realistic Scenario

Hashtag is a heterogenous collection of tweets

Hashtag reflects a story better

Hashtag is a collection of tweet-chains



Hashtag diffusion in social Media

Real world reinforcement factors

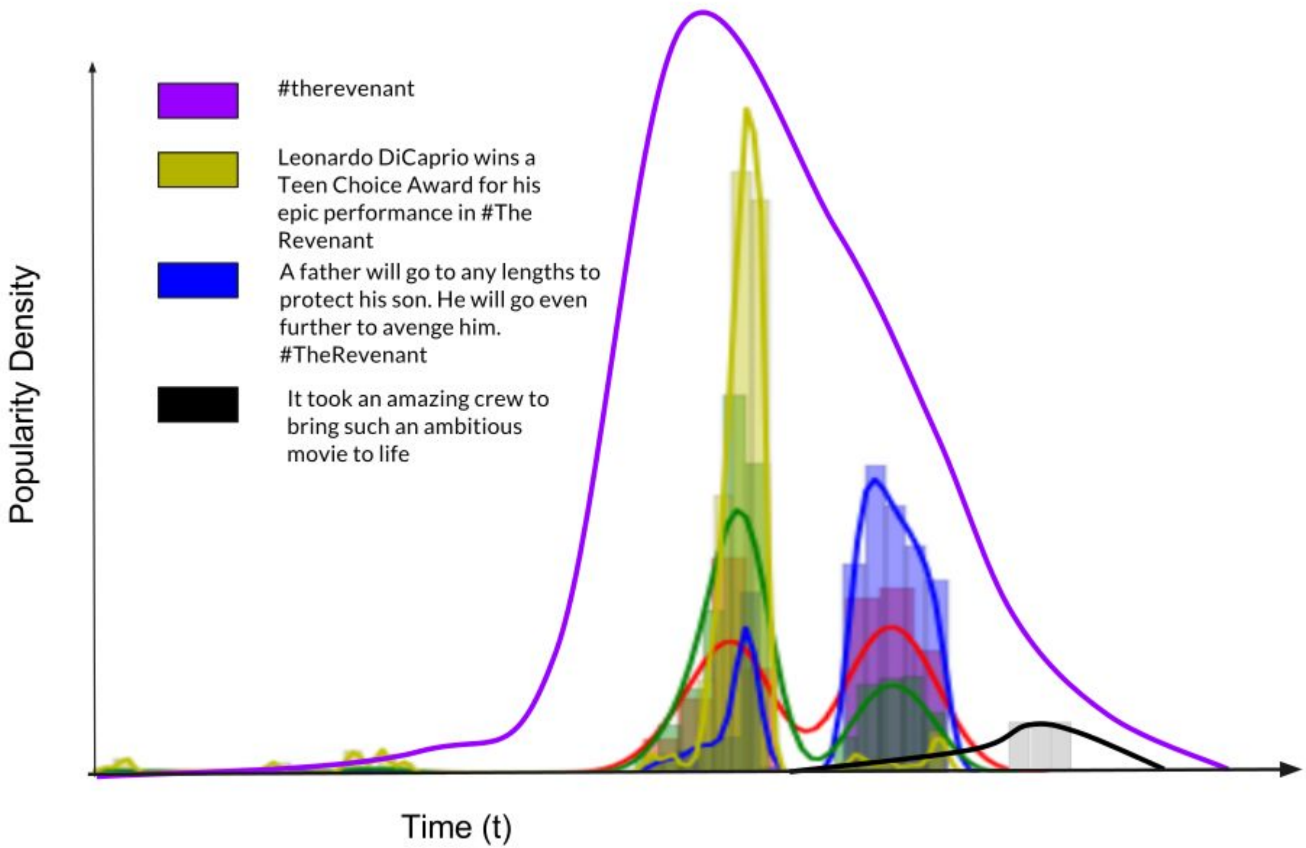
Intra-hashtag

Hashtag-tweet reinforcement

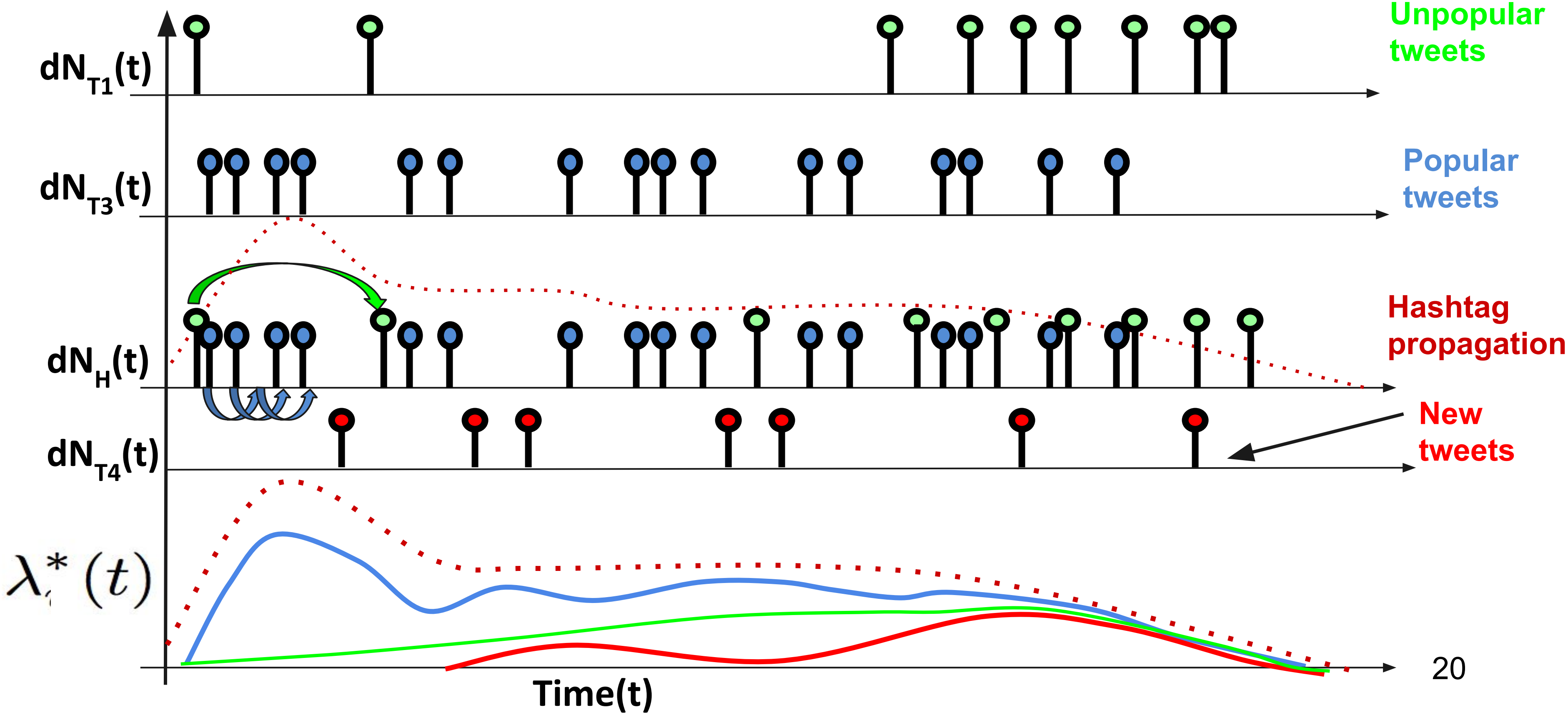
Inter-hashtag

Hashtag-hashtag competition

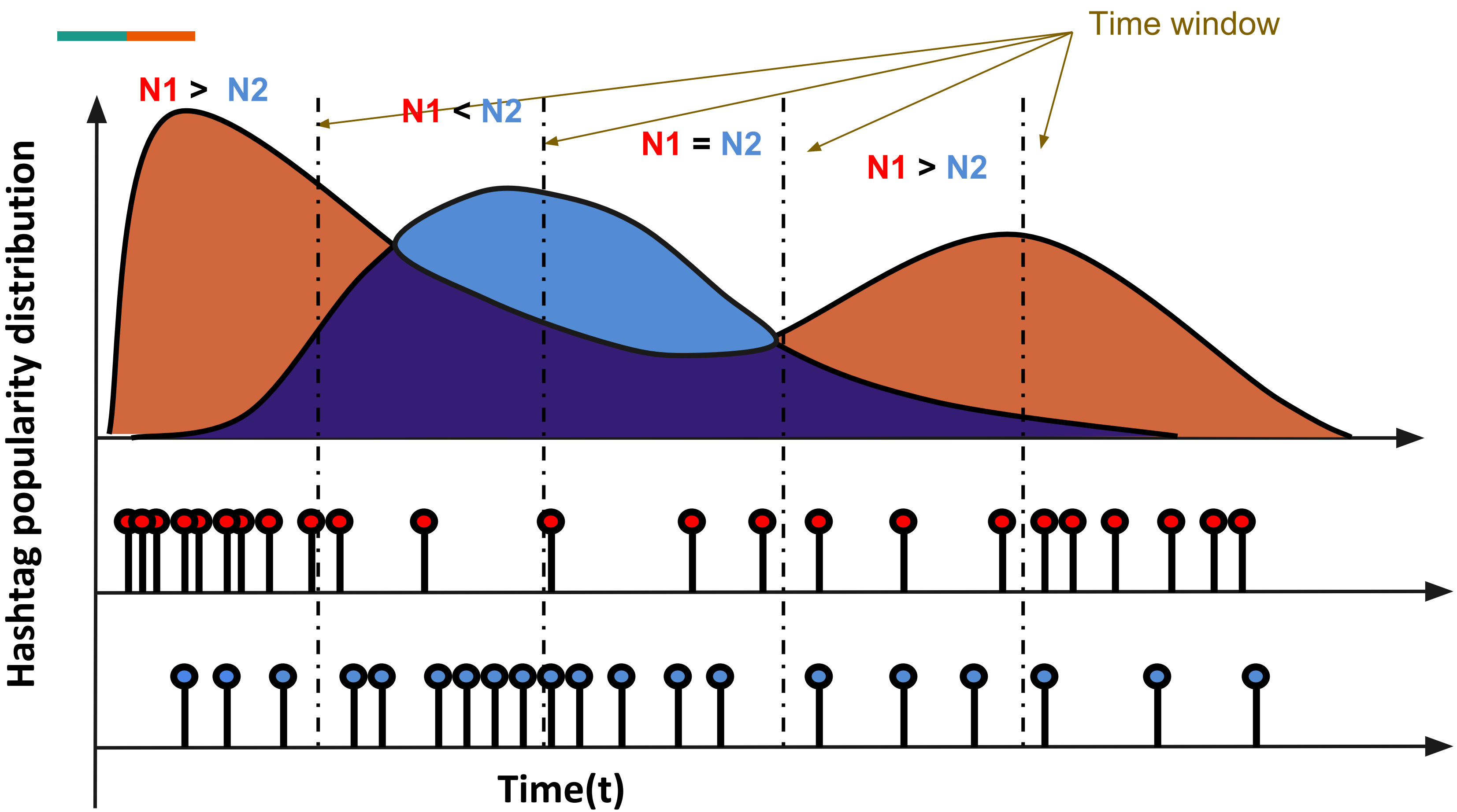
Some example from the real data



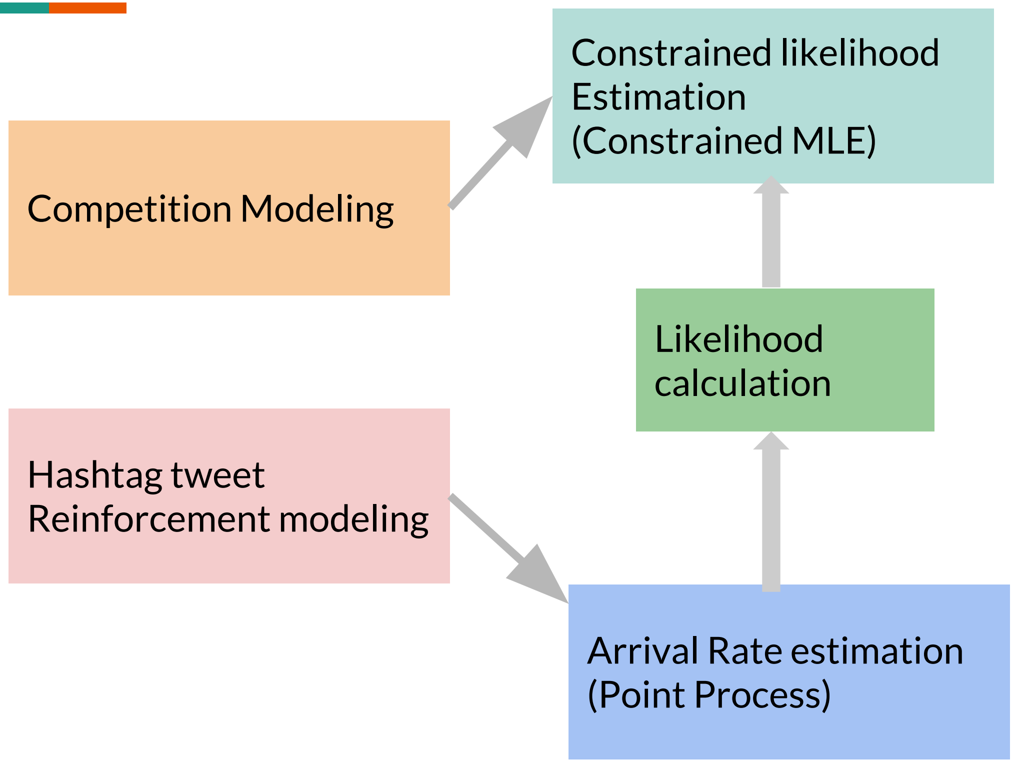
Hashtag-tweet reinforcement



Inter-hashtag competition

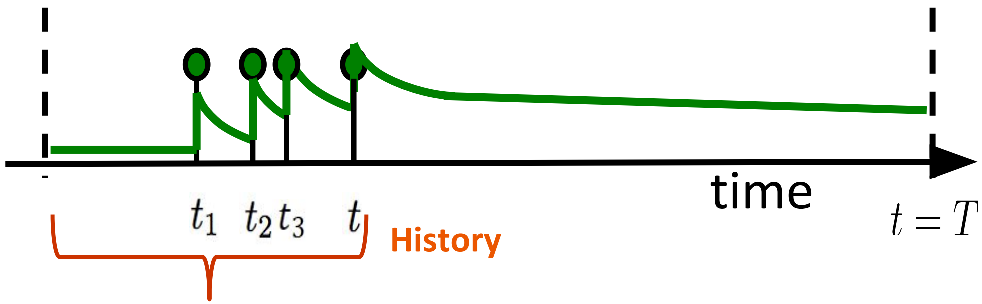


LMPP (Large Margin Point Process)



Arrival rate formulation → Likelihood estimation → Model competition as constraint

Arrival rate formulation (Hawkes process) captures the tweet reinforcement



$$\lambda_{\boldsymbol{H}}(t; \boldsymbol{k}(t)) = \underbrace{\lambda_{\boldsymbol{H},0} e^{-\epsilon t}}_{\text{Background rate}} + \sum_{j=1}^{M} \beta_{\boldsymbol{H}}^{j} \sum_{t_i \in \mathcal{H}_{\boldsymbol{H}}(t)} e^{-\underbrace{\left(\omega_j + \frac{\omega}{\boldsymbol{k}(t_i)}\right)(t - t_i)}_{\text{Triggering kernel: Effect of the previous history}}}$$

Popularity index

Triggering kernel
Effect of the previous history

Background rate

Arrival rate formulation

Likelihood estimation

Model competition as constraint

Log likelihood of observing events till time T

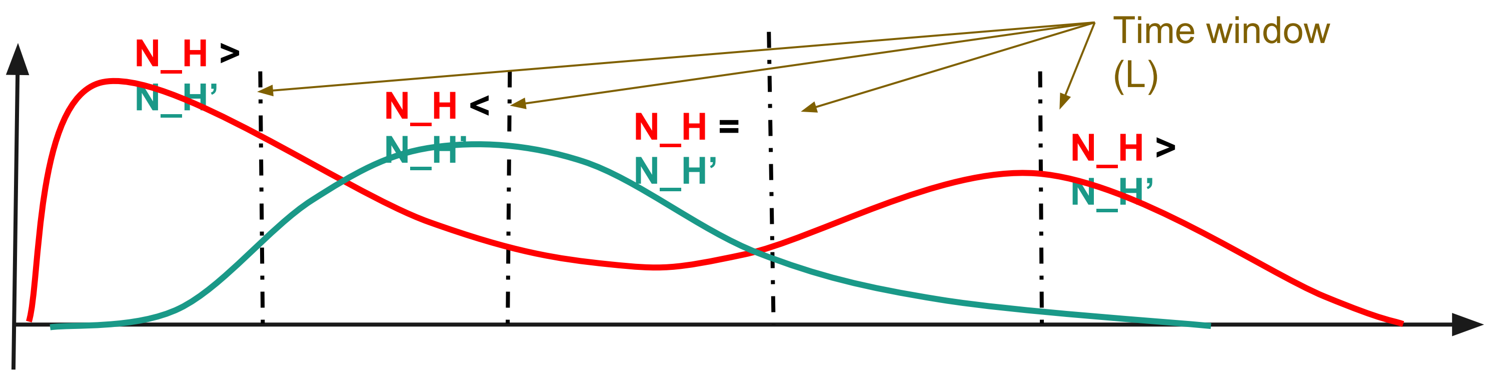
$$
\log[L(\boldsymbol{\lambda}_{\mathbb{H},0}, \boldsymbol{\beta} \,|\, \epsilon, \boldsymbol{\omega}, \omega)]
= \sum_{\boldsymbol{H} \in \mathbb{H}} \sum_{t_i \in \mathcal{H}_{\boldsymbol{H}}(T)} \log \lambda_{\boldsymbol{H}}(t_i) - \sum_{\boldsymbol{H} \in \mathbb{H}} \int_0^T \lambda_{\boldsymbol{H}}(t)\, dt
$$

Parameters of the model we want to derive

Arrival rate formulation → Likelihood estimation → Model competition as constraint

Difference in total number of posts corresponding to two hashtags when N_H > N_H'

$$\int_{iT_s}^{(i+1)T_s} (\lambda_{\boldsymbol{H}}(t) - \lambda_{\boldsymbol{H}'}(t))dt \geq 1;\ \boldsymbol{H}, \boldsymbol{H}' \in \mathbb{H},\ 0 \leq i \leq L-1$$



Dataset

- **Oscars**: 2016 Academy Award ceremony, data collected from Feb 24 to Feb 29, 2016. #leonardodicaprio,#bearstory
- **Dem-Primary**: Campaign for Hillary Clinton for USA election. Data collected from Feb to June, 2016. #nohillary #standwithhillary
- **MTV-Awards**: 2016 MTV Award Star program, collected from April 3 to April 12, 2016. #mtvawardsstar
- **Nepal-earthquake**: Earthquake in Nepal, from April 25 to May 1, 2015. ,#nepalrelief #helpnepal
- **BBD**: Bigbillionday sale of Flipkart India, from October 6 to October 8, 2014. #bigbillionsale
- **Copa**: Copa America Football tournament, from June 3 to June 26, 2016 #copaamericaentd, #aocopa
- **TWC20**: ICC World Cup T20, India, from March 8 to April 3, 2016 #indvspak #indvsban

Evaluation protocol

Forecasting hashtag popularity

- Learn the model parameters and forecast

Rank prediction of competing hashtags

- Predicting jump in rank

Results for jump prediction

JUMP A sudden change in position in the rank-list by

	Avg. Precision					
Datasets	**LMPP**	**HTR**	**RPP**	**Hawkes**	**SEISMIC**	**SpikeM**
Oscars	**0.74**	0.54	0.32	*0.38*	0.31	0.33
Nepal-Earthquake	**0.61**	0.60	*0.37*	0.28	0.40	0.44
Dem-Primary	**0.69**	0.56	*0.48*	0.30	0.45	0.34
BBD	**0.66**	0.48	0.32	*0.55*	0.31	0.43
Copa	**0.72**	0.29	0.42	*0.60*	0.29	0.34
T20WC	**1.0**	0.32	0.64	0.10	0.29	*0.65*

LMPP outperforms other methods however there is no clear 2^{nd} best

Learning Opinion Dynamics in Social Networks

NIPS'16

Opinions in Social media

People's opinion about political discourse

CNN

How social media is revolutionizing debates

The New York Times

Campaigns Use Social Media to Lure Younger Voters

Investors' sentiment about stocks

Nasdaq

Leveraging Social Media

Business Insider UK

Startups are setting up funds based on what is trending on Twitter

Brand sentiment and reputation

TC

Twitter Unveils A New Set Of Brand-Centric Analytics

The New York Times

Social Media Are Giving a Voice to Taste Buds

Opinion modeling in social media - is it new?

