Machine learning for Dynamic Social Network Analysis

Applications: Models

Manuel Gomez Rodriguez
Max Planck Institute for Software Systems

IJCAI TUTORIAL, AUGUST 2017

Outline of the Seminar

REPRESENTATION: TEMPORAL POINT PROCESSES

1. Intensity function
2. Basic building blocks
3. Superposition
4. Marks and SDEs with jumps

APPLICATIONS: MODELS

1. Information propagation
2. Information reliability
3. Knowledge acquisition

Next

APPLICATIONS: CONTROL

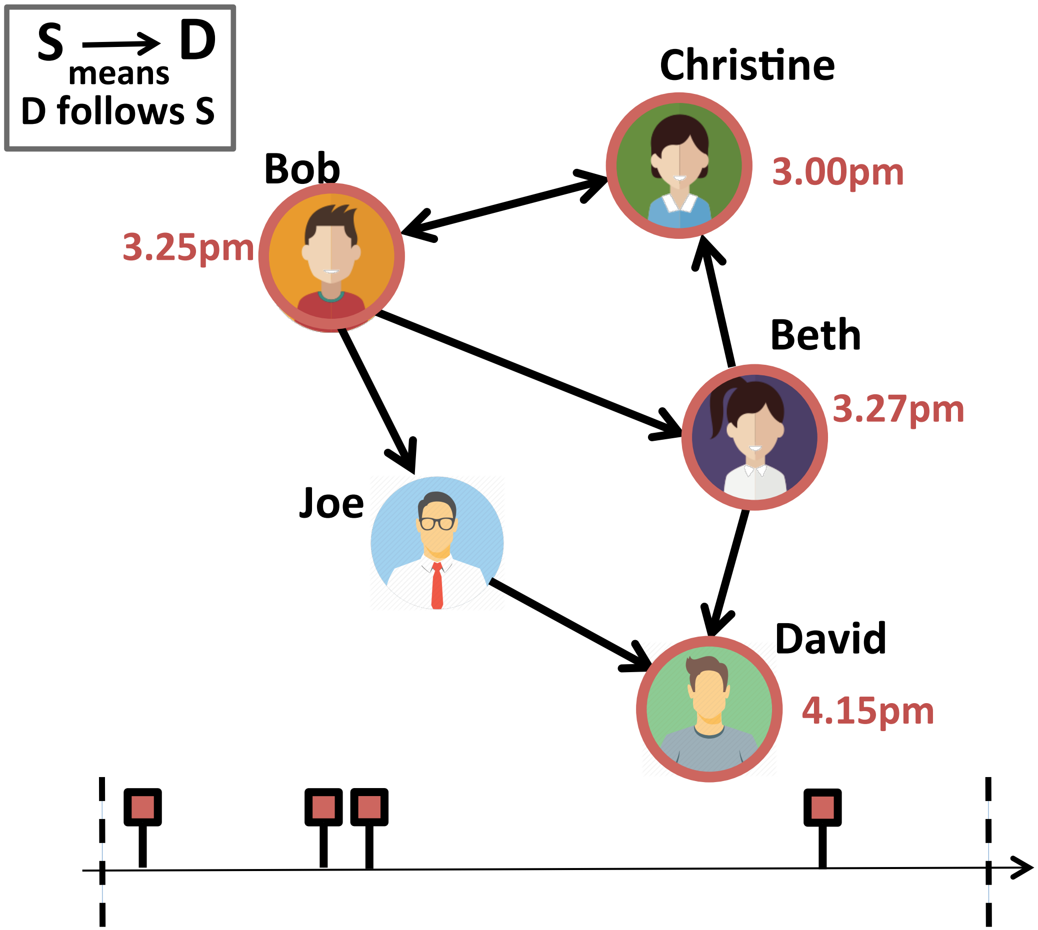
1. Activity shaping
2. When-to-post

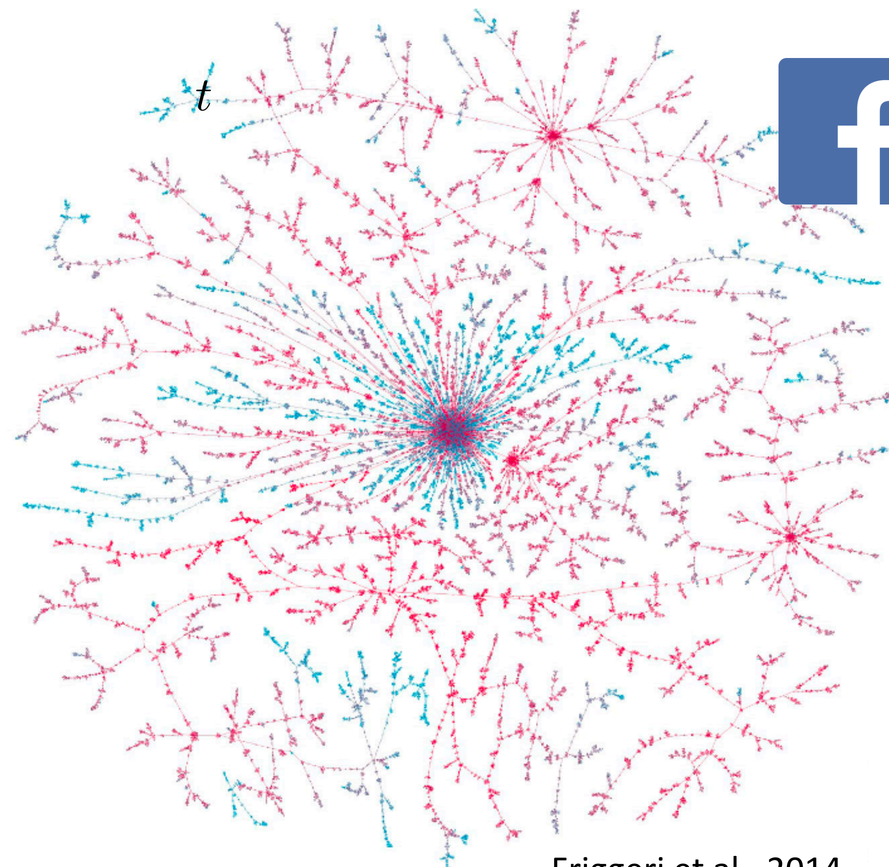
Slides/references: learning.mpi-sws.org/ijcai-2017-tutorial