There are a lot of theoretical models of opinion dynamics, but…

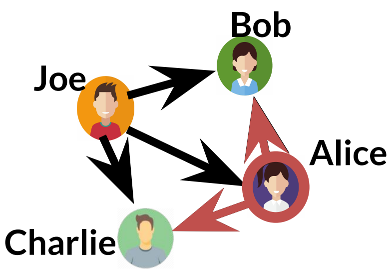
- Updated in discrete time
- Do not distinguish between latent and expressed opinions
- Difficult to learn from data
- Focus on steady state, neglecting transient behavior

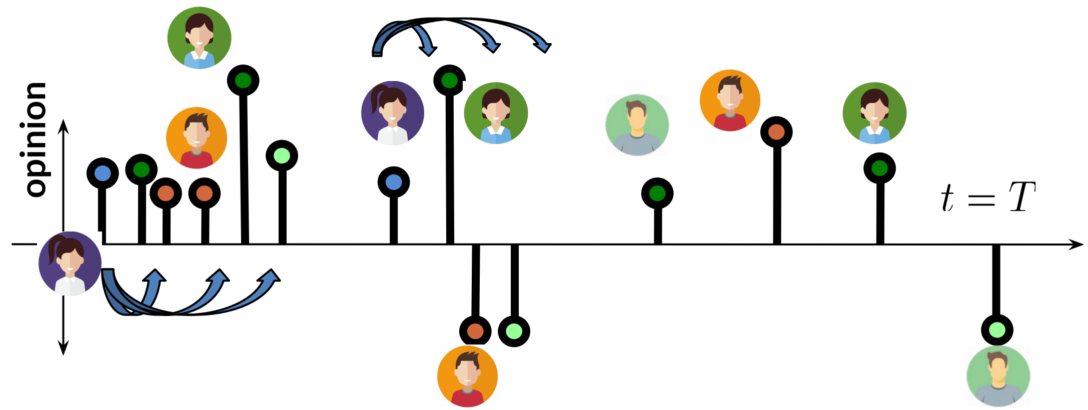
Difference with the hashtag diffusion model

User interaction graph

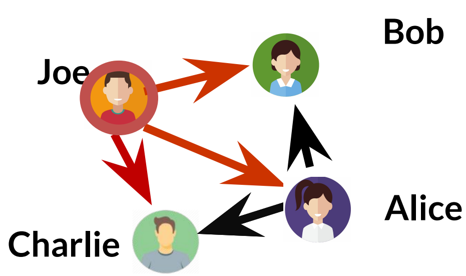
Utilizing text of the message (sentiment)

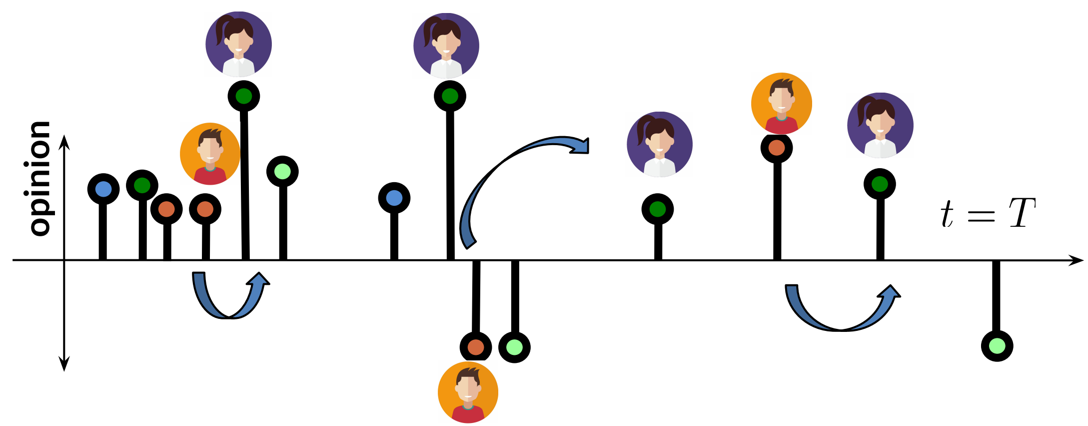
Key Idea: Temporal and Informational influence





Alice is highest temporal influence





Joe has highest informational influence

Learning opinion dynamics (SLANT)

Difference between the Hashtag and Opinion Dynamics Modeling

Previous Model

$\lambda^*(t)$

Current Model

Intensity of the arrival

$\lambda_u^*(t)$

Opinion

$x_u^*(t)$

Learning opinion dynamics (SLANT)

Model intensity of the messages

Exponential kernel (memory)

$$\underbrace{\lambda_u^*(t)}_{\text{User's message intensity}} = \underbrace{\mu_u}_{\text{Messages on her own initiative}} + \sum_{v\in u\cup\mathcal{N}(u)} b_{vu} \underbrace{\sum_{e_i\in\mathcal{H}_v(t)} \kappa(t-t_i)}_{\text{Previous messages by user v}}$$

Temporal influence from user v on user u

Learning opinion dynamics (SLANT)

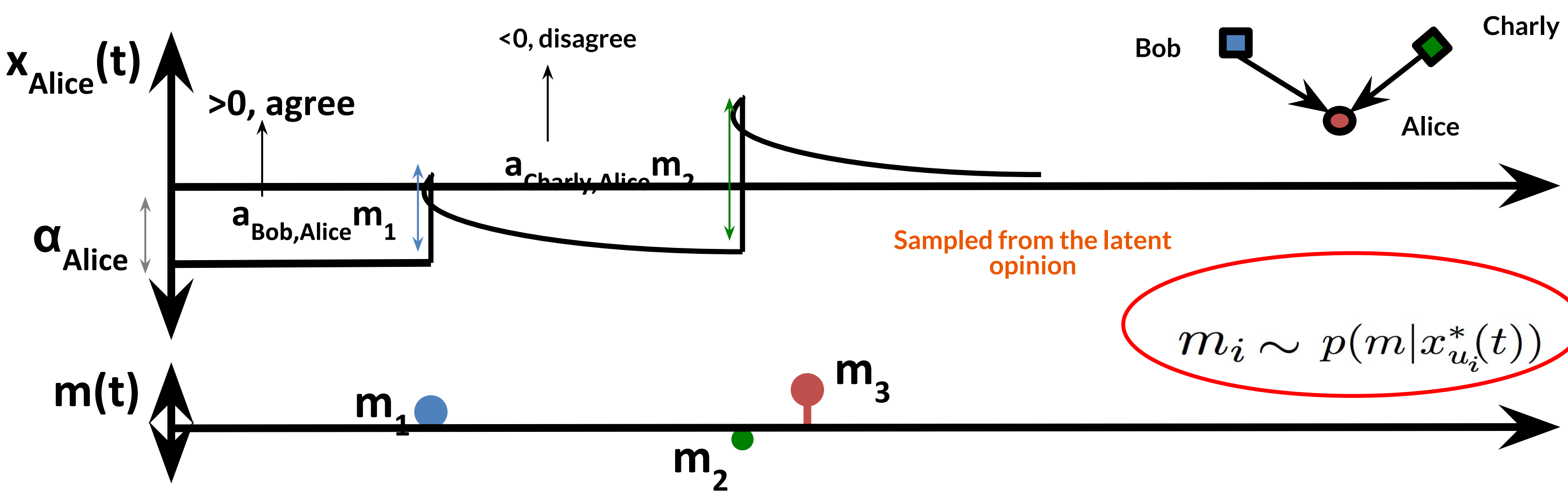
Stochastic Process for opinions

Exponential kernel (memory)

$$\underbrace{x_u^*(t)}_{\text{User's latent opinion}} = \underbrace{\alpha_u}_{\text{User's initial opinion}} + \sum_{v \in \mathcal{N}(u)} a_{vu} \underbrace{\sum_{e_i \in \mathcal{H}_v(t)} m_i\, g(t - t_i)}_{\text{Previous expressed opinions y user v}}$$

Informational influence from user v on user u

Not captured previously



Sampled from the latent opinion

$$m_i \sim p(m \mid x_{u_i}^*(t))$$

Learning opinion dynamics (SLANT)

Efficient parameter estimation using MLE

Find *optimal* parameters using **maximum likelihood estimation (MLE)**:

$$\underset{\boldsymbol{\alpha},\boldsymbol{\mu}\geq 0,\boldsymbol{A},\boldsymbol{B}\geq 0}{\text{maximize}} \underbrace{\sum_{e_i\in\mathcal{H}(T)} \log p(m_i|x^*_{u_i}(t_i))}_{\textbf{Message sentiments}} + \underbrace{\sum_{e_i\in\mathcal{H}(T)} \log \lambda^*_{u_i}(t_i) - \sum_{u\in\mathcal{V}} \int_0^T \lambda^*_u(\tau)\, d\tau}_{\textbf{Message times}}$$

For a large class of sentiment distributions, the MLE problem is **parallelizable and convex**.

Markov property allows us to **compute sums and integrals in linear time**!

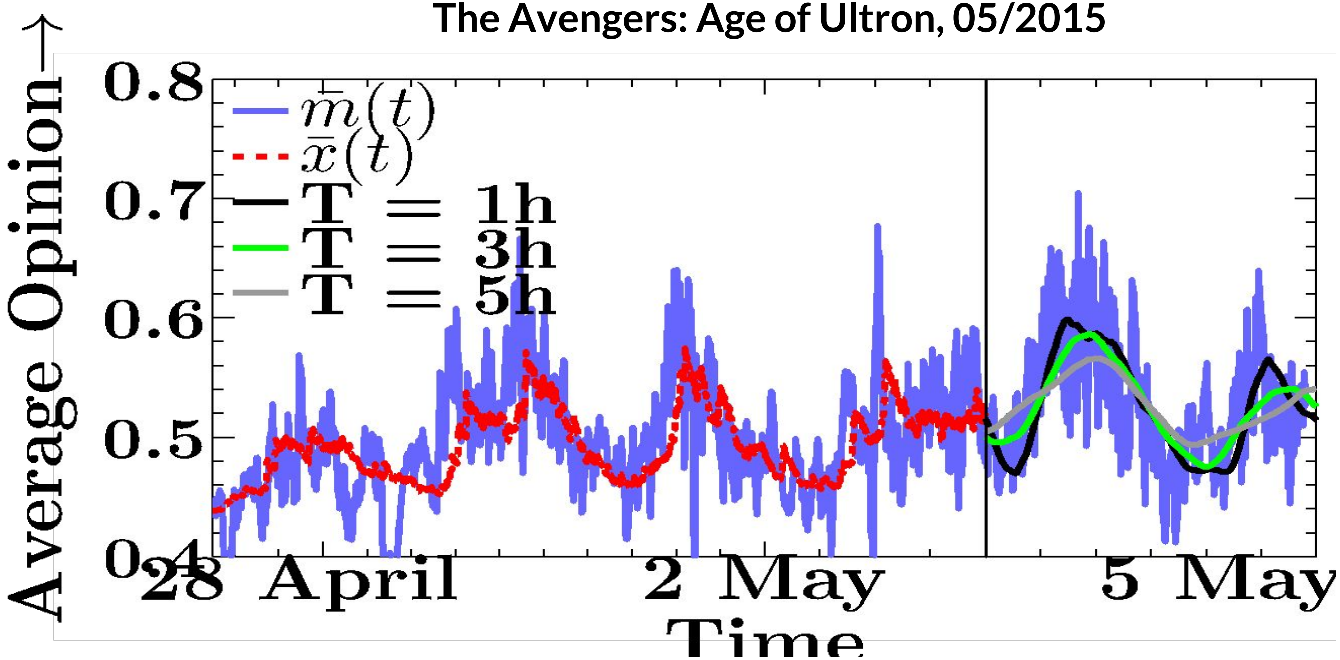
Real Data Experiments

- Delhi Assembly Election, 12/2013
- The Avengers: Age of Ultron, 05/2015
- Mayweather vs Pacquiao, 05/2015
- Bollywood star incident, 05/2015
- US elections, 04/2016

Real Data Experiments

Macroscopic Forecasting

The Avengers: Age of Ultron, 05/2015



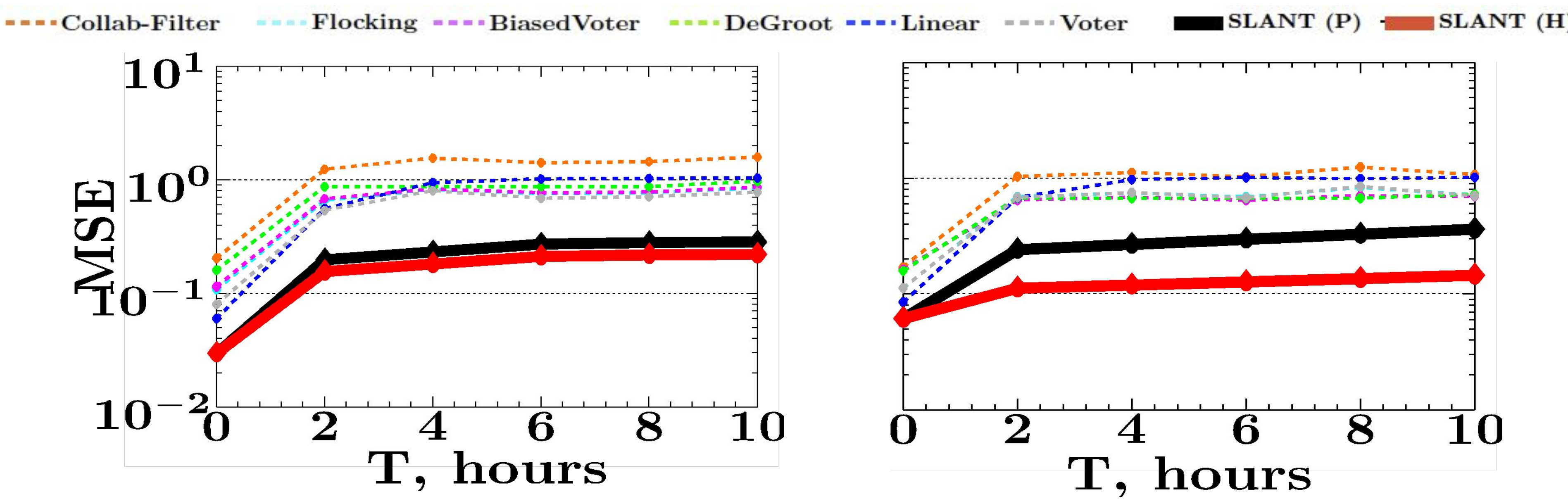
The forecasted opinion becomes less accurate as T increases, as one may expect.

Real Data Experiments

Microscopic forecasting, MSE

Delhi Assembly Election, 12/2013

US elections, 04/2016



Our model (in red) outperforms state of the art in terms of MSE, often by orders of magnitude

Learning **Nonlinear** Opinion Dynamics in Social Networks

ICDM 17

SLANT+: A nonlinear departure from SLANT

SLANT

$$\lambda_u^*(t) = f_u(\cup_{v\in\mathcal{N}(u)}\mathcal{H}_v(t))$$

Fixed, linear, semi-coupled

Set of past events

$$x_u^*(t) = g_u(\cup_{v\in\mathcal{N}(u)}\mathcal{H}_v(t))$$

Informational Influence

Coupling

Temporal Influence

Nonlinear Modeling

SLANT+

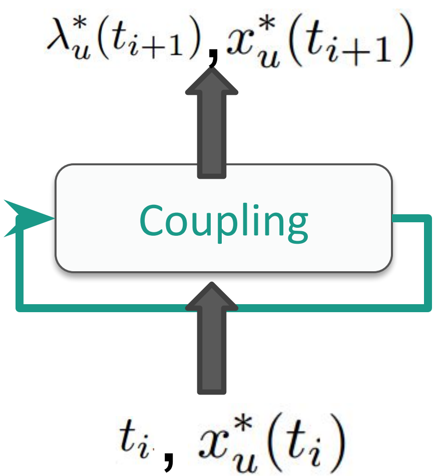
$$\lambda_u^*(t) = f_u(\cup_{v\in\mathcal{N}(u)}\mathcal{H}_v(t))$$

Nonlinear Modeling

Coupling

$$x_u^*(t) = g_u(\cup_{v\in\mathcal{N}(u)}\mathcal{H}_v(t))$$

SLANT+: An intuitive approach

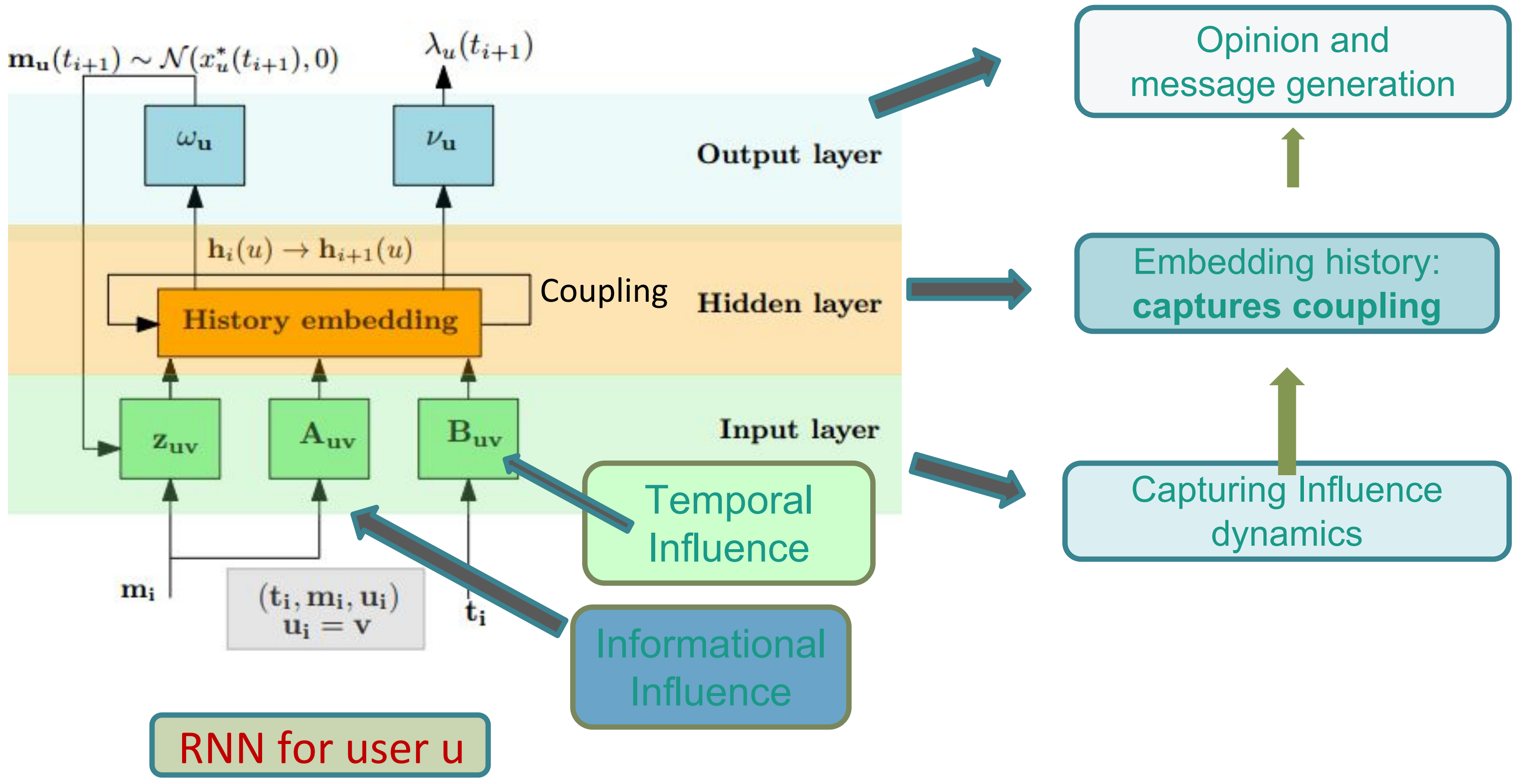


RMTPP, Du et al. 16, In KDD 2016

Capturing nonlinearity via recurrent neural network

A single RNN for whole network	Small RNNs per user
1. Expressivity issues 2. Training issues	1. Explainable 2. Decentralized

Nonlinear Modeling:
A networked guided RNN approach



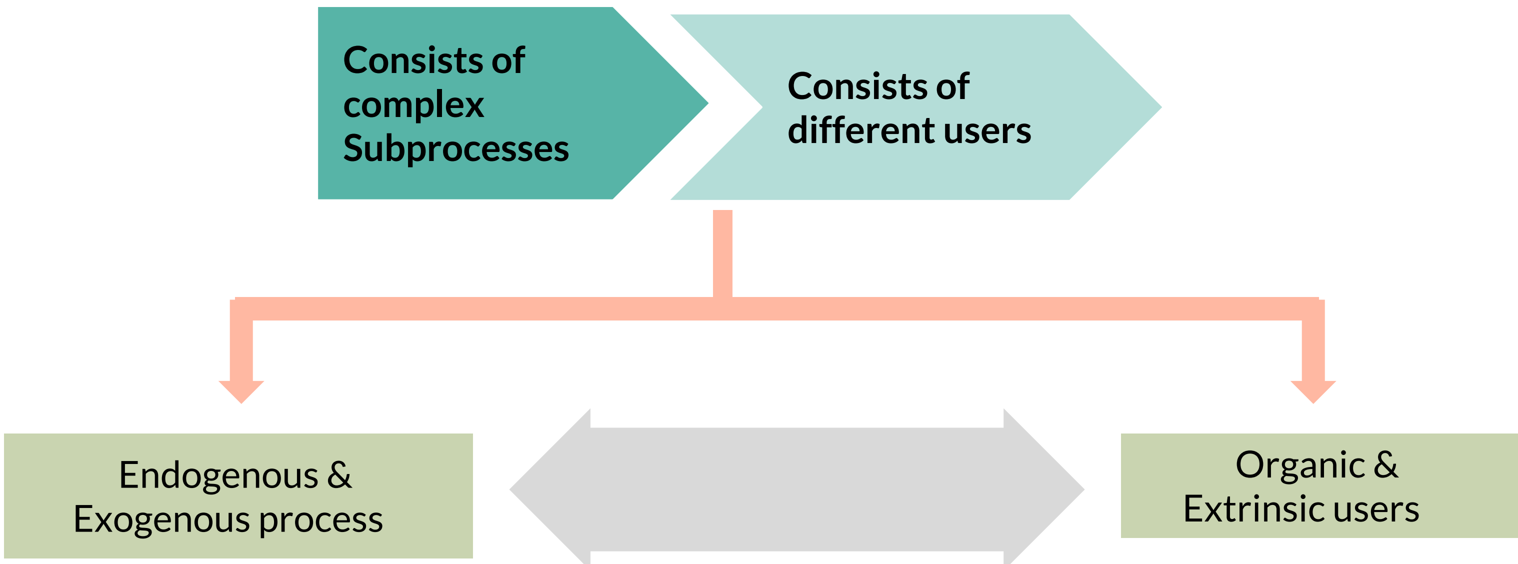
Experimental Results

	Mean Squared Error						
Dataset	SLANT+	SLANT	BVoter	Voter	AsLM	DeGroot	Flocking
Movie	0.007 (90.79)	0.076	0.755	0.822	1.367	0.499	0.69
Politics	0.038 (82.16)	0.213	0.771	0.670	1.023	0.875	0.76
Fight	0.045 (79.82)	0.223	1.351	1.477	1.514	0.963	1.31
Bollywood	0.049 (88.71)	0.434	2.015	2.132	3.579	1.724	1.94
Series	0.049 (32.88)	0.073	0.287	0.536	0.796	0.533	0.49
	Failure Rate						
Movie	0.00 (–)	0.0	0.0	0.0	0.0	0.0	0.0
Politics	0.03 (80.0)	0.15	0.51	0.51	0.51	0.46	0.58
Fight	0.06 (53.85)	0.13	0.59	0.59	0.54	0.43	0.54
Bollywood	0.01 (93.33)	0.15	0.43	0.44	0.50	0.42	0.43
Series	0.01 (66.67)	0.03	0.31	0.41	0.33	0.47	0.48

Summary

- Generalised framework that captures popular dynamics on social media
- Besides modeling, rigorous experimentation over diverse set of data shows the performance at critical junctions are significantly better

Subprocesses in information and opinion dynamics



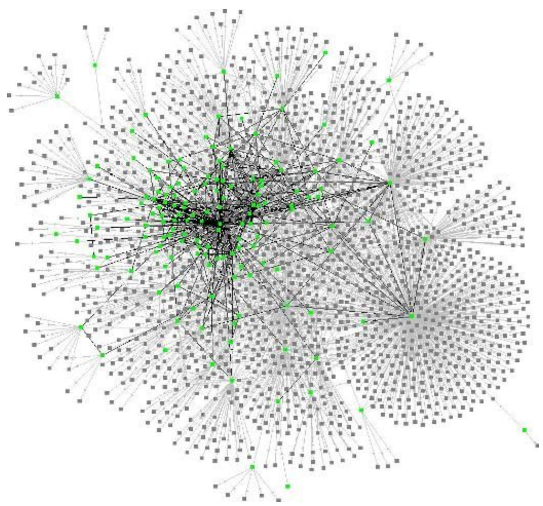
Demarcating Endogenous and Exogenous Opinion Diffusion Process in Social Networks

WWW 18

Opinion dynamics….

Initiated by sociologists and physicists

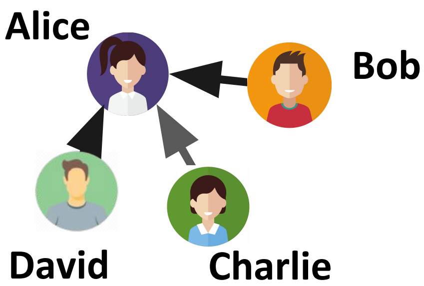
- ❑ Voter Model (Clifford and Sudbery 1973, Cox 1991, etc.)
- ❑ DeGroot Model (DeGroot, 1974)

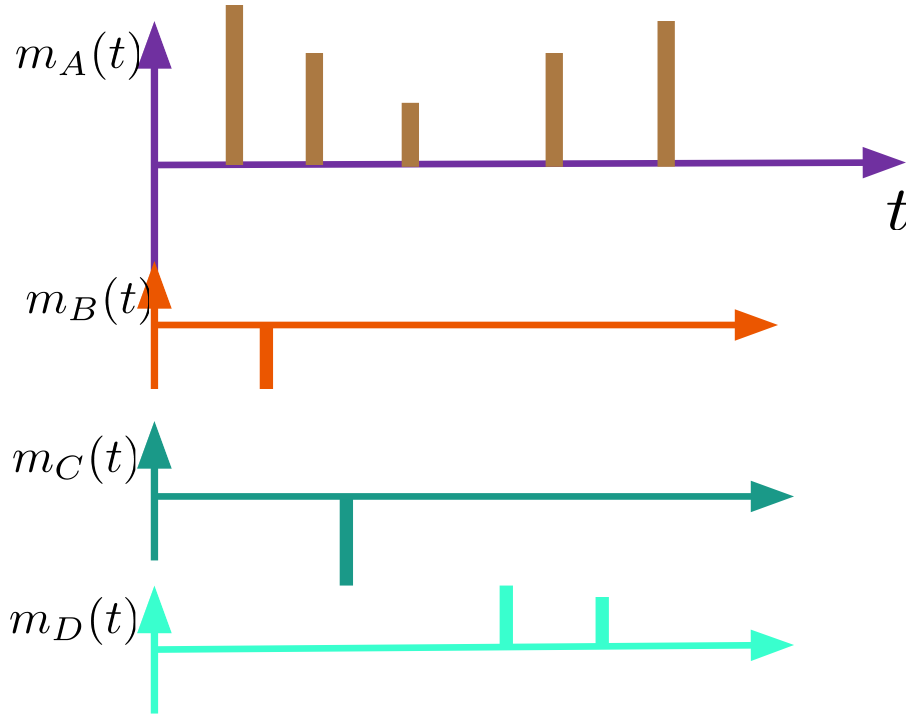


Data driven models

- ❑ Biased Voter Model (Das et al 2014)
- ❑ Linear Model (De et al 2014)
- ❑ SLANT, SLANT+ (De et al 2016, 2017)

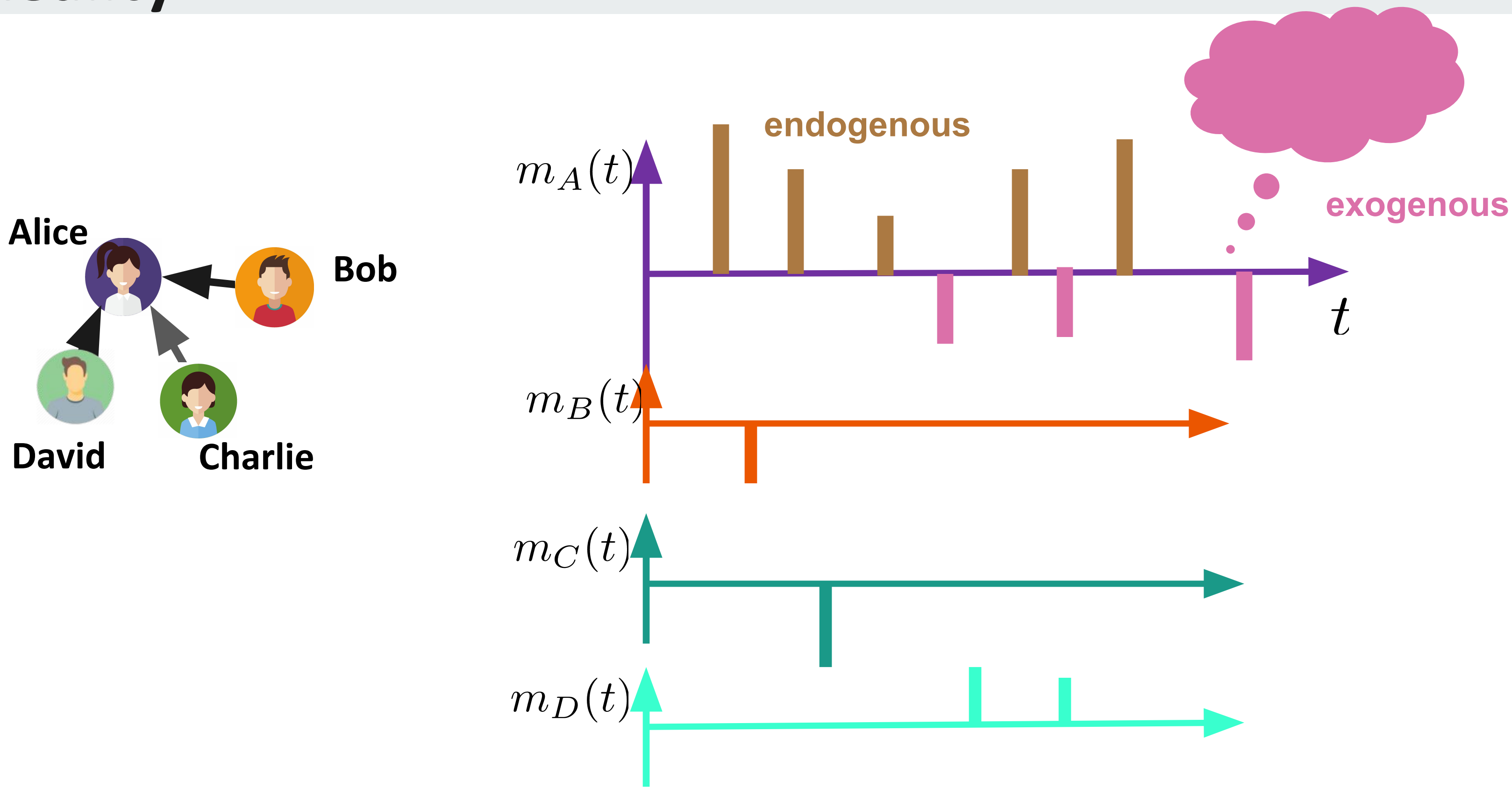
The modus operandi….





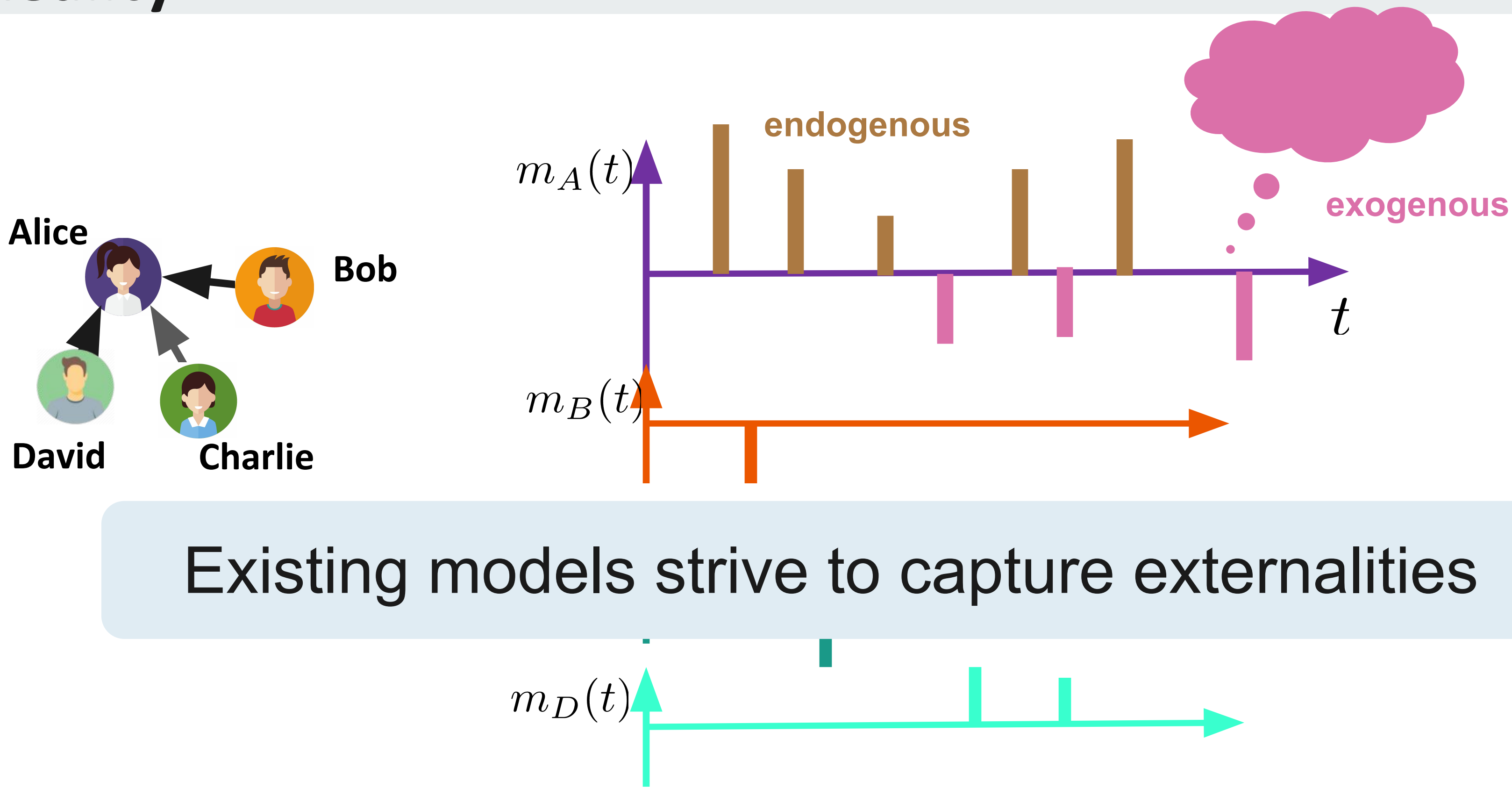
Users' opinions are influenced by neighbors

Reality



Users also express opinions gathered from external sources

Reality



Existing models strive to capture externalities

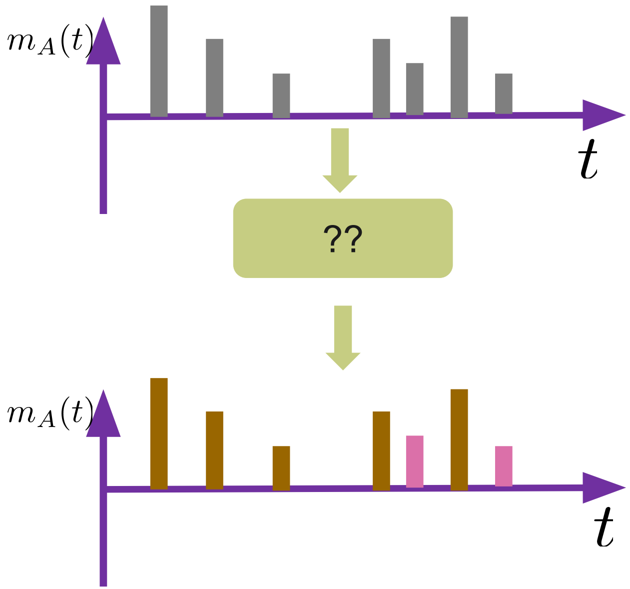
Users also express opinions gathered from external sources

Why externalities have been a big challenge for data-driven models?

The roadblocks

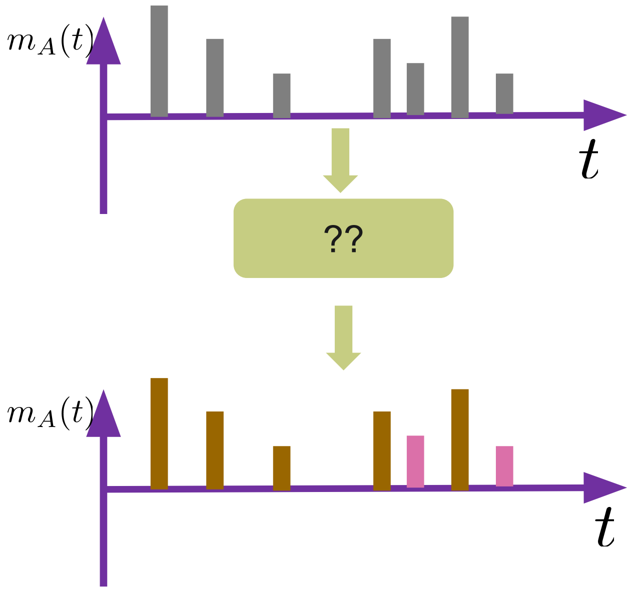
Messages are unlabeled

No supervised techniques work



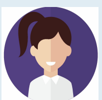
A simple solution...

Content based clustering

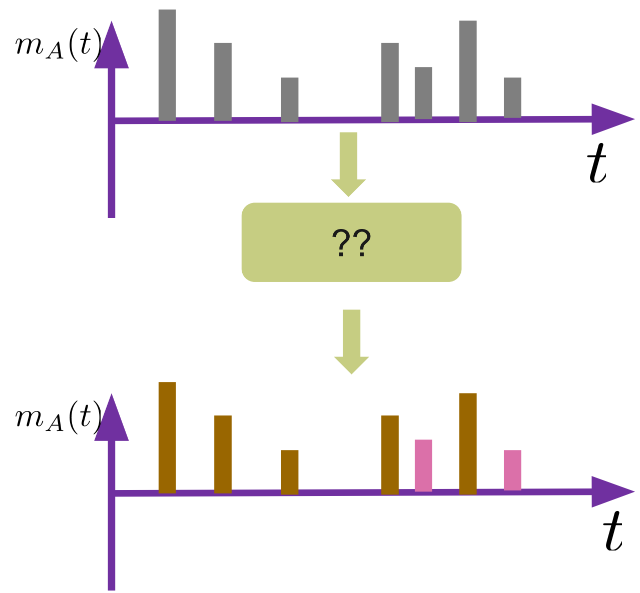


A simple solution...

Content based clustering

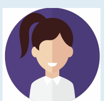
 NY Times hints at a new trade war

 A timely article in NY Times…..



A simple solution...

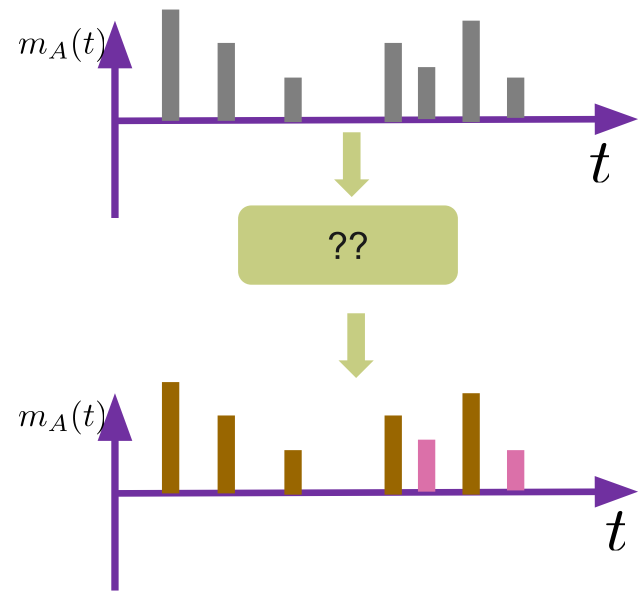
Content based clustering



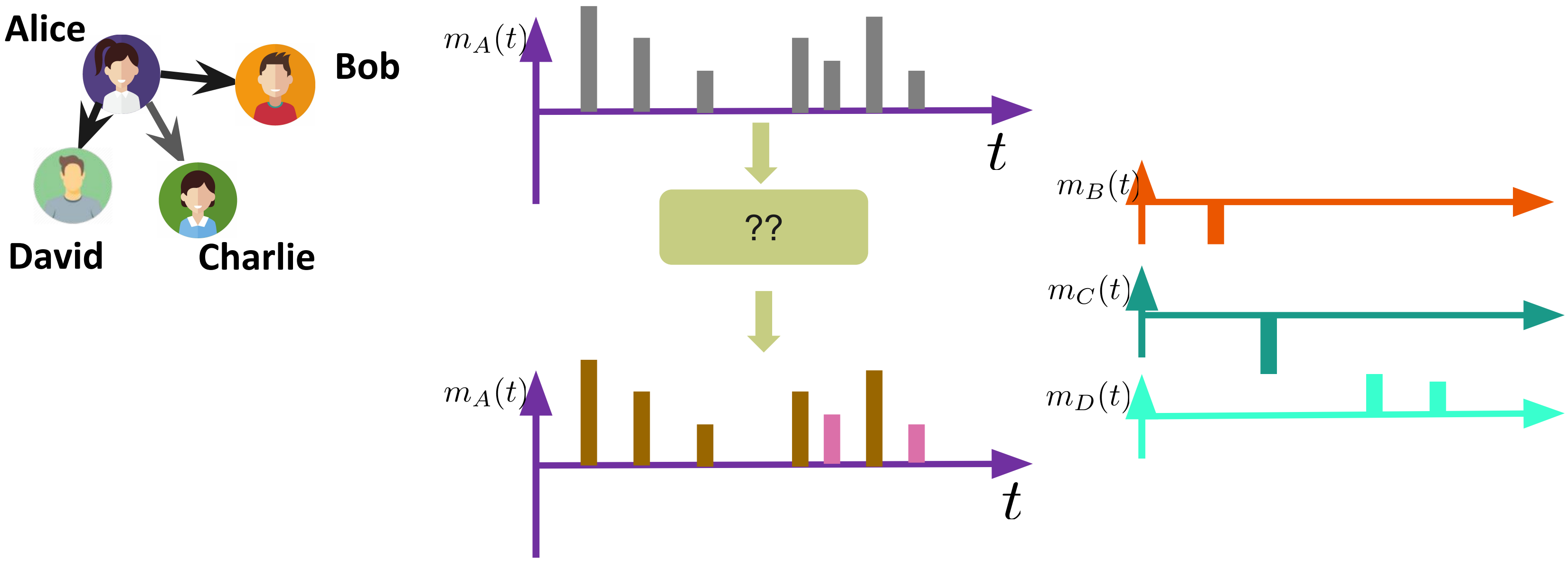
NY Times hi

…n NY Times…..

Same content can be endogenous and exogenous



Our work..



Can we demarcate endogenous and exogenous messages

Models for endogenous dynamics

Voter model — Clifford et al 1973

Degroot model — DeGroot et al 1974

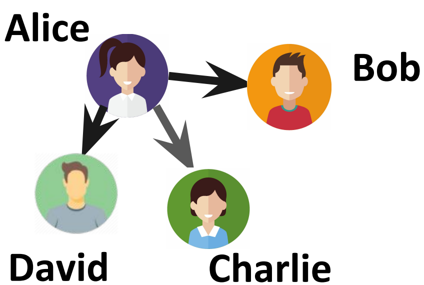
Asynchronous linear model — De et al 2014