Applications: Models

1. **Idea adoption**
2. Information reliability
3. Knowledge acquisition

Idea adoption: an example





Friggeri et al., 2014

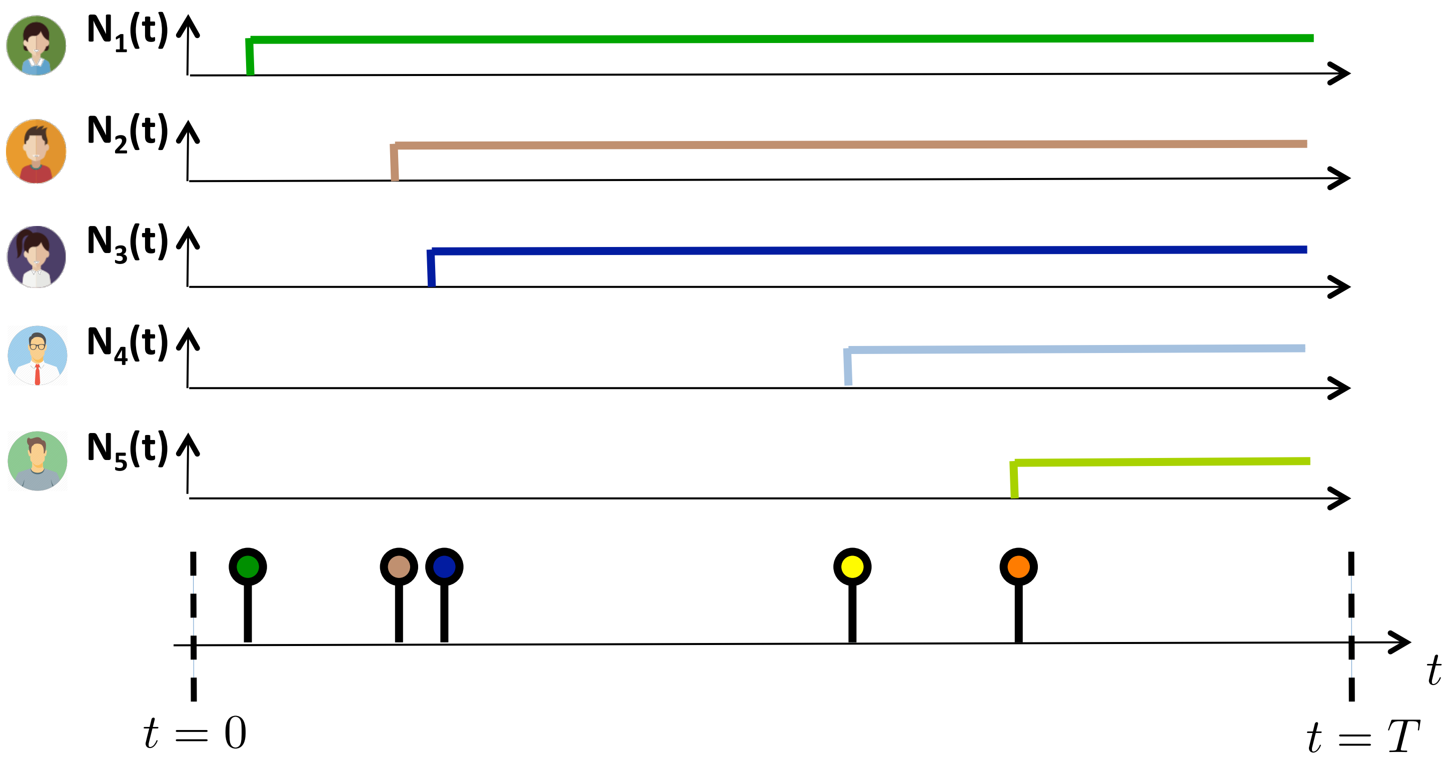
They can have an impact in the off-line world

theguardian

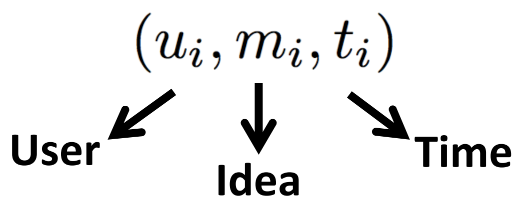
Click and elect: how fake news helped Donald Trump win a real election

Idea adoption representation

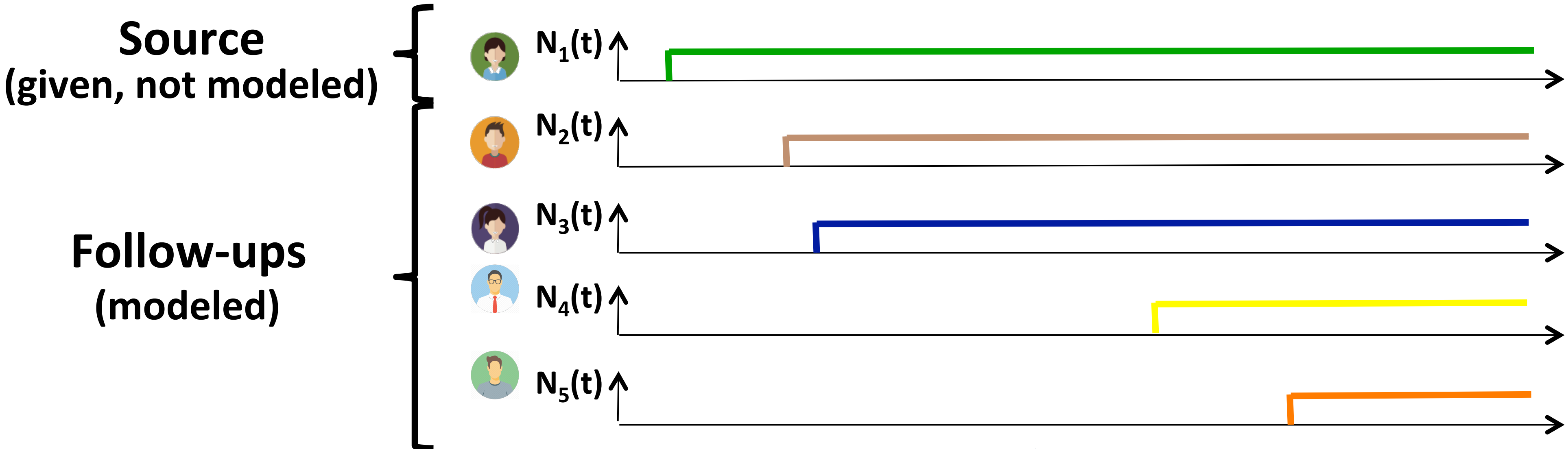
We represent an idea adoptions using **terminating temporal point processes**:

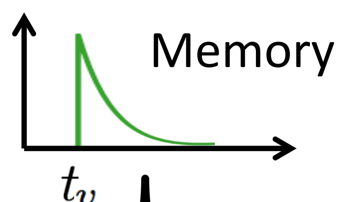


Idea adoption:



Idea adoption intensity





$$\lambda^*_u(t) = \underbrace{(1 - N(t))}_{\text{Adopt idea only once}} \sum_{v \in [m]} b_{vu} \, \kappa(t - t_v)$$

Influence from user v on user u

Time of message by user v

[Gomez-Rodriguez et al., ICML 2011]

Model inference from multiple adoptions

Conditional intensities

$$\lambda_u^*(t)$$

Idea adoption log-likelihood

$$\mathfrak{L} = \sum_{u=1}^{n} \log \lambda_u^*(t_u) - \int_0^T \lambda_u^*(\tau)\, d\tau$$

Sum up log-likelihoods of multiple ideas!

Maximum likelihood approach to find model parameters!

Theorem. For any choice of parametric memory, the **maximum likelihood** problem is **convex in B**.

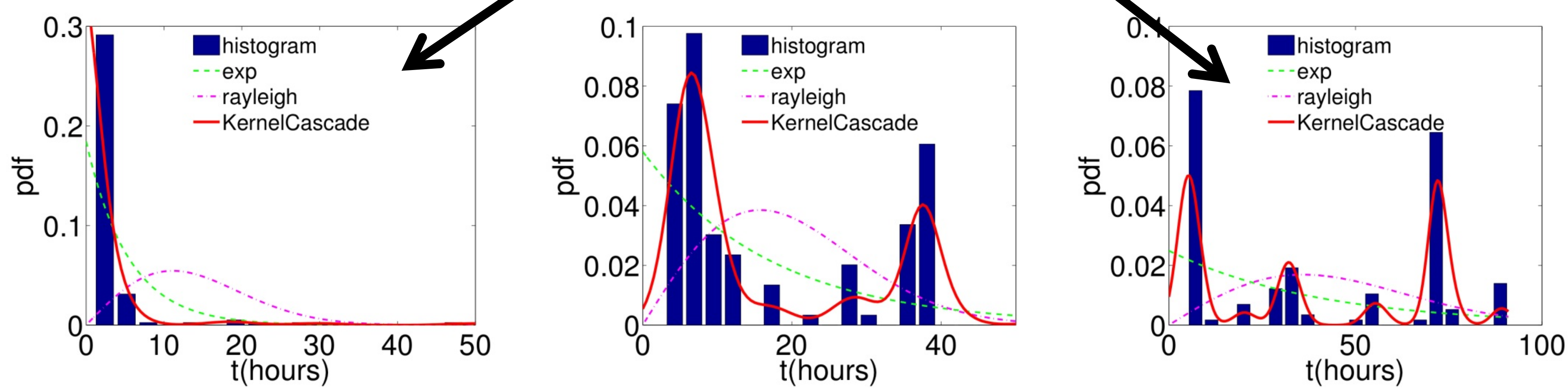
[Gomez-Rodriguez et al., ICML 2011]

Nonparametric kernels

Multimodal influence/memory:

$$\lambda_u^*(t) = (1 - N(t)) \sum_{v \in [m]} \kappa(t - t_v)$$

Mixture of kernels



Pair 1 Pair 2 Pair 3

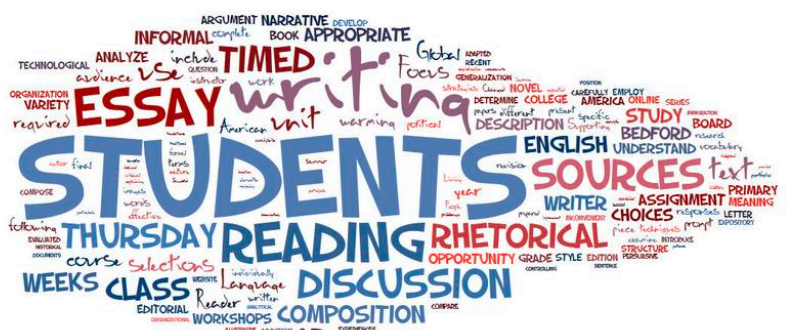
[Du et al., NIPS 2012]

Topic-sensitive rates

Topic-modulated influence:

$$b_{vu} = \sum_{l=1}^{K} b_{vu}^{l} \, m_l$$

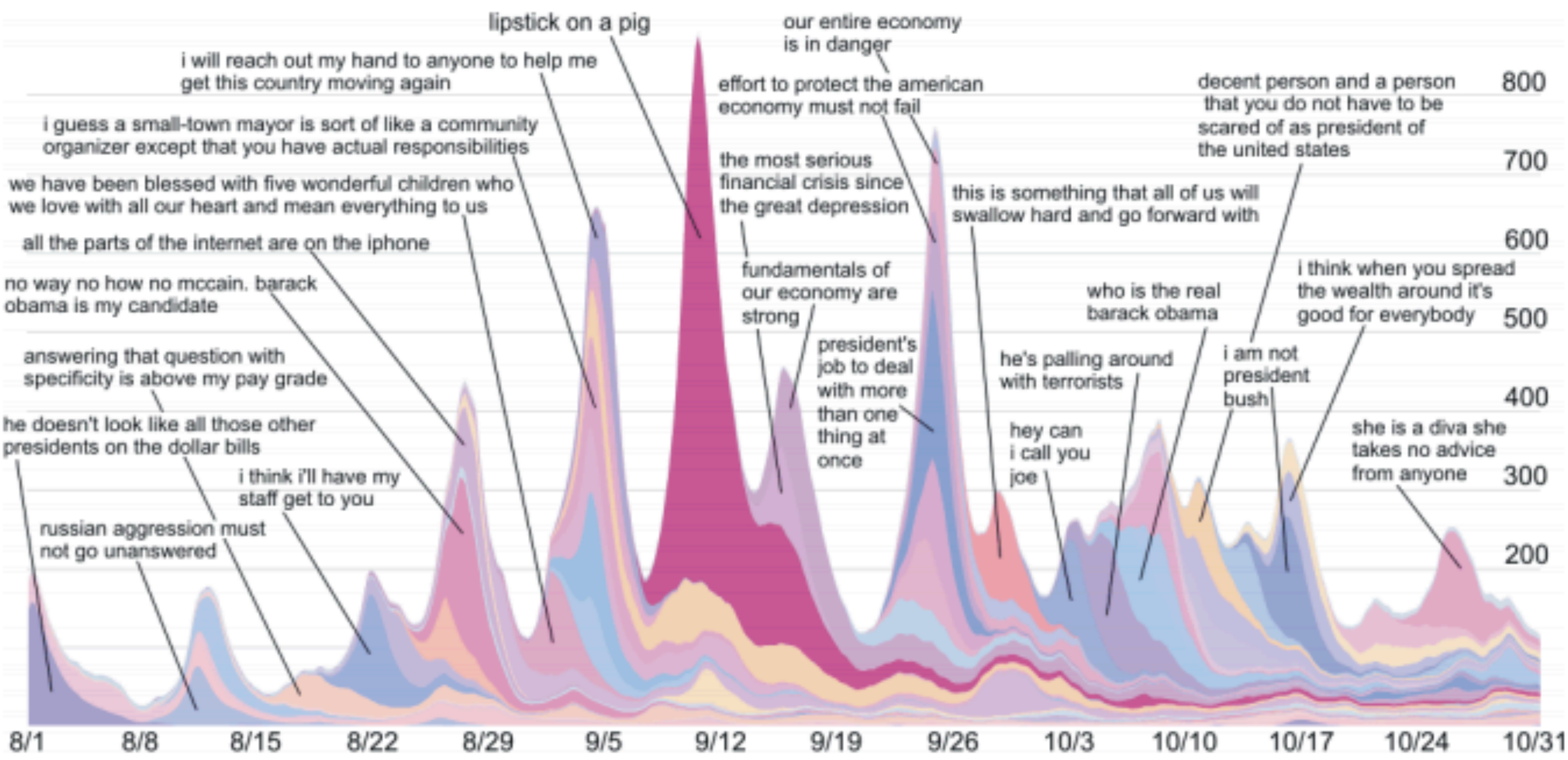
LDA weight for topic l



tagxedo.com

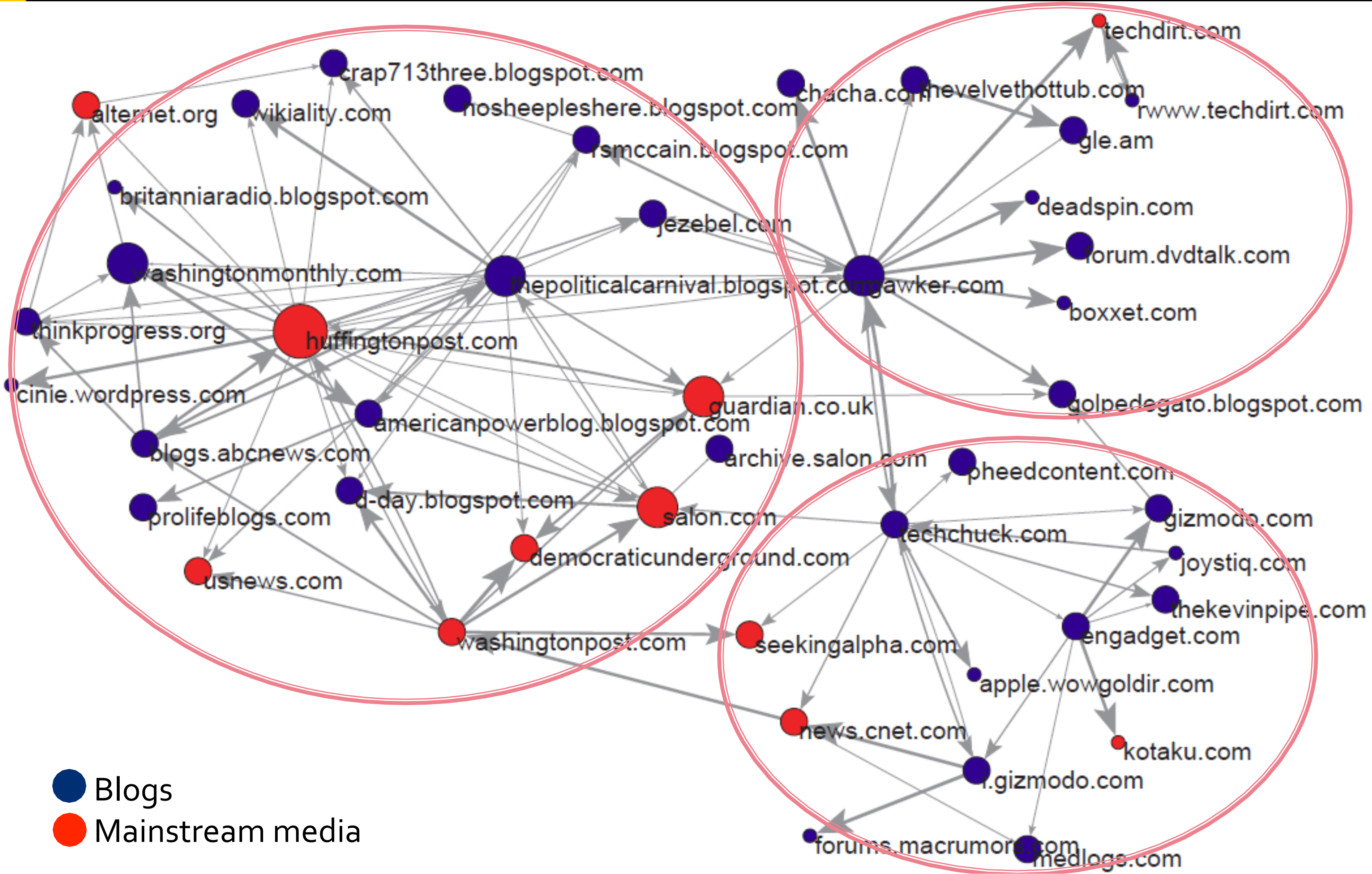
[Du et al., AISTATS 2013]