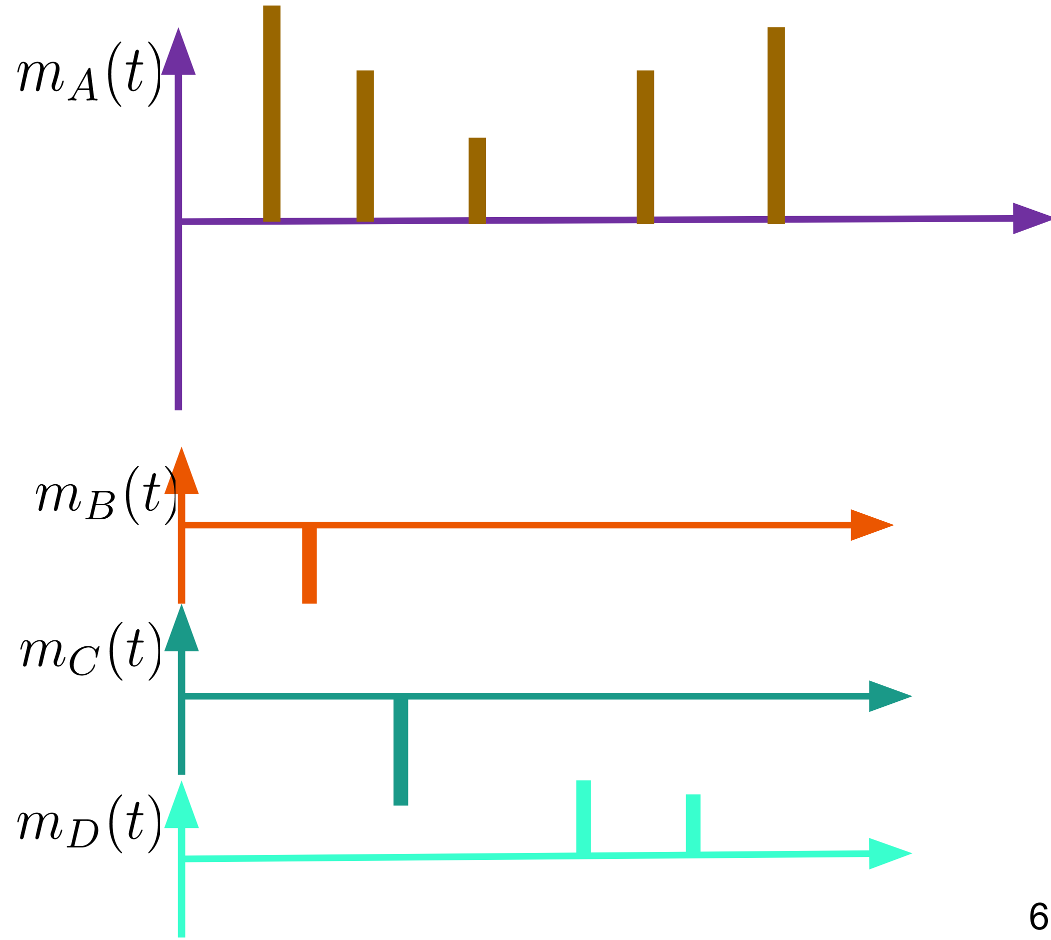
SLANT — De et al 2016

SLANT

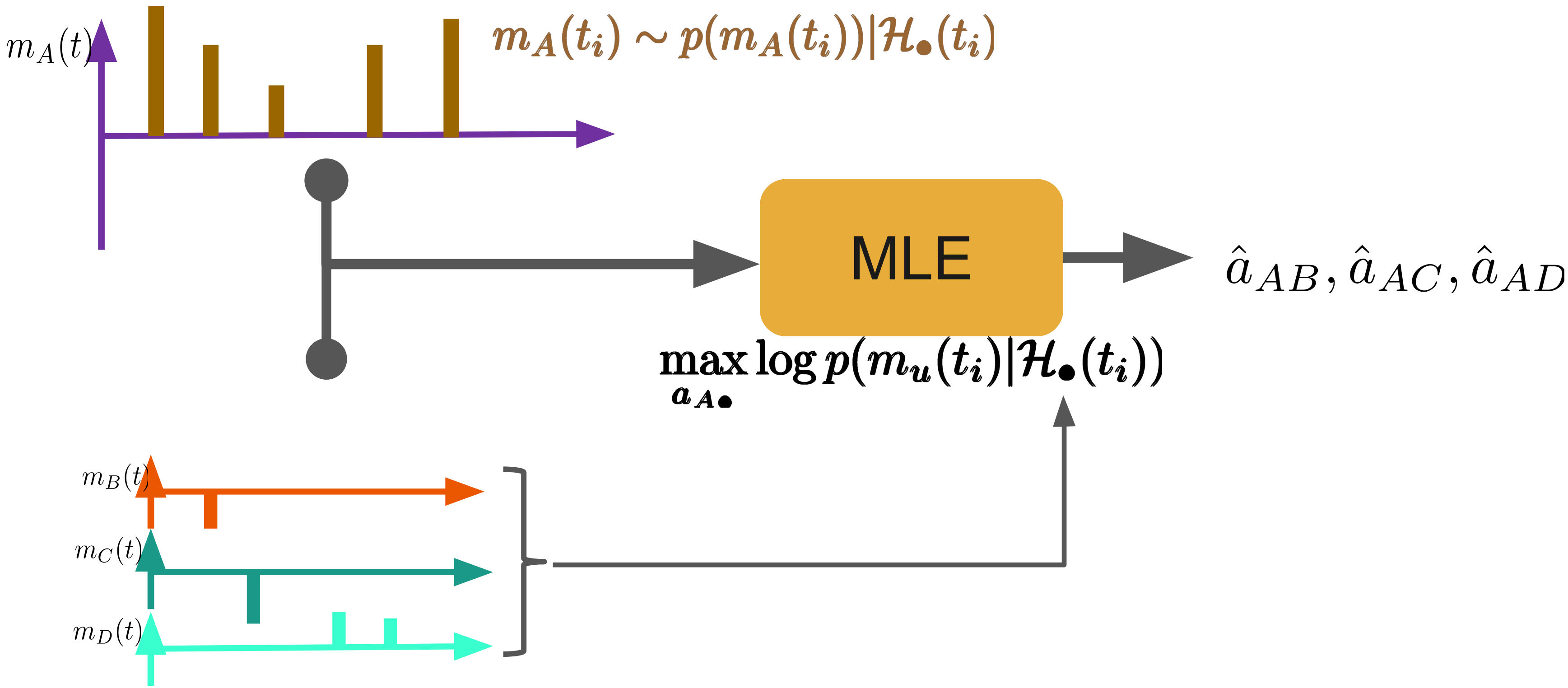
$$p(m_A(t_i))|\mathcal{H}_\bullet(t_i)$$

$$a_{AB}, a_{AC}, a_{AD}$$

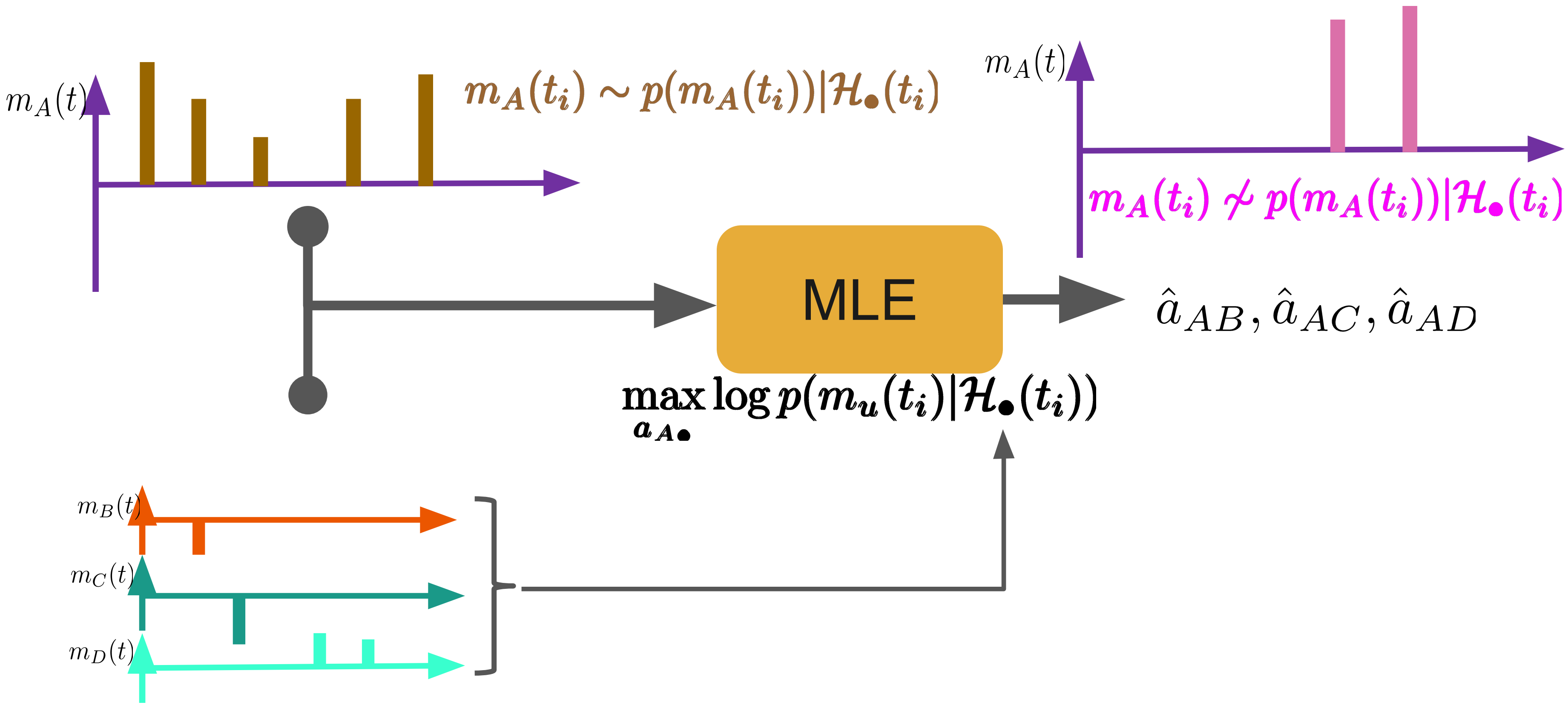




SLANT: Parameter estimation

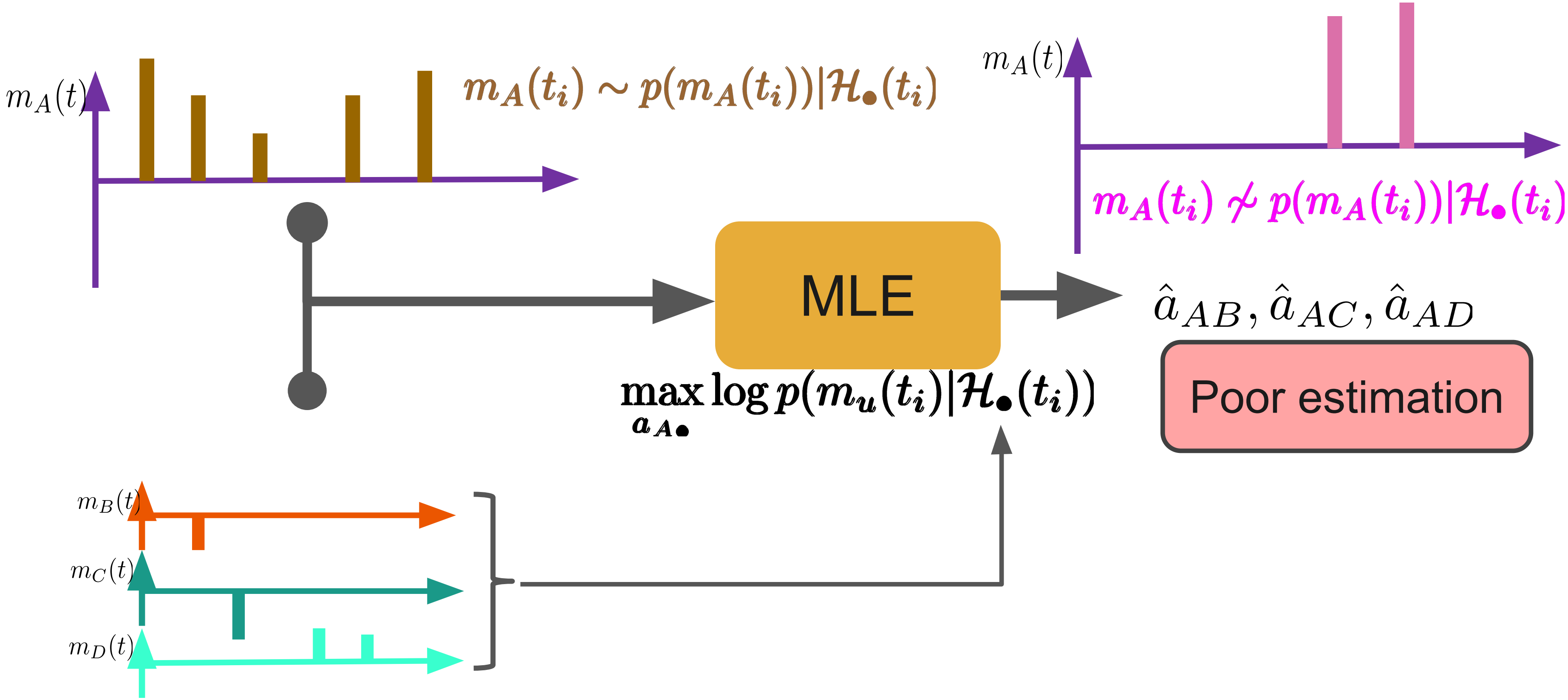


Reality



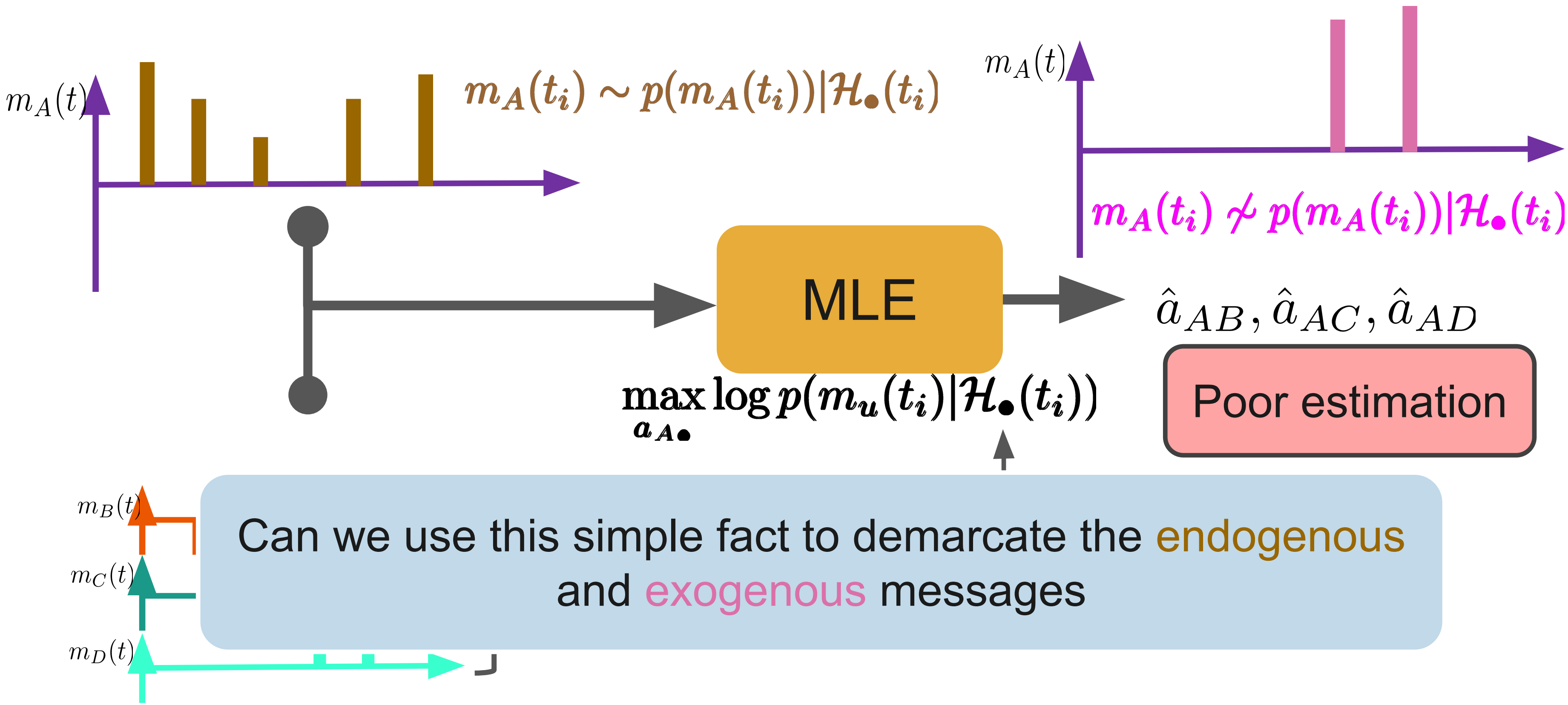
$m_A(t)$
$m_A(t_i) \sim p(m_A(t_i))|\mathcal{H}_\bullet(t_i)$
$m_A(t)$
$m_A(t_i) \not\sim p(m_A(t_i))|\mathcal{H}_\bullet(t_i)$
MLE
$\hat{a}_{AB}, \hat{a}_{AC}, \hat{a}_{AD}$
$\max_{a_{A\bullet}} \log p(m_u(t_i)|\mathcal{H}_\bullet(t_i))$
$m_B(t)$
$m_C(t)$
$m_D(t)$

Reality



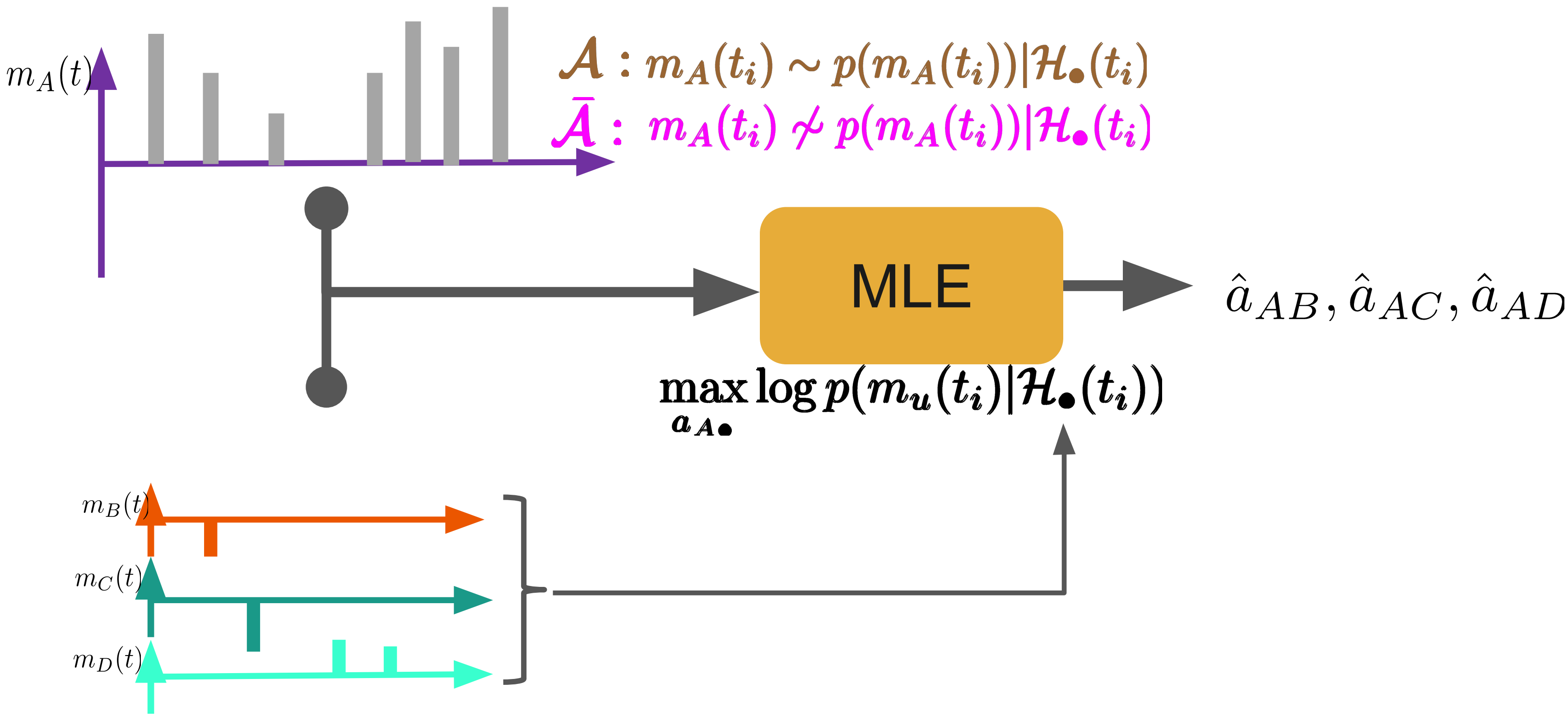
$m_A(t)$
$m_A(t_i) \sim p(m_A(t_i))|\mathcal{H}_\bullet(t_i)$
$m_A(t)$
$m_A(t_i) \nsim p(m_A(t_i))|\mathcal{H}_\bullet(t_i)$
MLE
$\hat{a}_{AB}, \hat{a}_{AC}, \hat{a}_{AD}$
Poor estimation
$\max_{a_{A\bullet}} \log p(m_u(t_i)|\mathcal{H}_\bullet(t_i))$
$m_B(t)$
$m_C(t)$
$m_D(t)$

Reality



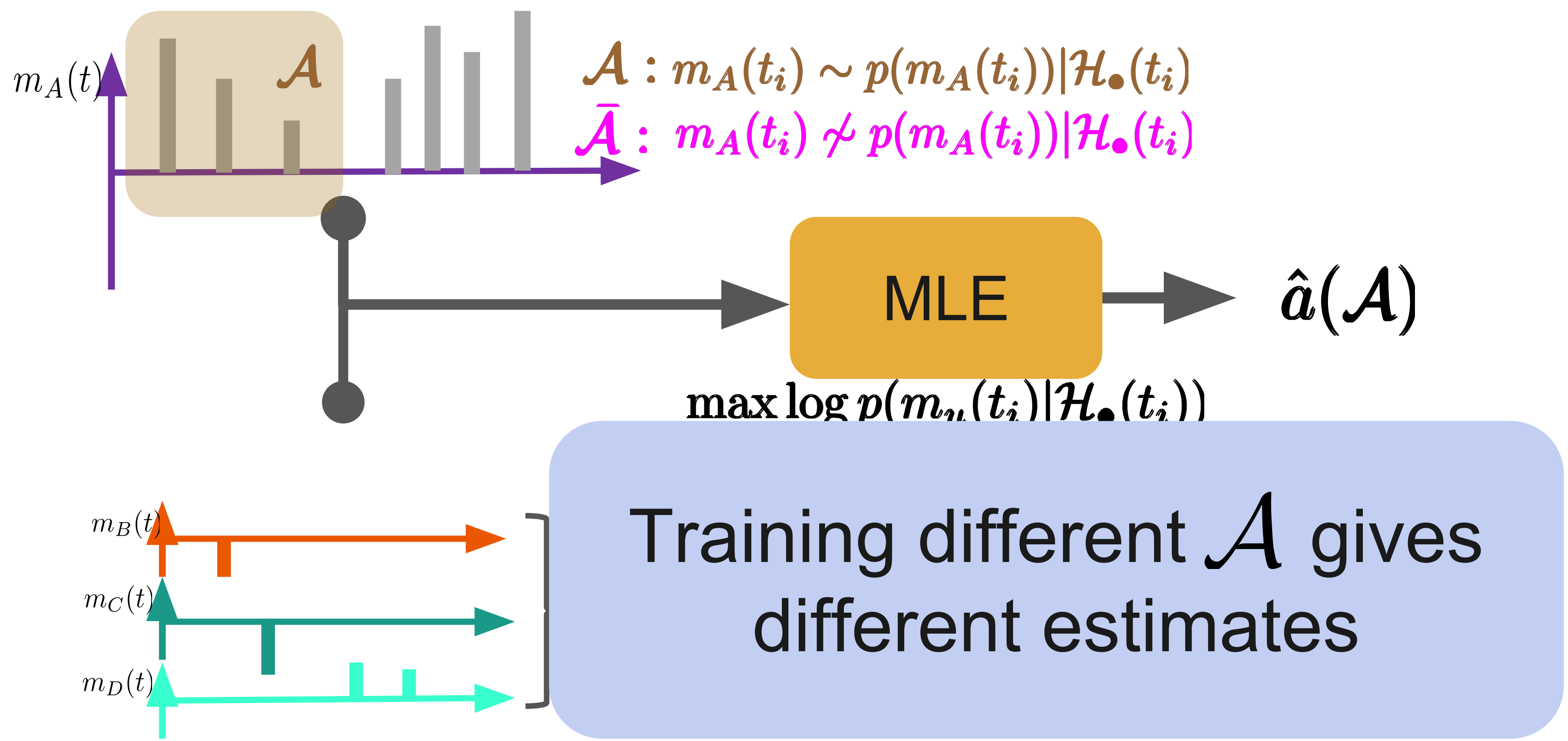
Can we use this simple fact to demarcate the endogenous and exogenous messages