Memetracker



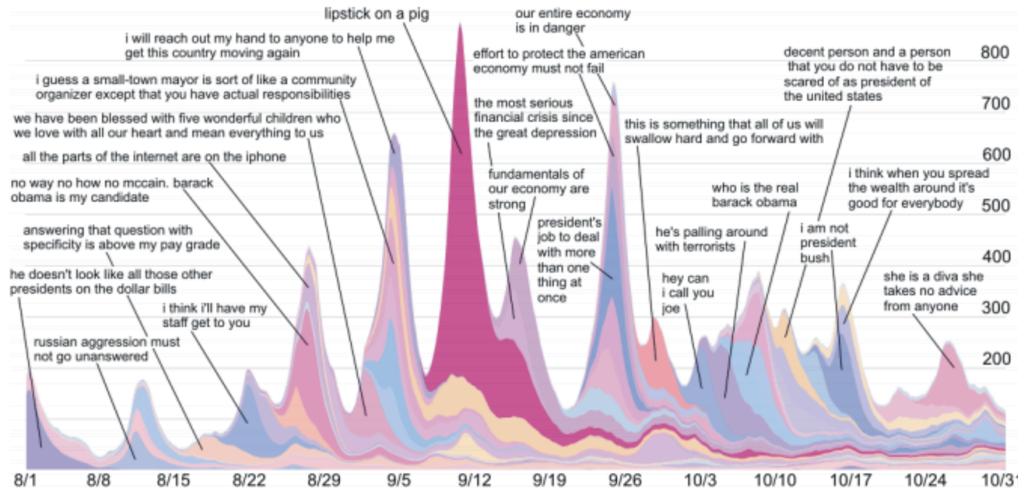
[Leskovec et al., KDD '09]

Diffusion Network (small part)



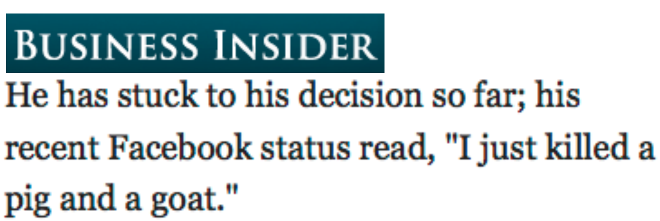
Recurrent events: beyond cascades

Up to this point, we have assumed we can map each event to a cascade



In general, especially in social networks:

Difficult **to distinguish cascades** in event data

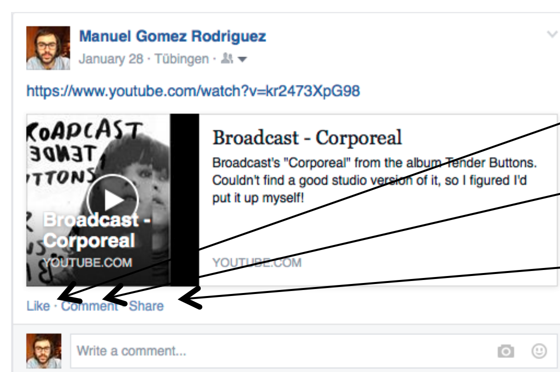

BUSINESS INSIDER

He has stuck to his decision so far; his recent Facebook status read, "I just killed a pig and a goat."


GAWKER

Mark Zuckerberg Is Killing Progressively Larger Animals

Most cascades are single nodes (or *forests*)