Setup



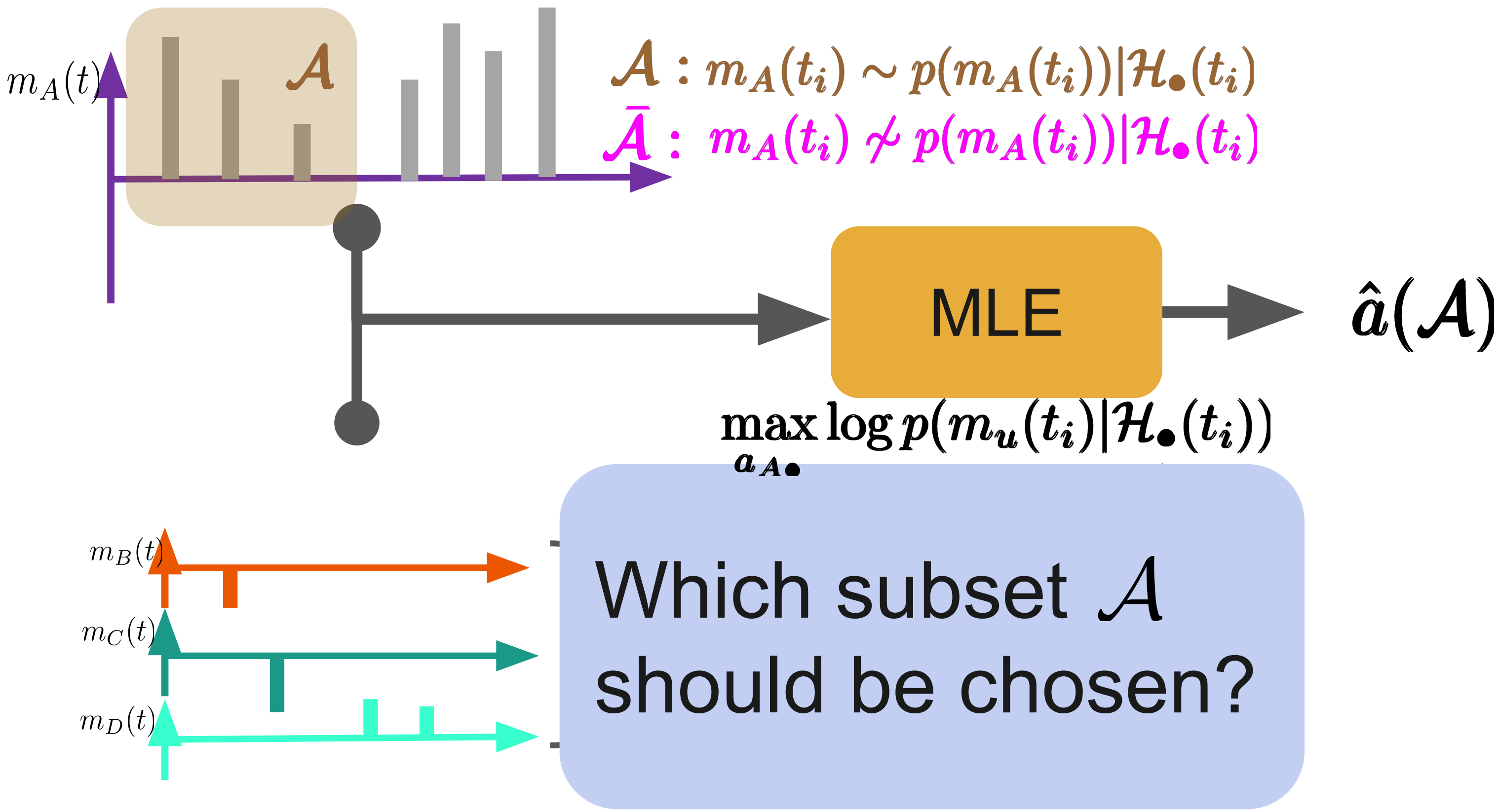
$m_A(t)$
$\mathcal{A} : m_A(t_i) \sim p(m_A(t_i))|\mathcal{H}_\bullet(t_i)$
$\bar{\mathcal{A}} : m_A(t_i) \not\sim p(m_A(t_i))|\mathcal{H}_\bullet(t_i)$
MLE
$\hat{a}_{AB}, \hat{a}_{AC}, \hat{a}_{AD}$
$\max_{a_{A\bullet}} \log p(m_u(t_i)|\mathcal{H}_\bullet(t_i))$
$m_B(t)$
$m_C(t)$
$m_D(t)$

Setup

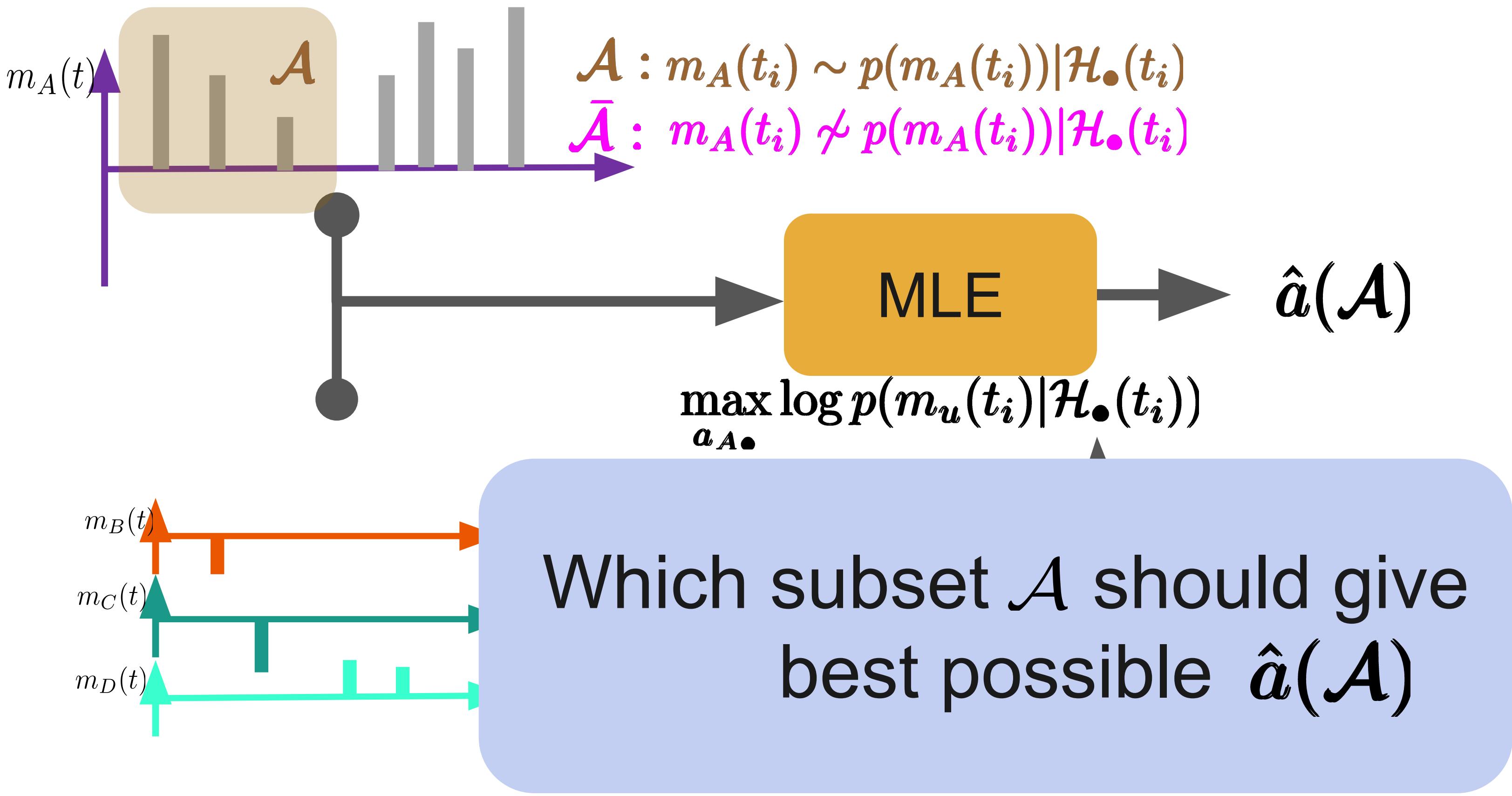


$m_A(t)$
$\mathcal{A}$
$\mathcal{A} : m_A(t_i) \sim p(m_A(t_i))|\mathcal{H}_\bullet(t_i)$
$\bar{\mathcal{A}} : m_A(t_i) \nsim p(m_A(t_i))|\mathcal{H}_\bullet(t_i)$
MLE
$\hat{a}(\mathcal{A})$
$\max \log p(m_u(t_i)|\mathcal{H}_\bullet(t_i))$
$m_B(t)$
$m_C(t)$
$m_D(t)$
Training different $\mathcal{A}$ gives different estimates

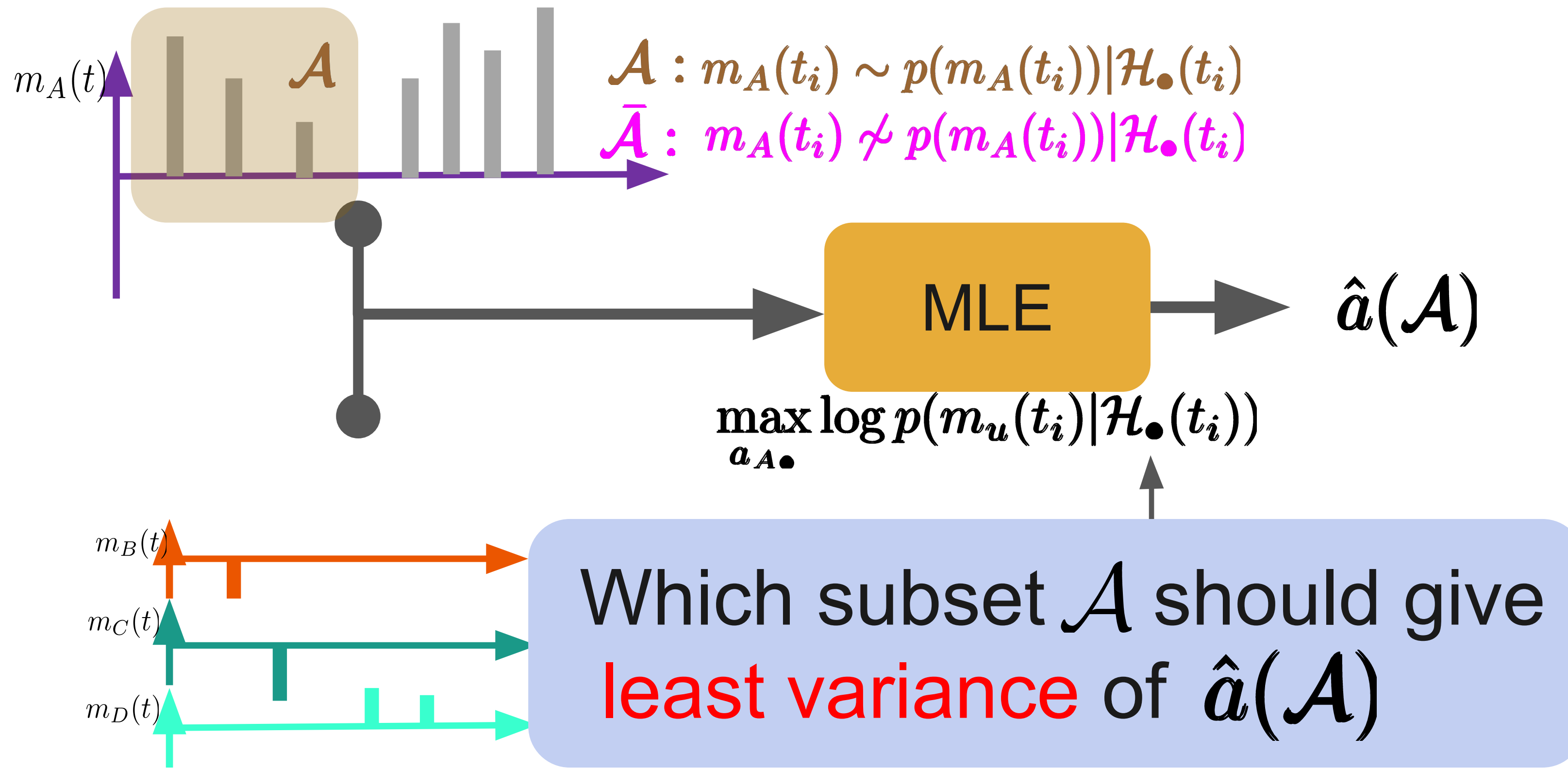
The core problem



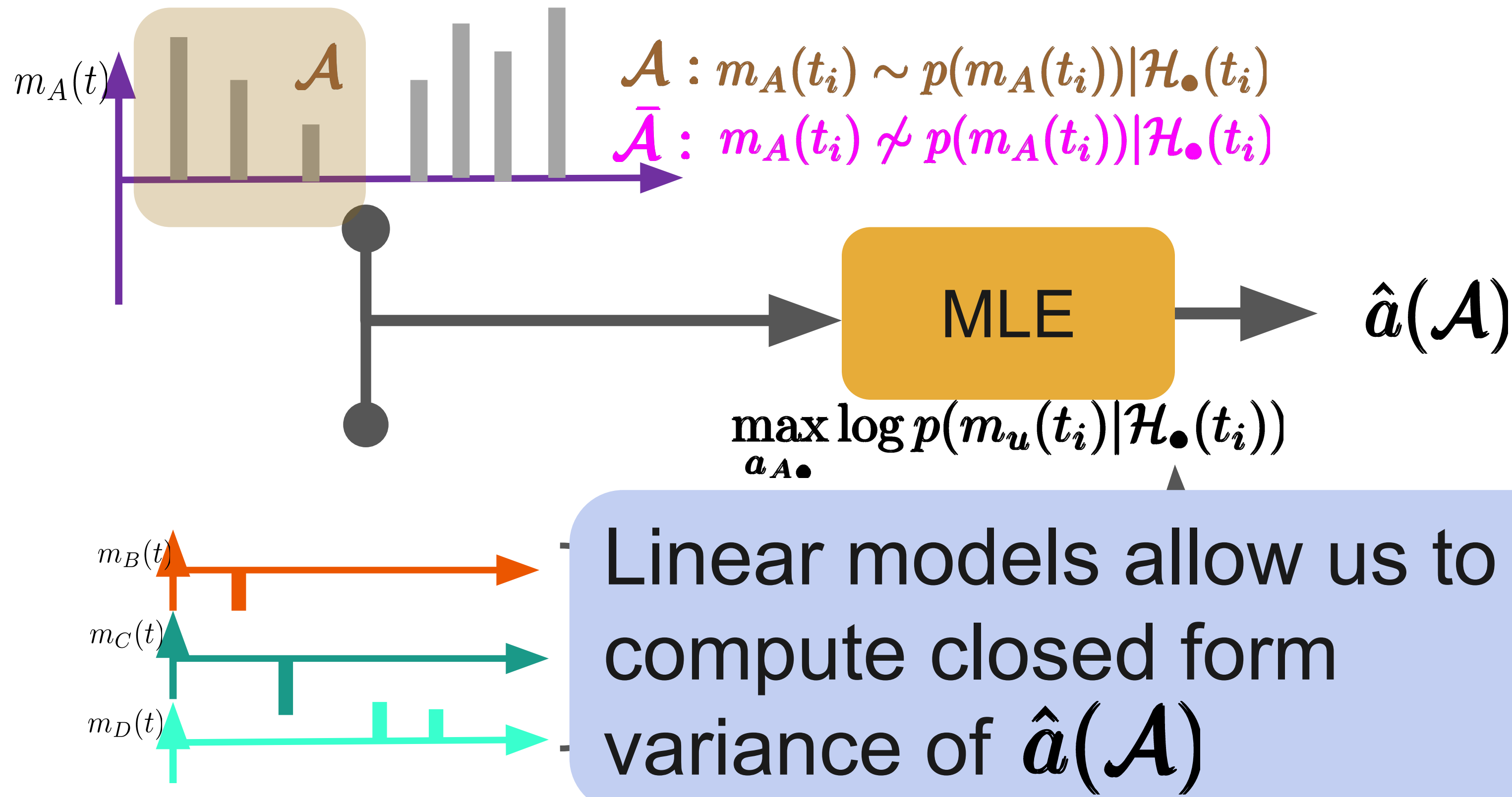
The core problem



The core problem



The core problem



Linear models allow us to compute closed form variance of $\hat{a}(\mathcal{A})$

CherryPick

$$\max_{\mathcal{A}} -\operatorname{tr}\log \boldsymbol{\Sigma}(\mathcal{A})$$

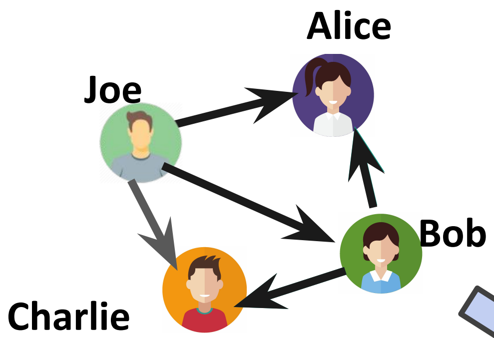
Submodular in $\mathcal{A}$

CherryPick

Greedy approach gives approximation guarantee

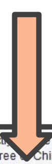
Submodular in $\mathcal{A}$

Organic and Extrinsic Users

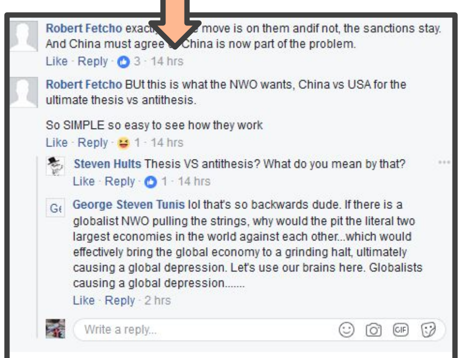


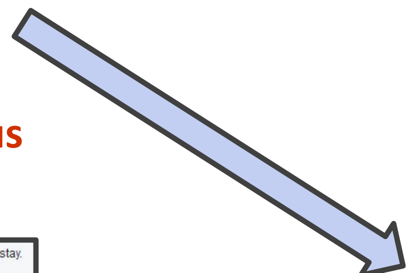
Extrinsic: Influenced more by externalities

Organic: Influenced more by other users



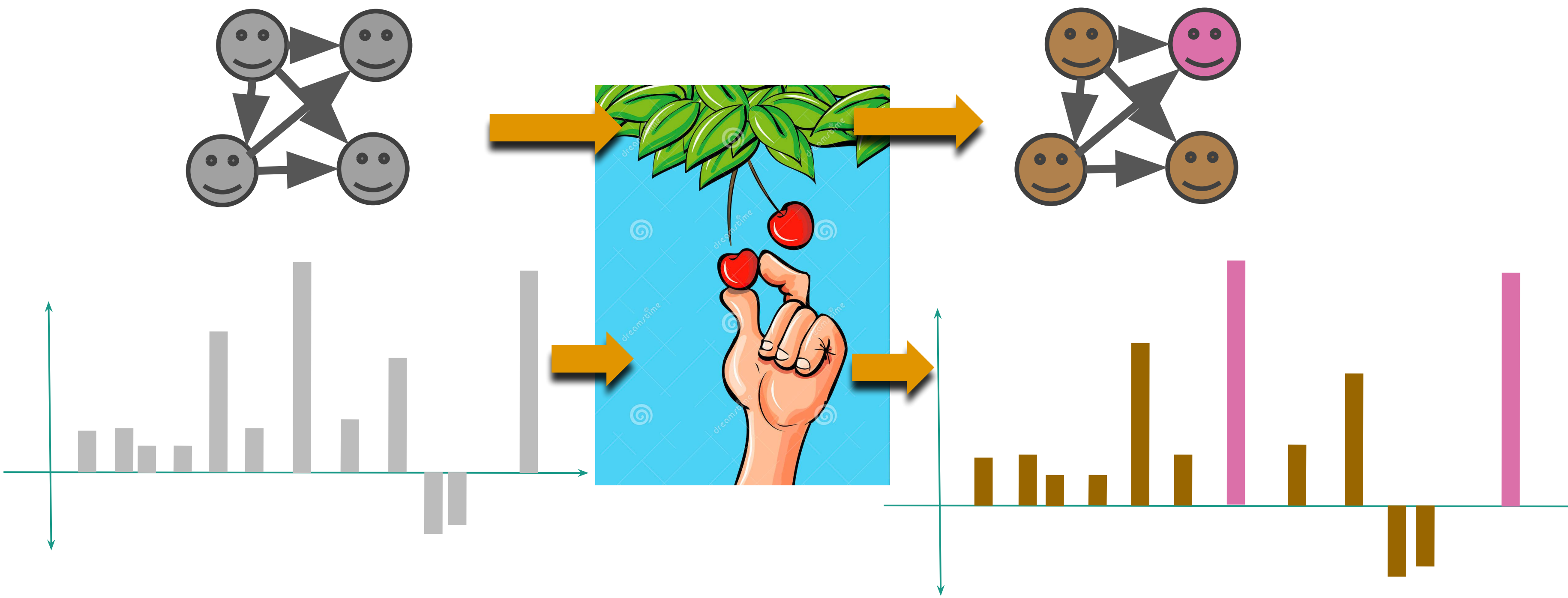
Endogenous opinions





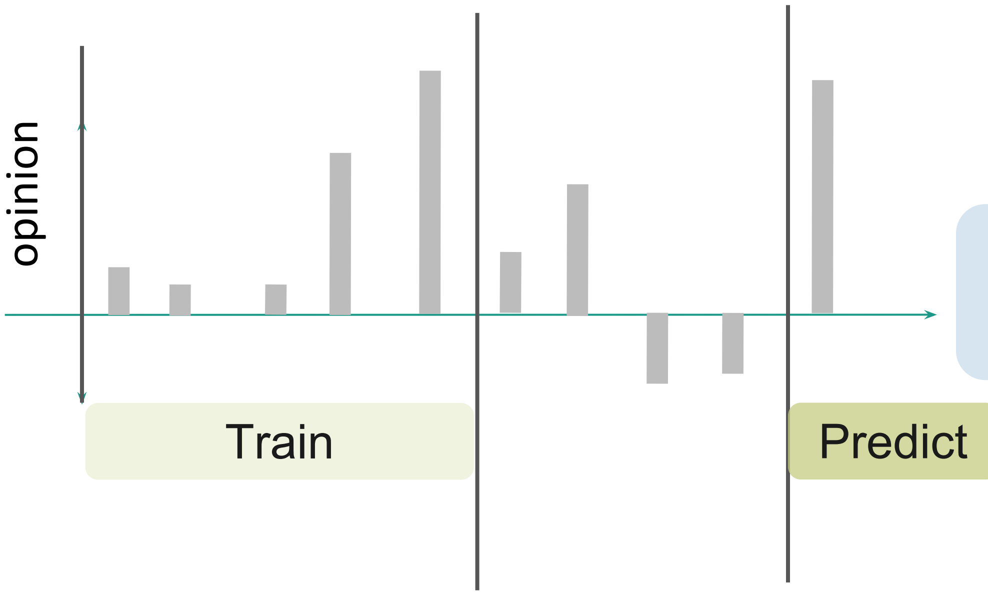
Exogenous opinions

Summary of the approach..