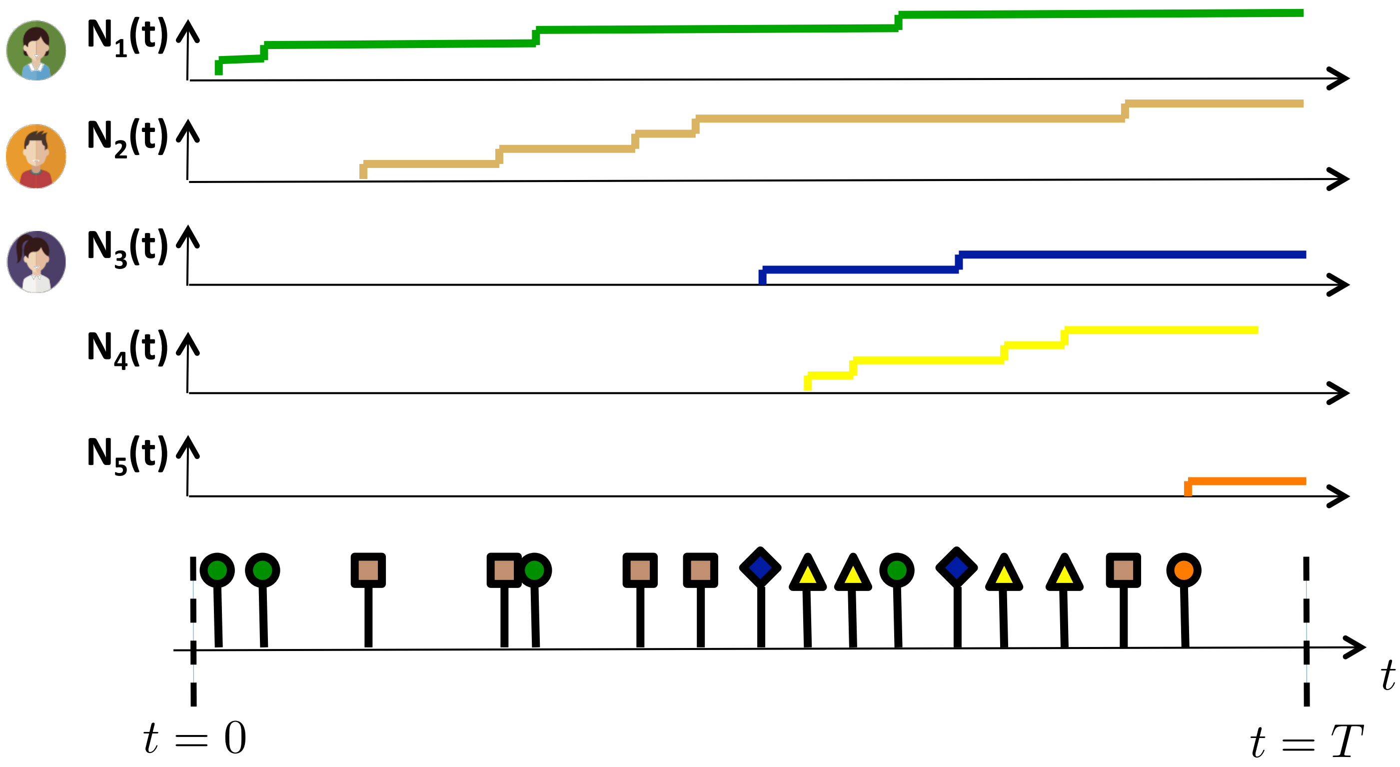
no likes

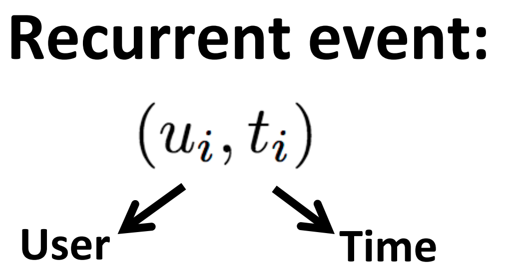
no comments

no shares

Recurrent events representation

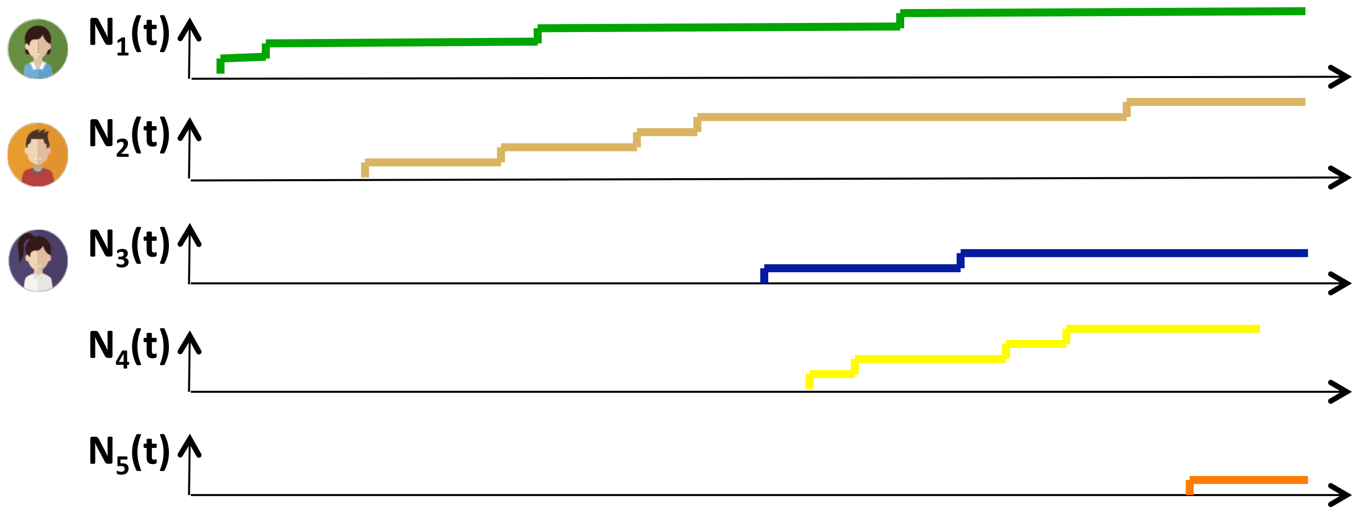
We represent messages using **nonterminating temporal point processes**:



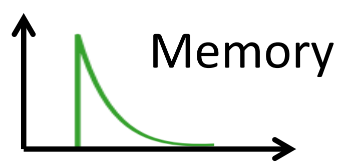


[Farajtabar et al., NIPS 2014]

Recurrent events intensity



Cascade sources!