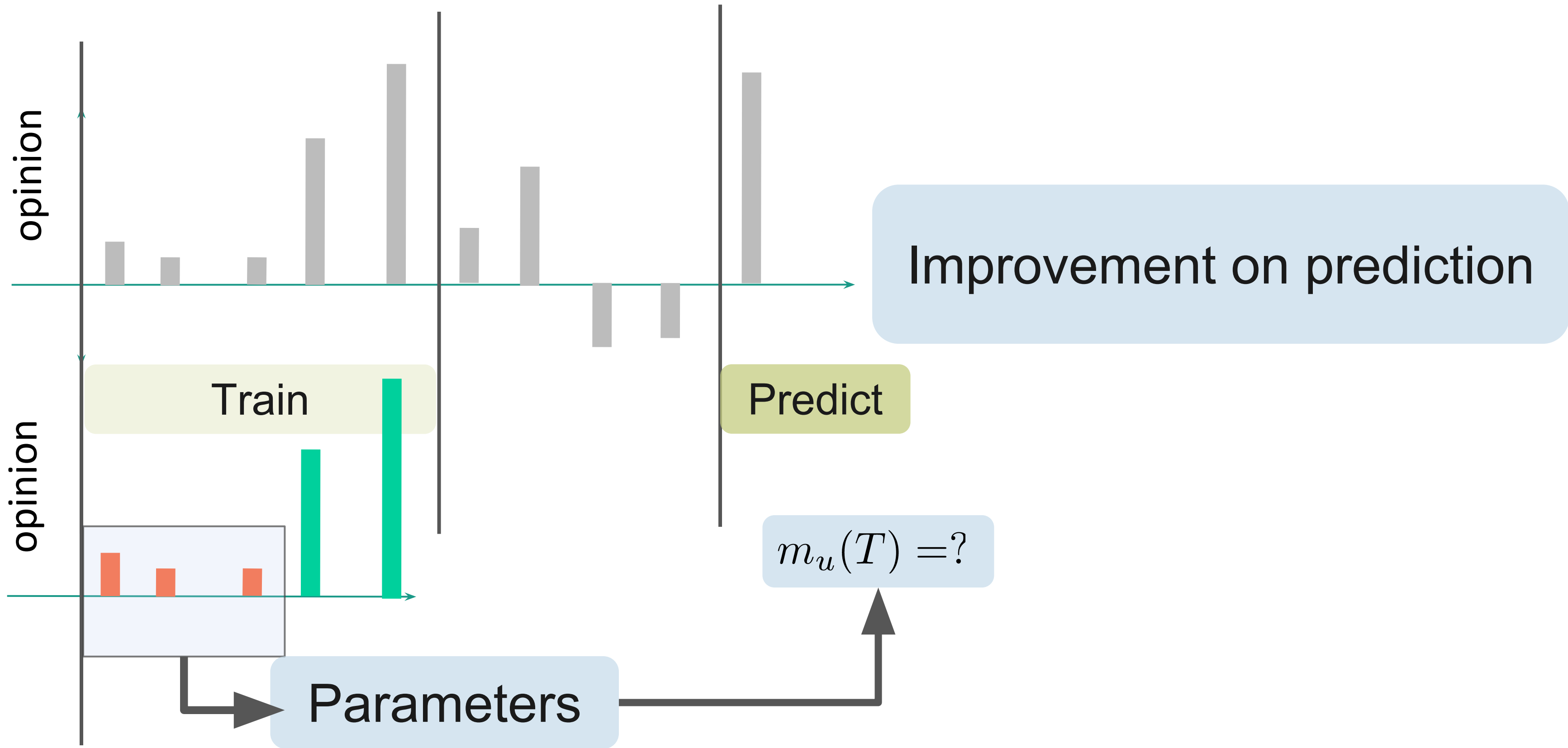
Experiments

Evaluation protocol



Improvement on prediction

Evaluation protocol



Datasets

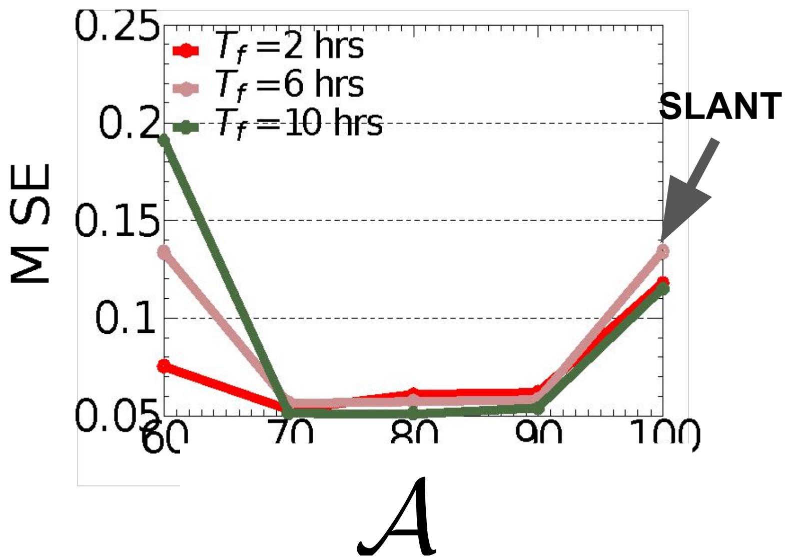
twitter

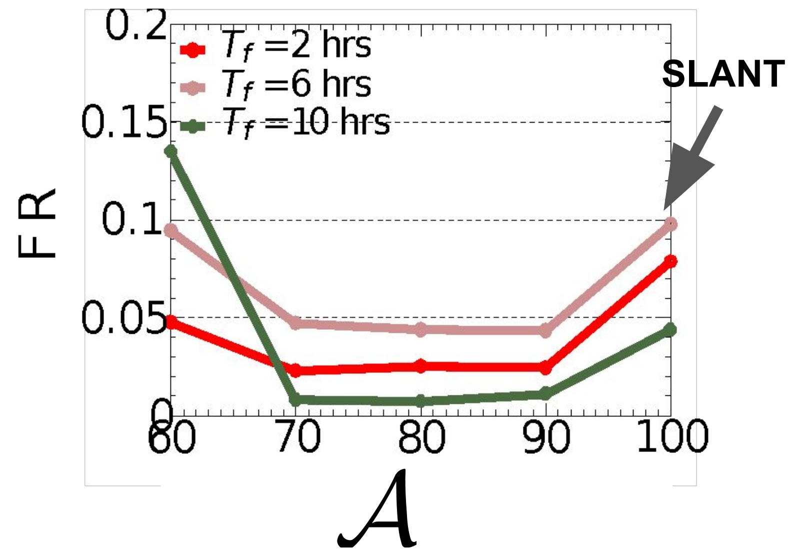
- Bollywood
- Sports
- Series
- Verdict
- US elections

Comparative Analysis

	MSE: $\mathbb{E}(m-\hat{m})^2$		FR: $\mathbb{P}(\text{sign}(m) \neq \text{sign}(\hat{m}))$	
	CherryPick	Slant	CherryPick	Slant
Elections	0.146	0.193	0.073	0.098
Series	0.110	0.213	0.097	0.125
Verdict	0.060	0.090	0.057	0.073

Impact of pre-specified size on $\mathcal{A}$





Verdict datasets

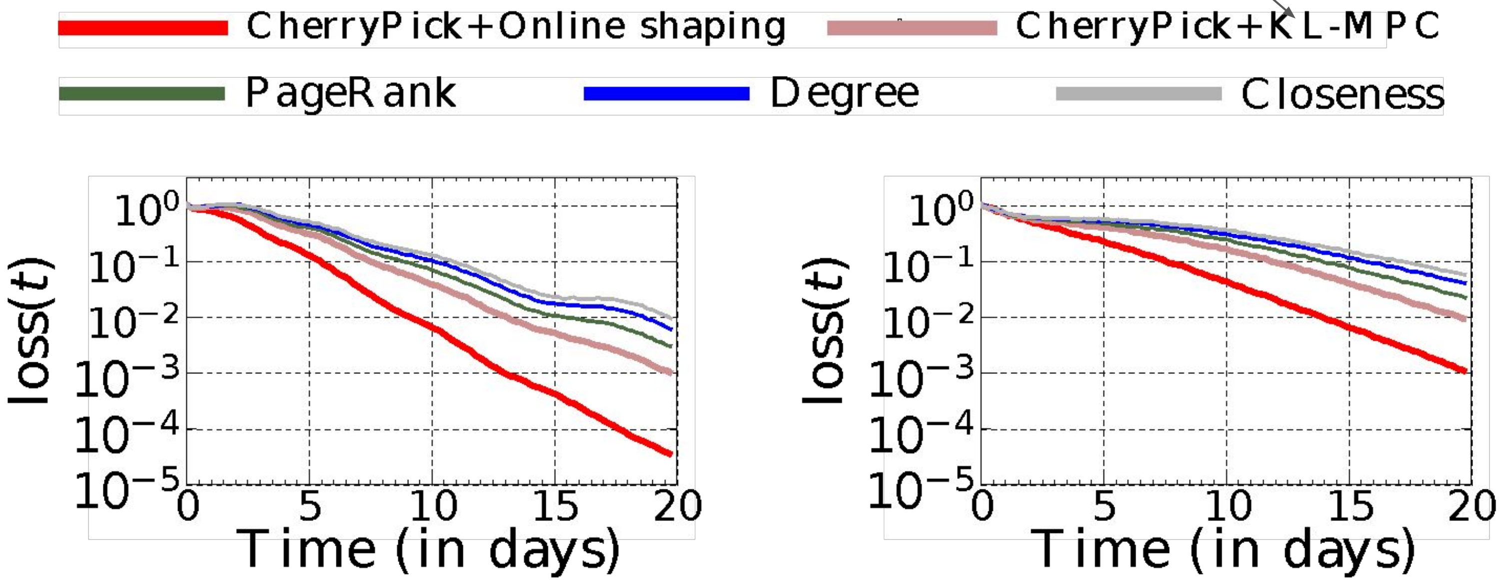
Examples from US Election dataset

Exogenous	#BREAKING: [..] campaign manager will not be prosecuted. A total joke.
Endogenous	Sad! [..], quit your whining
Endogenous	[..] campaign manager won’t be prosecuted. Such a huge disaster.
Exogenous	[…] shows farsightedness by shaking hands with Russia..

t

Effect of extrinsic user identification

Wang et al. ICML 2017



(a) Bollywood

(b) Series

Conclusion

- Models for information and opinion dynamics
- Capture the external effects
- Control the opinion dynamics

Complex Network Research Group (CNReG)

iitkgpcnerg

@cnerg

www.cnerg.org

Shaping Opinion Dynamics in Social Networks

AAMAS

18

Shaping opinion dynamics

Can a small set of incentivized users shape other users' opinions?

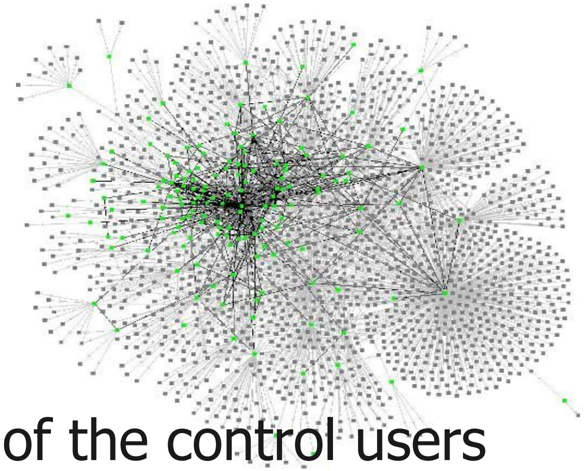
Why this goal?

Brand management

Awareness on orchestrated campaigns

Our goal

- Steady state opinion control
  - Find optimal incentivized users
  - Compute the optimal message rate of the control users
  - Robust opinion shaping



Steady state shaping

- $$\lim_{t \to \infty} \mathbb{E}_{\mathcal{H}_t}[\boldsymbol{x}^*_{\mathcal{I}^c}(t)]$$

-- the steady state opinion of “non incentivized” user $\mathcal{I}^c$

- Select the control nodes $\mathcal{I}$

Differential dynamics

Influence among non-control users

$$\frac{d\mathbb{E}_{\mathcal{H}_t}[\boldsymbol{x}^*_{\mathcal{I}^c}(t)]}{dt} = [-\omega I + \boldsymbol{A}_1 \boldsymbol{\Lambda}_{\mathcal{I}^c}(t)]\mathbb{E}_{\mathcal{H}_t}[\boldsymbol{x}^*_{\mathcal{I}^c}(t)] + \boldsymbol{A}_2(\boldsymbol{\eta}^+_{\mathcal{I}} - \boldsymbol{\eta}^-_{\mathcal{I}}) + \omega \boldsymbol{\alpha}_{\mathcal{I}^c},$$

Influence of control users on the network

Message rates for positive and negative opinions

Steady state opinion

Influence among non-control users

Influence of control users on the network

$$(\boldsymbol{A}_1\boldsymbol{\Lambda}_1 - \omega I)\mathbf{x}_{\mathcal{I}^c} + \boldsymbol{A}_2\left(\boldsymbol{\eta}_{\mathcal{I}}^{+} - \boldsymbol{\eta}_{\mathcal{I}}^{-}\right) + \omega\boldsymbol{\alpha}_{\mathcal{I}^c} = 0$$

where $\boldsymbol{x}_{\mathcal{I}^c} = \lim_{t\to\infty} \mathbb{E}_{\mathcal{H}_t}[\boldsymbol{x}^*_{\mathcal{I}^c}(t)]$.

Message rates for positive and negative opinions

In terms of “unknown” incentivized users

$$\mathcal{S} = 1[u \in \mathcal{I}]_{u \in \mathcal{V}}$$

One hot representation of Incentivized users

$$(\boldsymbol{A}_1 \boldsymbol{\Lambda}_1 - \omega I)\mathbf{x}_{\mathcal{I}^c} + \boldsymbol{A}_2 \left(\boldsymbol{\eta}_{\mathcal{I}}^+ - \boldsymbol{\eta}_{\mathcal{I}}^-\right) + \omega \boldsymbol{\alpha}_{\mathcal{I}^c} = 0$$

$$(\boldsymbol{A}\boldsymbol{\Lambda} - \omega I)[(\mathbf{1} - \boldsymbol{\mathcal{S}}) \odot \boldsymbol{x}] + \boldsymbol{A}[\boldsymbol{\mathcal{S}} \odot (\boldsymbol{\eta}^+ - \boldsymbol{\eta}^-)] + \omega(\mathbf{1} - \boldsymbol{\mathcal{S}}) \odot \boldsymbol{\alpha} = 0$$

Activity shaping problem

Once we know that $\mathbf{x}_{\mathcal{I}^c} = \left(I - \frac{\mathbf{A}_1 \mathbf{\Lambda}_1}{\omega}\right)^{-1} \left(\boldsymbol{\alpha}_{\mathcal{I}^c} + q\boldsymbol{A}_2 \left(\boldsymbol{\mu}_{\mathcal{I}}^{+} - \boldsymbol{\mu}_{\mathcal{I}}^{-}\right)\right)$

we can find $\boldsymbol{\mu}^{+}$ and $\boldsymbol{\mu}^{-}$ to satisfy **many different goals**:

OPINIO

We can solve this problem **efficiently** for **a large family of utilities!**

maxi

subjec $\quad + \omega \boldsymbol{\alpha}_{\mathcal{I}^c} = 0$

$\boldsymbol{c}^{T} \left(\boldsymbol{\mu}_{\mathcal{I}}^{+} + \boldsymbol{\mu}_{\mathcal{I}}^{-}\right) \leq C$

$\boldsymbol{\mu}_{\mathcal{I}}^{+} \geq 0 \quad \boldsymbol{\mu}_{\mathcal{I}}^{-} \geq 0$

Cost for incentivizing **Budget**

SmartShape-Basic

$$\underset{\boldsymbol{x},\, \boldsymbol{\mathcal{S}},\, \boldsymbol{\eta}^{\pm}}{\text{maximize}} \quad U((\mathbf{1} - \boldsymbol{\mathcal{S}}) \odot \boldsymbol{x})$$

So that,

$$(\boldsymbol{A}\boldsymbol{\Lambda} - \omega I)[(\mathbf{1} - \boldsymbol{\mathcal{S}}) \odot \boldsymbol{x}] + \boldsymbol{A}[\boldsymbol{\mathcal{S}} \odot (\boldsymbol{\eta}^{+} - \boldsymbol{\eta}^{-})] + \omega(\mathbf{1} - \boldsymbol{\mathcal{S}}) \odot \boldsymbol{\alpha} = 0$$

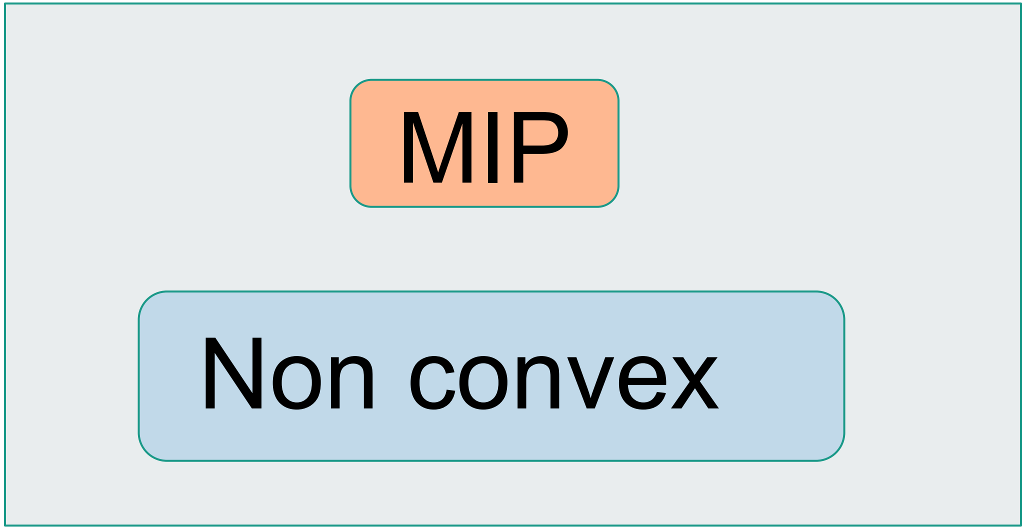
Steady state characterization

$$\boldsymbol{c}^{T}\,(\boldsymbol{\eta}^{+} + \boldsymbol{\eta}^{-}) \odot \boldsymbol{\mathcal{S}} \leq C, \text{and } \boldsymbol{\eta}^{\pm} \geq 0.$$

Budget constraint

Well posed rate

SmartShape-Basic



SmartShape-Basic

Linearization of variables

Linearization of variables

$$\texttt{\% Linearization of } (1-\mathcal{S}) \odot \boldsymbol{x}$$

$$\underline{\boldsymbol{x}} \leq \boldsymbol{z} \leq \overline{\boldsymbol{x}}$$

$$\underline{\boldsymbol{x}} \odot (\mathbf{1}-\boldsymbol{\mathcal{S}}) \leq \boldsymbol{z} \leq \overline{\boldsymbol{x}} \odot (\mathbf{1}-\boldsymbol{\mathcal{S}})$$

$$\boldsymbol{x}-\boldsymbol{\mathcal{S}} \odot \overline{\boldsymbol{x}} \leq \boldsymbol{z} \leq \boldsymbol{x}-\boldsymbol{\mathcal{S}} \odot \underline{\boldsymbol{x}}$$

$$\boldsymbol{z} \leq \boldsymbol{x}+\boldsymbol{\mathcal{S}} \odot \overline{\boldsymbol{x}}$$

$$\texttt{\% Linearization of } \mathcal{S} \odot \eta^{\pm}$$

$$\boldsymbol{c} \odot \boldsymbol{\xi}^{\pm} \leq C\boldsymbol{\mathcal{S}}$$

$$0 \leq \boldsymbol{\xi}^{\pm} \leq \boldsymbol{\eta}^{\pm}$$

$$\boldsymbol{c} \odot \boldsymbol{\eta}^{\pm}-(\mathbf{1}-\boldsymbol{\mathcal{S}})C \leq \boldsymbol{c} \odot \boldsymbol{\xi}^{\pm}$$

$$0 \leq \boldsymbol{\mathcal{S}} \leq 1$$

Francisco E Torres. Linearization of mixed-integer products. *Mathematical programming*, 49(1):427–428, 1990

SmartShape-Basic

$$\underset{z,x,\eta^{\pm},\xi^{\pm},\mathcal{S}}{\textbf{maximize}} \quad U(z)$$

$$\text{subject to:} \quad (\boldsymbol{A\Lambda} - \omega I)\boldsymbol{z} + \boldsymbol{A}\left(\boldsymbol{\xi}^{+} - \boldsymbol{\xi}^{-}\right)] + \omega(\mathbf{1} - \boldsymbol{\mathcal{S}}) \odot \boldsymbol{\alpha} = 0$$

$$\mathbf{c}^{T}\left(\boldsymbol{\xi}^{+} + \boldsymbol{\xi}^{-}\right) \leq C, \;\; \boldsymbol{\xi}^{+} \geq 0 \quad \boldsymbol{\xi}^{-} \geq 0$$

% Linearization of $(1 - \mathcal{S}) \odot \boldsymbol{x}$

$$\underline{\boldsymbol{x}} \leq \boldsymbol{z} \leq \overline{\boldsymbol{x}}$$

$$\underline{\boldsymbol{x}} \odot (\mathbf{1} - \boldsymbol{\mathcal{S}}) \leq \boldsymbol{z} \leq \overline{\boldsymbol{x}} \odot (\mathbf{1} - \boldsymbol{\mathcal{S}})$$

$$\boldsymbol{x} - \boldsymbol{\mathcal{S}} \odot \overline{\boldsymbol{x}} \leq \boldsymbol{z} \leq \boldsymbol{x} - \boldsymbol{\mathcal{S}} \odot \underline{\boldsymbol{x}}$$

$$\boldsymbol{z} \leq \boldsymbol{x} + \boldsymbol{\mathcal{S}} \odot \overline{\boldsymbol{x}}$$

% Linearization of $\mathcal{S} \odot \eta^{\pm}$

$$\boldsymbol{c} \odot \boldsymbol{\xi}^{\pm} \leq C\boldsymbol{\mathcal{S}}$$

$$0 \leq \boldsymbol{\xi}^{\pm} \leq \boldsymbol{\eta}^{\pm}$$

$$\boldsymbol{c} \odot \boldsymbol{\eta}^{\pm} - (\mathbf{1} - \boldsymbol{\mathcal{S}})C \leq \boldsymbol{c} \odot \boldsymbol{\xi}^{\pm}$$

$$0 \leq \boldsymbol{\mathcal{S}} \leq 1$$

SmartShape-Robust

$$\underset{\boldsymbol{z},\boldsymbol{x},\boldsymbol{\eta}^{\pm},\boldsymbol{\xi}^{\pm},\mathcal{S},\Gamma_{\boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{S}}}}{\textbf{maximize}} \qquad U(\boldsymbol{z}) - \gamma \operatorname{Tr}(\boldsymbol{\Gamma}_{\boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{S}}})$$

Examples of utility U(.)