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

Influence from user v on user u

Hawkes process

[Farajtabar et al., NIPS 2014]

Applications: Models

1. Information propagation
2. Information reliability
3. Knowledge acquisition

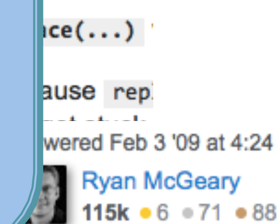
Information reliability: an example

Learning from the crowd ('crowdlearning') has become very popular:

Users learn from the kno

Knowledge is reviewed ... n ... :

Refutations and verifications depend on the source trustworthiness



03:21, 20 September 2016

is a $\boxed{\text{Kenyan}}$ politician

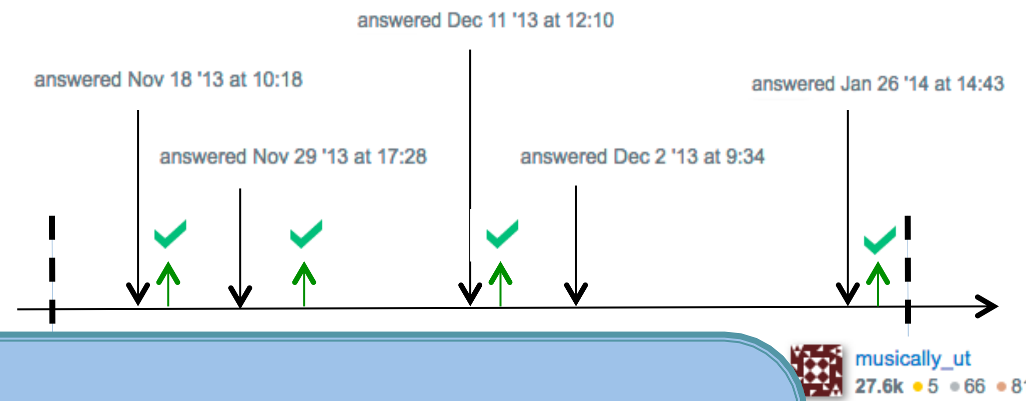
✗ ↓ *possible vandalism by MLM2016*

is an American politician

Information reliability: key, simple idea

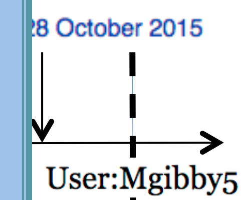
A **source is trustworthy** if:

Its contributions are
verified more frequently



Its c
refu

Over time, each *document* has a different level of ***inherent* unrealibility**



Challenge

At a time t, a ***document*** **may be *disputed***

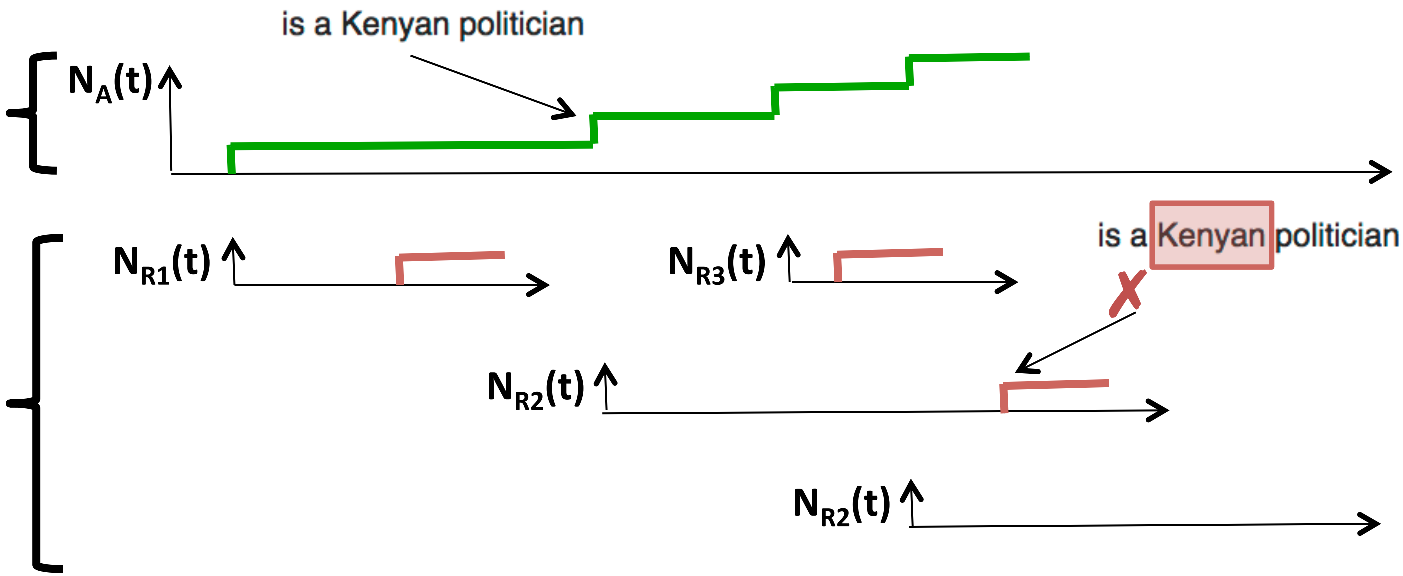
Verifications: rarer
Refutations: more frequent

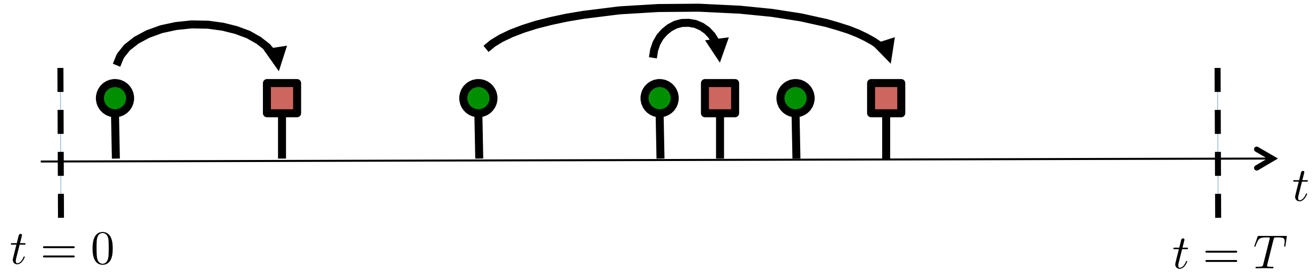
[Tabibian et al., WWW 2017]