MMOSH-1: $$U(\mathbf{x}_{\mathcal{I}^c}) = -\max_{u\in[m]} x_{\mathcal{I}^c,u}$$

MMOSH-2: $$U(\mathbf{x}_{\mathcal{I}^c}) = \min_{u\in[m]} x_{\mathcal{I}^c,u}.$$

AOSH-1: $$U(\mathbf{x}_{\mathcal{I}^c}) = -\sum x_u.$$

AOSH-2: $$U(\mathbf{x}_{\mathcal{I}^c}) = \sum_{u\in\mathcal{I}^c} x_u.$$

Top-k-opinion-shaping

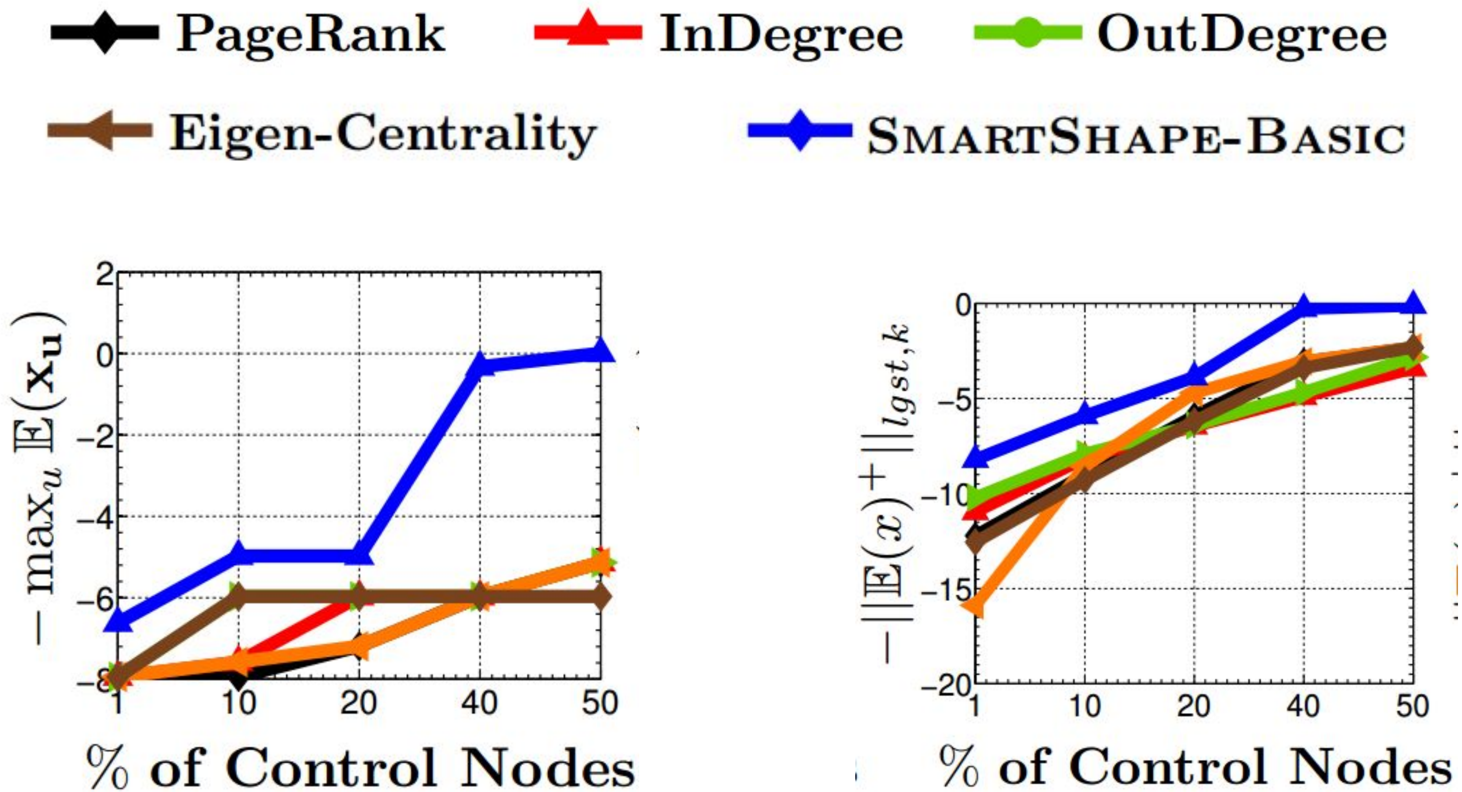
Top-k—OSH-{1,2}: $$U(\mathbf{y}_{\mathcal{I}^c}) = -\sum_{i=1}^{k} |y_{[i]}|,$$

Top-k—OSH-1: $y_{\mathcal{I}^c,u} \geq \max(x_{\mathcal{I}^c,u}, 0) \quad \forall u \in \mathcal{I}^c$

Top-k—OSH-2: $y_{\mathcal{I}^c,u} \leq \min(x_{\mathcal{I}^c,u}, 0) \quad \forall u \in \mathcal{I}^c$

$y_{[i]}$ = ***i-th*** largest component of ***y***

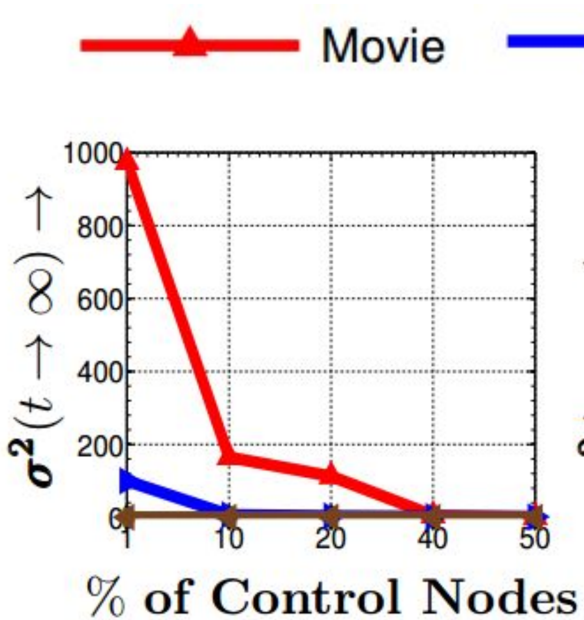
Experiments on Politics dataset



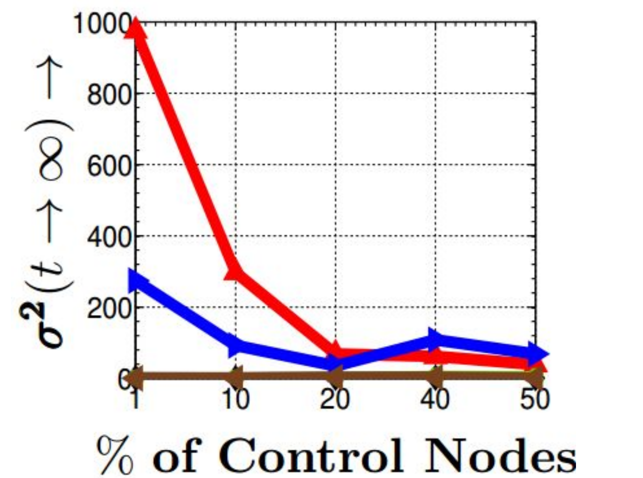
MMOSH-1

Top-k-OSH-1

Experiments on SmartShape-Robust



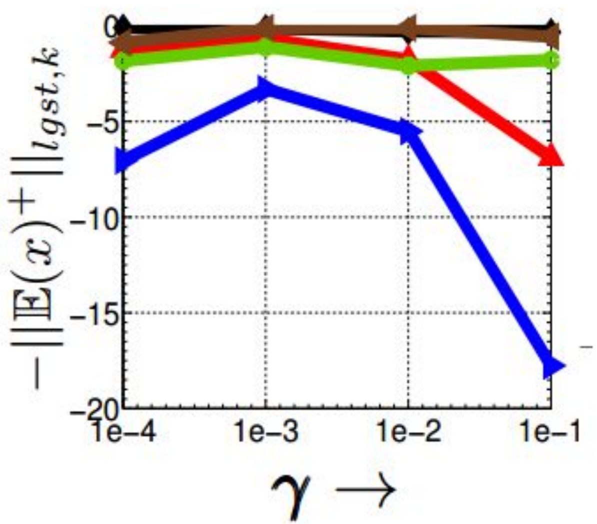
MMOSH-1

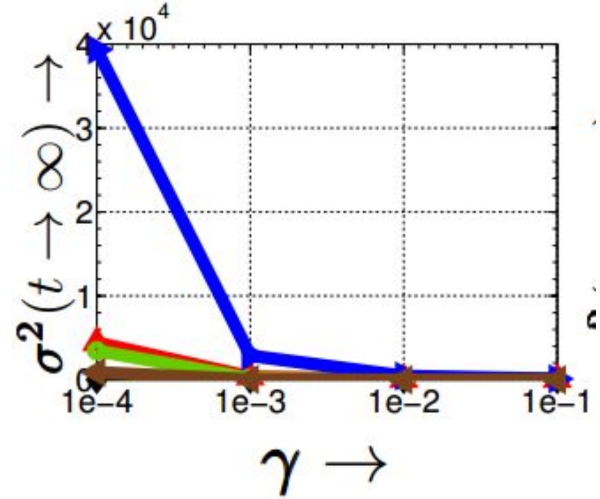


Top-k-OSH-1

Experiments on SmartShape-Robust

Top-k-OSH-1



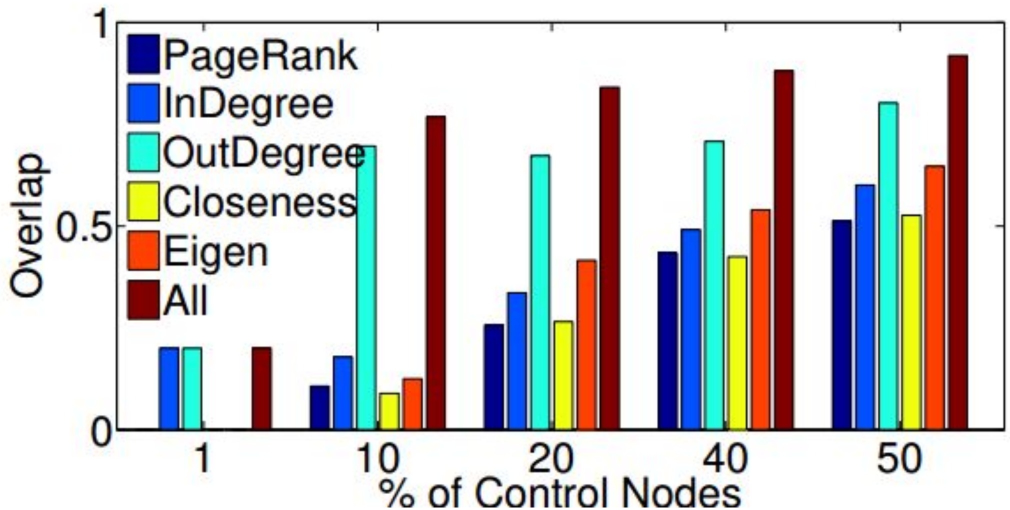


Objective variation

Variance variation

Characterization of incentivized users

Movie dataset



Overlap= Jaccard Coefficient

Conclusion

- Models for information and opinion dynamics
- Capture the external effects
- Control the opinion dynamics

Few points to work in future

- How to capture competition via an deep adversarial network?
- Opinion regulation through deep learning
- Demarcation with partial labels

Complex Network Research Group (CNReG)

iitkgpcnerg

@cnerg

www.cnerg.org

Thank You

Accepted papers

- Bidisha Samanta, Abir De, Abhijnan Chakrabarti, Niloy Ganguly, LMPP: A Large Margin Point Process Combining Reinforcement and Competition for Modeling Hashtag Popularity, **IJCAI**, Melbourne, August 2017
- Bidisha Samanta, Abir De, Niloy Ganguly, STRM: A Sister Tweet Reinforcement Process for Modeling Hashtag Popularity, **Infocom'17**, Atlanta, GA.
- Abir De, Isabel Valera, Niloy Ganguly, Sourangshu Bhattacharya, Manuel Gomez Rodrigue Learning and Forecasting Opinion Dynamics in Social Network **NIPS'16,** Barcelona, Spain
- Abir De, Sourangshu Bhattacharya, Parantapa Bhattacharya, Niloy Ganguly, Soumen Chakrabarti: Learning a Linear Influence Model from Transient Opinion Dynamics. **CIKM** 2014

Outline of the talk

Overview and motivation
- Background of information diffusion
- Challenges
- Introduction to point process

Hashtag Diffusion
- Key Ideas behind developing model
- Detailed Modeling
- Results and explanation

Opinion Dynamics
- Key Ideas behind developing model
- Detailed Modeling
- Results and explanation

Not only temporality but also informations hidden in text influences the propagation

Outline of the talk

- **Overview and motivation**
  - Background
  - Challenges
  - Introduction to point process
- **Hashtag Diffusion**
  - Key Ideas behind developing model
  - Detailed Modeling
  - Results and explanation
- **Opinion Dynamics**
  - Key Ideas behind developing model
  - Detailed Modeling
  - Results and explanation

Evaluation protocol

Forecasting hashtag popularity

- Learn the model parameters and forecast

Rank prediction of competing hashtags

- Predict popularity rankings of competing hashtags based on predicted count

Terminology

Competing hashtag

- Hashtags which belong to the same event
- Co-occur(start time lies within a window of few hours and a significant(60%) overlap of lifetime)

Metrics

Mean Absolute Percentage Error (MAPE)

$$\mathrm{MAPE}(\boldsymbol{H}) = \frac{1}{M_H} \sum_{i=0}^{M_H-1} \left| \frac{\hat{N}_H(t_i) - N_H(t_i)}{N_H(t_i)} \right|.$$

Forecasting

Rank Prediction

Spearman's Rank Correlation Coefficient (SRCC)

$$\rho(\hat{R}_{\mathbb{H}}, R_{\mathbb{H}}) = \frac{\mathbf{Cov}(\hat{R}_{\mathbb{H}}, R_{\mathbb{H}})}{\sqrt{\mathbf{Var}(\hat{R}_{\mathbb{H}})\mathbf{Var}(R_{\mathbb{H}})}}.$$

Avg. Precision & Avg. Recall of Jump Prediction

$$\text{Precision} = \frac{tp}{tp + fp} \qquad \text{Recall} = \frac{tp}{tp + fn}$$

Research problem

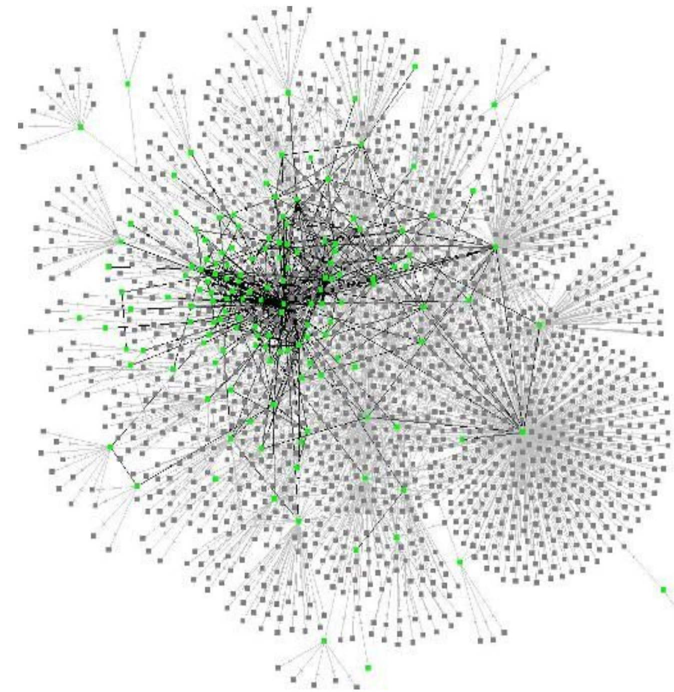
What we want to solve?

Can we **design a realistic model** that **fits real fine-grained opinion traces**?

Why we want to solve?

Predict (infer) opinions, even if not expressed!

What about the opinion and information propagation dynamics?



Complex stochastic Process
(Over a network)

Model Formulation → Simulation and Estimation → Real-data Experiment

Stubbornness, conformity, and compromise

Our model allows for:

Stubborn users

$$x^*_u(t) = \alpha_u$$

Compromised users

Conforming users

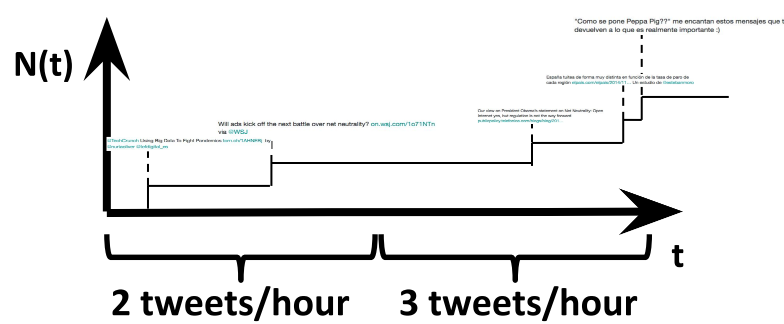
$$x^*_u(t) = \sum_{v \in \mathcal{N}(u)} a_{vu} \sum_{e_i \in \mathcal{H}_v(t)} m_i g(t - t_i)$$

Model Formulation → Simulation and Estimation → Real-data Experiment

Key idea: The counting process has **memory** and **is time-varying.**
Model message time as a counting process

$$\mathbb{E}[d\boldsymbol{N}(t)|\mathcal{H}_{t-}] = \boldsymbol{\lambda}(t)\,dt$$

- $d\boldsymbol{N}(t)$: Increase in # of messages (e.g., tweets) from t to t+dt
- $\mathcal{H}_{t-}$: History of messages up to t
- $\boldsymbol{\lambda}(t)$: Instantaneous message intensity (e.g., tweets / hour) → Instantaneous rate or frequency



Model Formulation | Simulation and Estimation | Real-data Experiment

Stubbornness, conformity, and compromise

Our model allows for:

Stubborn users

$$x_u^*(t) = \alpha_u$$

Compromised users

Conforming users

$$x_u^*(t) = \sum_{v \in \mathcal{N}(u)} a_{vu} \sum_{e_i \in \mathcal{H}_v(t)} m_i g(t - t_i)$$

Model Formulation

Simulation and Estimation

Real-data Experiment

Efficient model simulation

We adapt efficient sampling algorithm for multivariate Hawkes introduced by Farajtabar et al. (NIPS 2015). Two key ideas:

Markov property allows us to **update individual intensities and opinions in O(1)**.

Temporal and informational influence are sparse, whenever a node expresses its opinion, only a few opinions and intensity functions change.

Model Formulation → Simulation and Estimation → Real-data Experiment

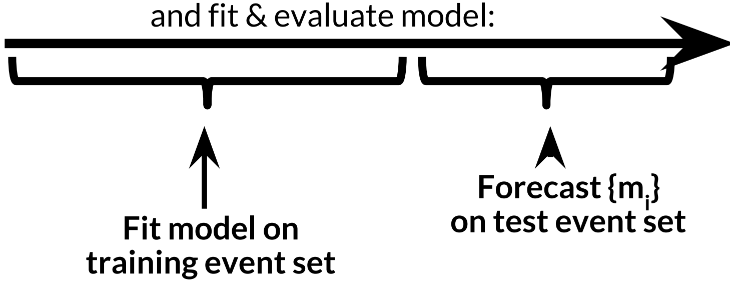
Evaluation of model on real data

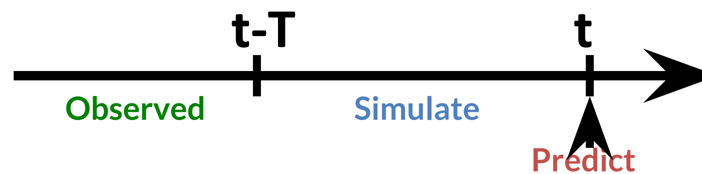
For each dataset, build collection of events $\{(t_i, m_i)\}$, where:

t_i: message times

mi: estimated sentiment from text (Hannak et al.)

and fit & evaluate model:





$$\hat{m} = E_{\mathcal{H}_t \setminus \mathcal{H}_{t-T}}[x_u^*(t) | \mathcal{H}_{t-T}]$$

Arrival rate formulation → Likelihood estimation → Model competition as constraint → Parameter estimation → Future arrival rate estimation

Find the optimal values of the parameters using **Joint Maximum Likelihood Estimation** subjected to a set of constraints

Slack Variables

$$\max_{\boldsymbol{\lambda}_{\mathbb{H},0},\, .\beta} \log[L(\boldsymbol{\lambda}_{\mathbb{H},0},\, .\beta \,|\, \epsilon, \boldsymbol{\omega}, \omega)] - C \sum_{i=0}^{L-1} \sum_{\boldsymbol{H},\boldsymbol{H}' \in \mathbb{H}} \zeta^{i}_{\boldsymbol{H},\boldsymbol{H}'}$$

$$\text{with, } y^{i}_{\boldsymbol{H},\boldsymbol{H}'} \int_{iT_s}^{(i+1)T_s} \Big(\lambda_{\boldsymbol{H}}(t) - \lambda_{\boldsymbol{H}'}(t)\Big) dt \geq 1 - \zeta^{i}_{\boldsymbol{H},\boldsymbol{H}'}$$

The MLE problem is convex and solvable easily

Arrival rate formulation → Likelihood estimation → Model competition as constraint → Parameter estimation → Future arrival rate estimation

Future expected rate is obtained by taking expectation on the popularity index K(t)

$$\tilde{\lambda}_{\boldsymbol{H}}(t) = \mathbb{E}_{\boldsymbol{k}}[\lambda_{\boldsymbol{H}}(t; \boldsymbol{k}(t))]$$

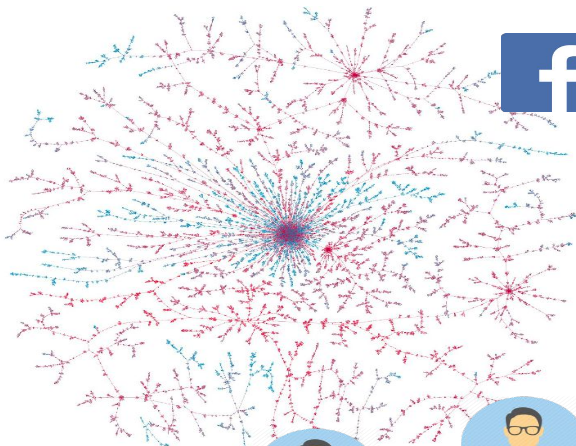
Future expected popularity of hashtag H

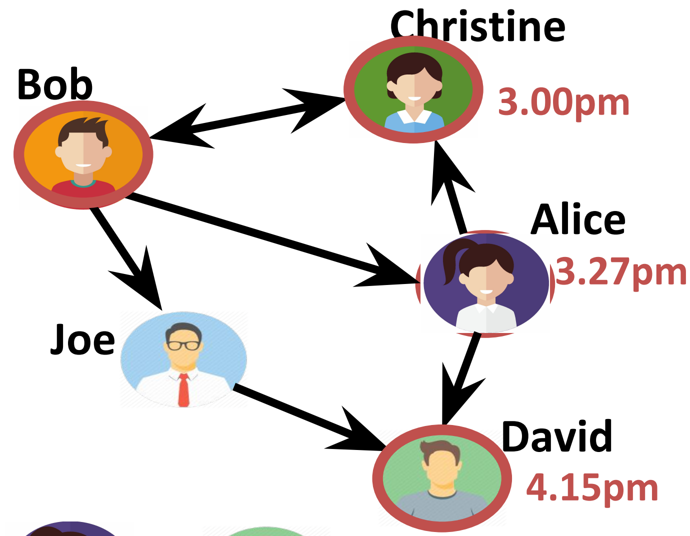
$$N_{\boldsymbol{H}}(t) = \int_0^t \tilde{\lambda}_{\boldsymbol{H}}(t)dt$$

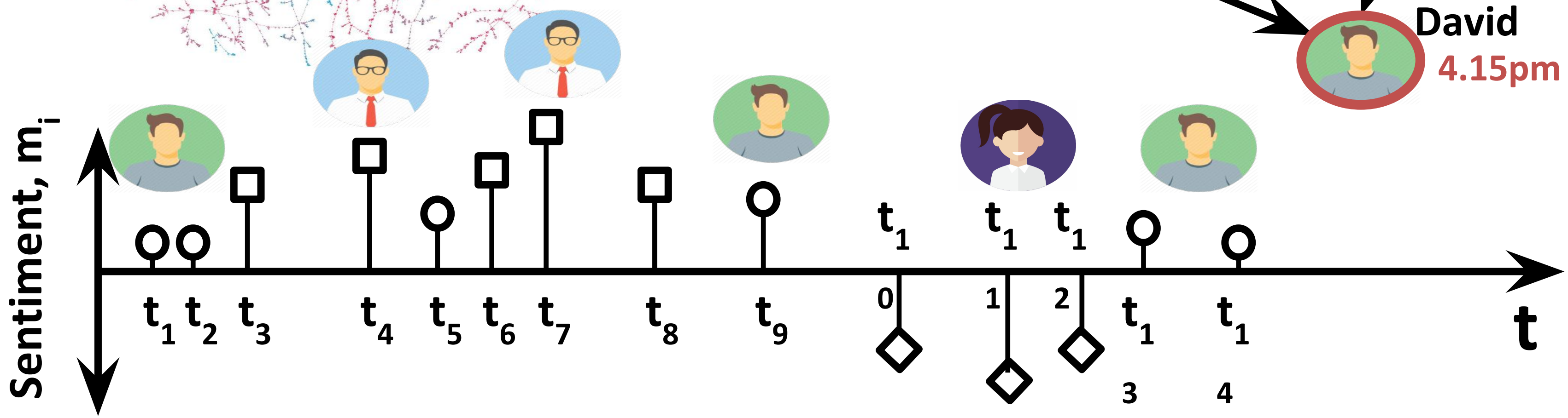
Results and discussions

- LMPP outperforms all the baselines in forecasting (MAPE : Mean Absolute Percentage Error) and rank prediction (SRCC: Spearman's Rank Correlation)

Example: Idea adoption or opinion influence

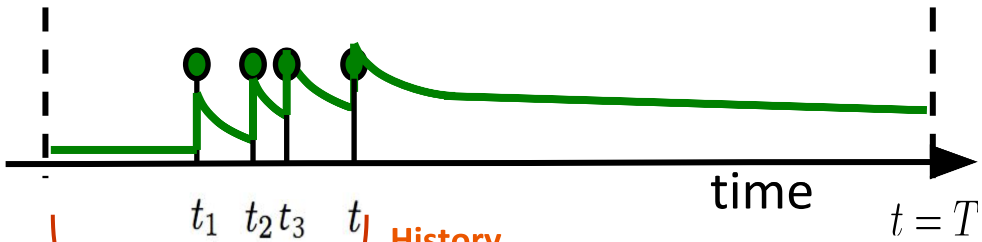






Arrival rate formulation → Likelihood estimation → Model competition as constraint

Triggering kernel formulation to capture the effect of the past events



$$\underbrace{t_1 \; t_2 \; t_3 \; t}_{\text{History}}$$

$$\kappa_k(t) = e^{-(\omega_0 + \frac{\omega}{k})t}$$

Popularity index

Arrival rate formulation (Hawkes process) captures the tweet reinforcement

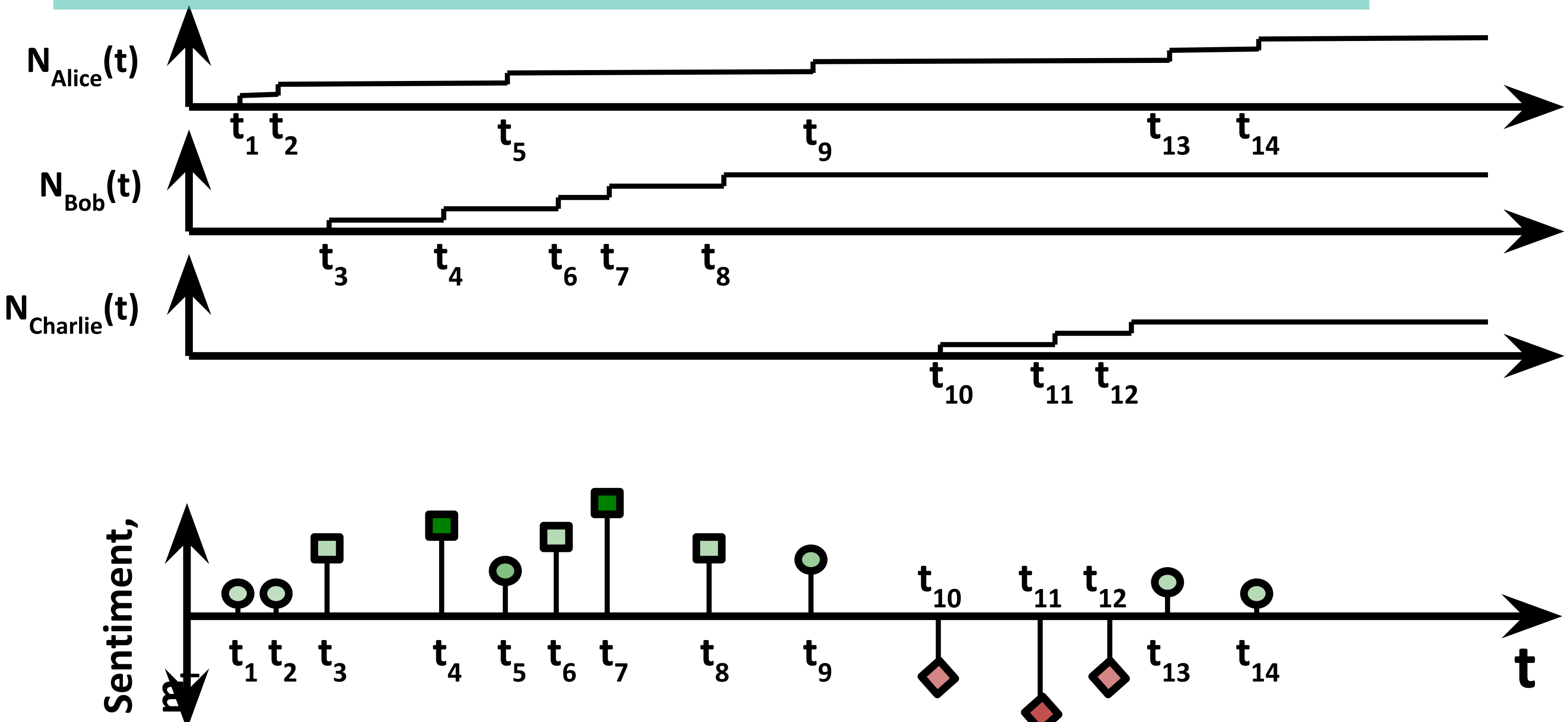
$$\lambda_{\boldsymbol{H}}(t; \boldsymbol{k}(t)) = \underbrace{\lambda_{\boldsymbol{H},0} e^{-\epsilon t}}_{\text{Background rate}} + \sum_{j=1}^{M} \beta^j_{\boldsymbol{H}} \sum_{t_i \in \mathcal{H}_{\boldsymbol{H}}(t)} e^{\underbrace{-(\omega_j + \frac{\omega}{k(t_i)})(t - t_i)}_{\text{Triggering kernel Effect of the previous history}}}$$

Model Formulation

Simulation and Estimation

Real-data Experiment

We model each user's messages as point processes $\{(t_i, m_i)\}$ with message sentiment



Exponential kernel
(memory)

$$\underbrace{\lambda_u^*(t)}_{\text{User's message intensity}} = \underbrace{\mu_u}_{\text{Messages on her own initiative}} + \sum_{v \in u \cup \mathcal{N}(u)} b_{vu} \underbrace{\sum_{e_i \in \mathcal{H}_v(t)} \kappa(t - t_i)}_{\text{Previous messages by user v}}$$

Temporal influence from user v on user u

Exponential kernel
(memory)

$$\underbrace{x_u^*(t)}_{\text{User's latent opinion}} = \underbrace{\alpha_u}_{\text{User's initial opinion}} + \sum_{v \in \mathcal{N}(u)} a_{vu} \underbrace{\sum_{e_i \in \mathcal{H}_v(t)} m_i g(t - t_i)}_{\text{Previous expressed opinions y user v}}$$

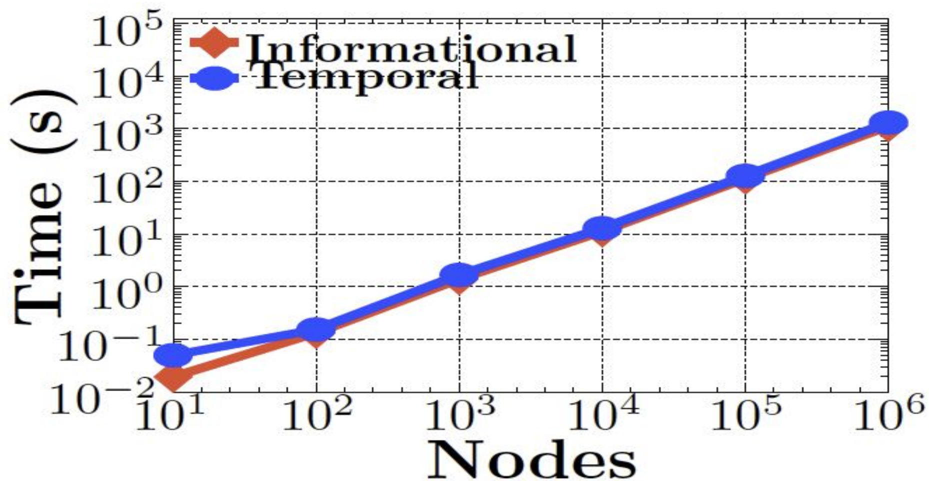
Informational influence from user v
on user u
Not captured previously

Model Formulation

Simulation and Estimation

Real-data Experiment

Parameter Estimation scalability



single machine with 24 cores

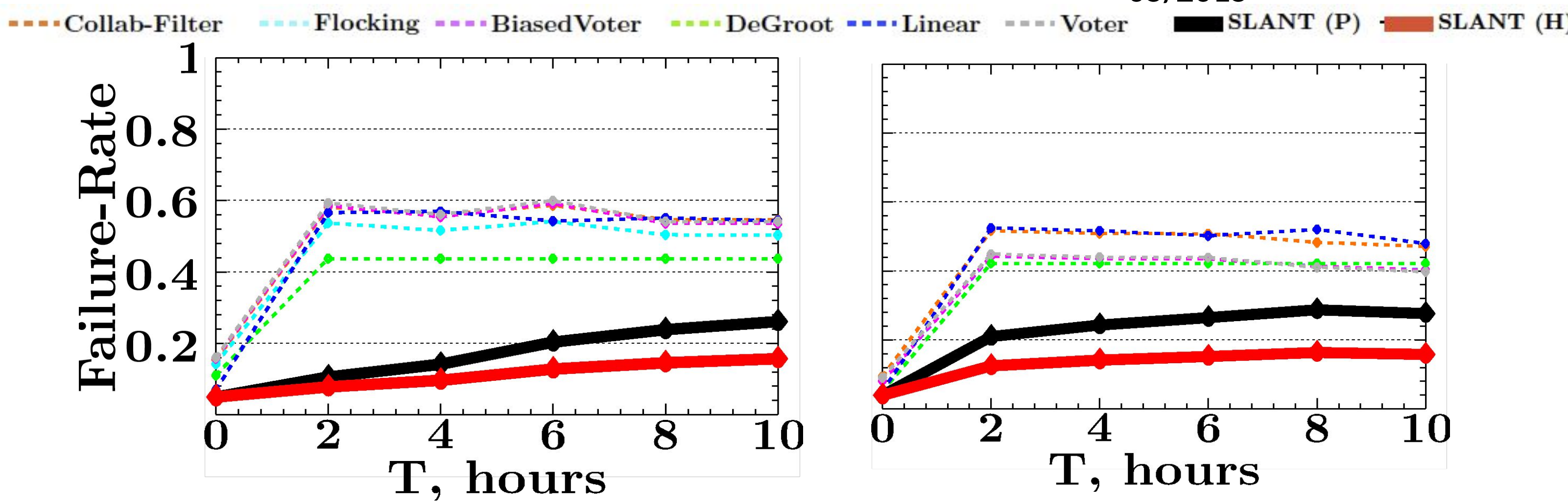
Our estimation method **scales to networks with million of nodes and events**

Model Formulation → Simulation and Estimation → Real-data Experiment

Microscopic prediction, Failure Rate

Delhi Assembly Election, 12/2013

Bollywood star incident 05/2015



Our model (in red) outperforms state of the art in terms of failure rate

$$\mathbb{P}(\mathrm{sign}(m) \neq \mathrm{sign}(\hat{m}))$$

Collaborators

Manuel Gomez Rodriguez

MPI SWS

Soumen Chakraborty

IIT Bombay

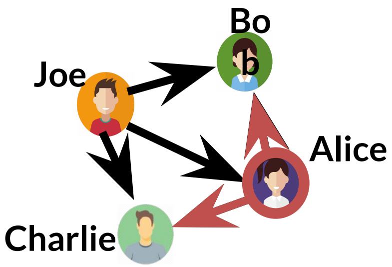
Isabel Valera

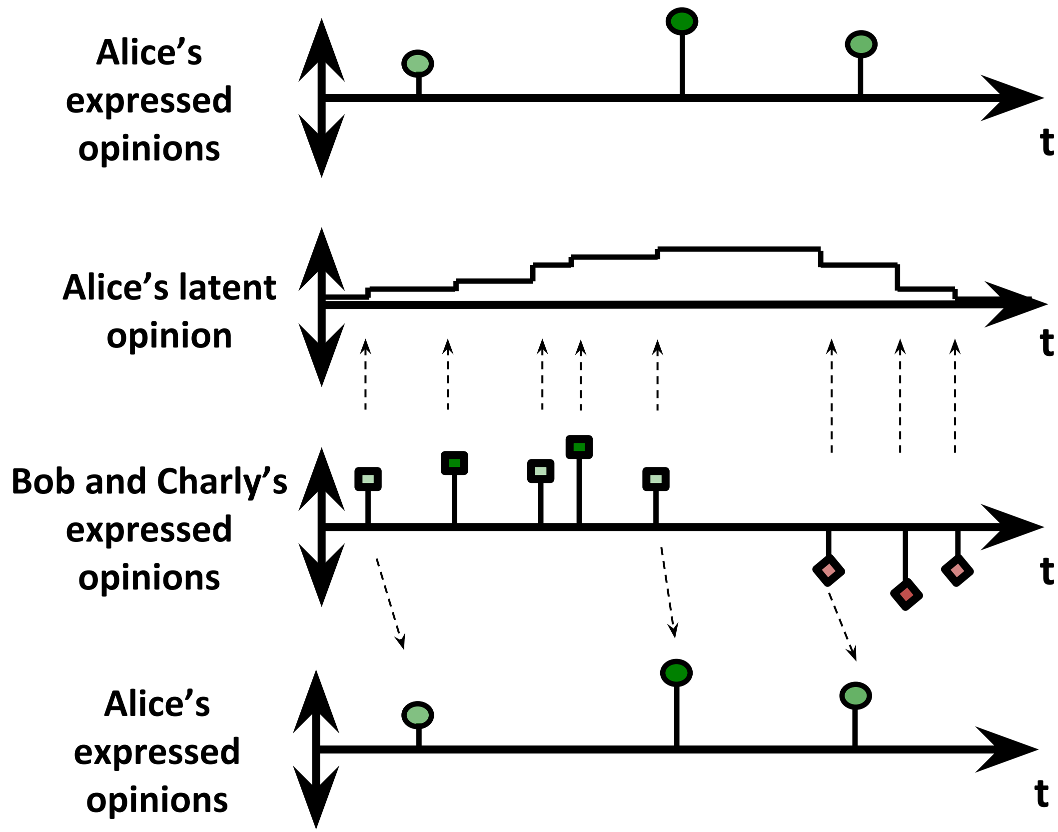
MPI SWS

Sourangshu Bhattacharya

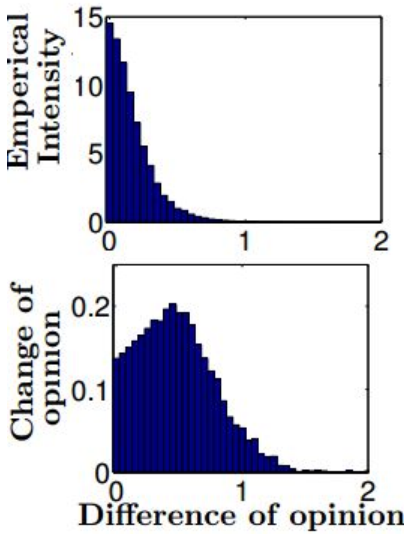
IIT Kharagpur

Key Idea: Latent vs expressed opinion



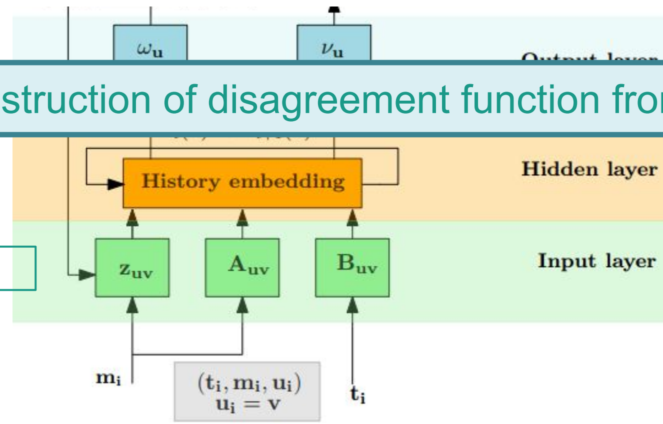


Explainability: A data driven construction



Analysis on conversations on Delhi Election 2015

Construction of disagreement function from data

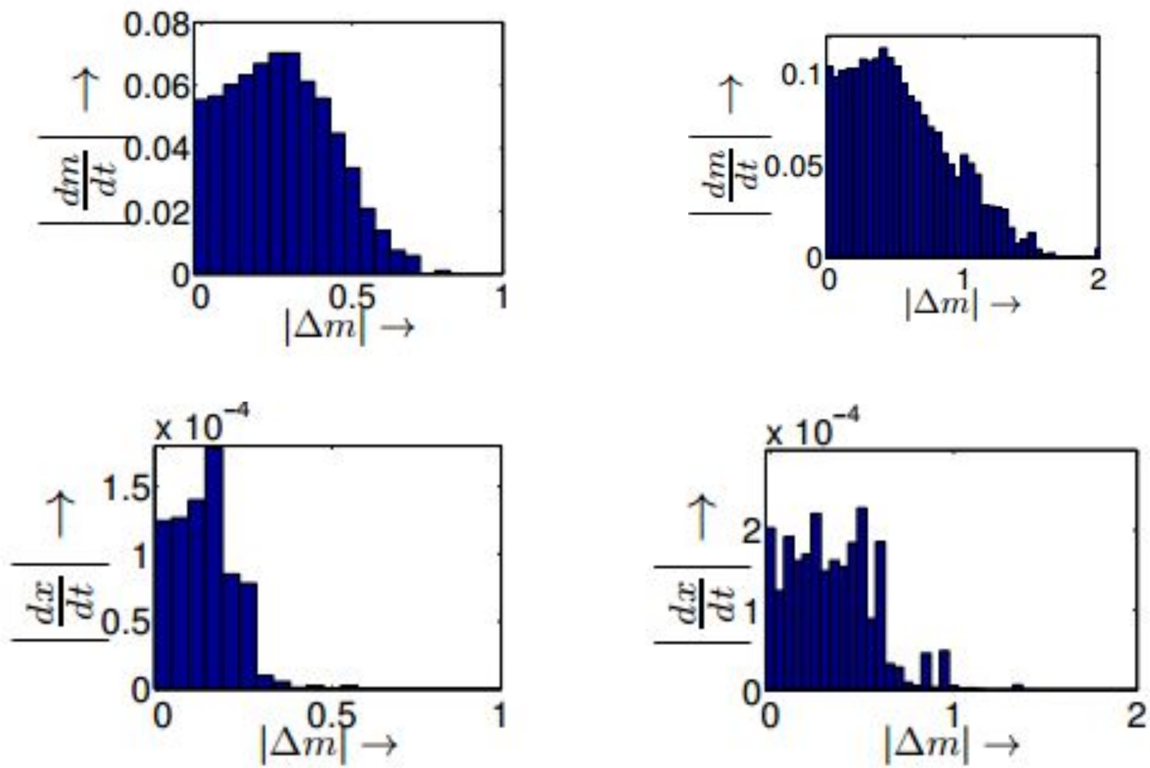


- Temporal data is usually limited, Twitter only allows 1% samples
- A prior embedding structure helps to learn from limited data.
  — works best for users who makes few comments

Disagreement fitting

Actual disagreement

Estimated Disagreement



(a) Movie

(b) Fight

CherryPick

$$\max_{\mathcal{A},\mathcal{O}} -\mathrm{tr}\ \log \boldsymbol{\Sigma}(\mathcal{A},\mathcal{O})$$

Doubly submodular

CherryPick

$$\max_{\mathcal{A},\mathcal{O}} -\mathrm{tr}$$

Greedy approach gives approximation guarantee

submodular