Representation: temporal point processes

Statement additions
(one process per document)

Statement refutations
(one process per statement)





source ↓ refutation time ↓

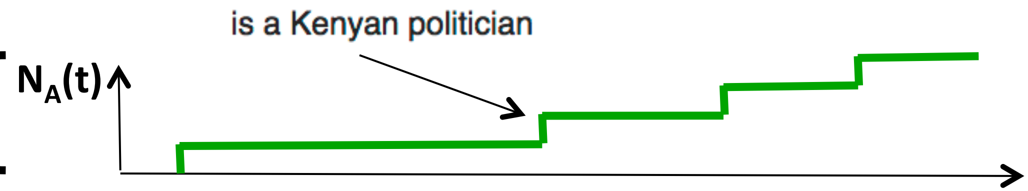
Statement: $e = (s, t, \tau)$

↑ addition time

[Tabibian et al., WWW 2017]

Intensity of statement additions

Statement additions
(one process per article)

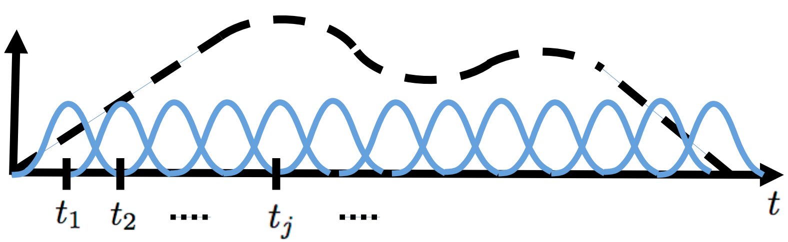


$$\lambda_d(t) = \sum_j \phi_{d,j} k(t - t_j) + \sum_{e_i \in \mathcal{H}_d(t)} \mathbf{w}_d^\top \gamma_{s_i} g(t - \tau_i)$$

Intensity or rate
(Statements per time unit)

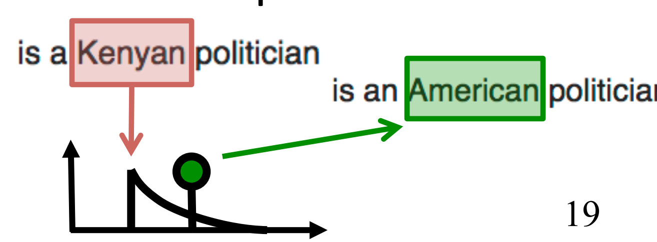
Article unreliability
(Mixture of Gaussians)

Temporal evolution of the *intrinsic* reliability of the article



Effect of past refutations
(topic dependent; topic weight w_d)

Refuted statements trigger the arrival of new statements to replace them

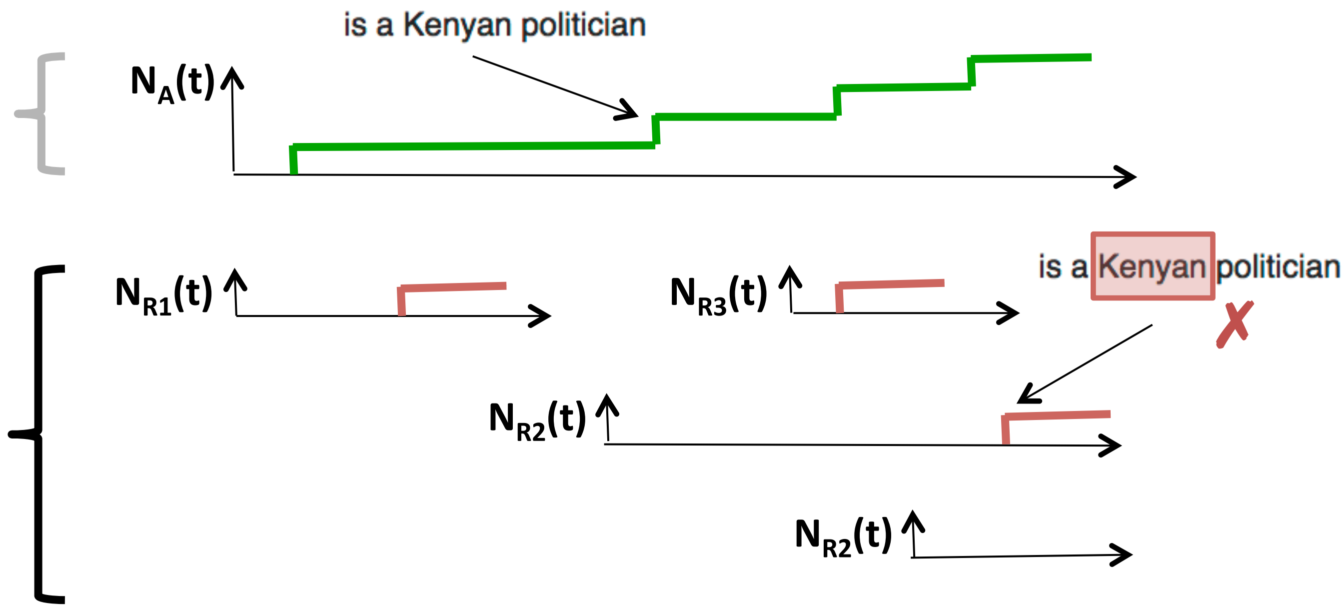


[Tabibian et al., WWW 2017]

Intensity of statement refutations

Statement additions
(one process per article)

Statement refutations
(one process per statement)



$$\mu_i(t) = \underbrace{(1 - N_i(t))}\left[\underbrace{\sum_j \beta_{d,j}\, k(t + t_i - t_j)} + \underbrace{\mathbf{w}_d^\top \boldsymbol{\alpha}_{s_i}}\right]$$

Intensity or rate
(Statements per time unit)

Refutations happen only once

Article unreliability
(Mixture of Gaussians)

Shared across statements of an article!

Source trustworthiness
(topic dependent; topic weight w_d)

The higher the parameter α_{s_i}, the quickest an article gets refuted

[Tabibian et al., WWW 2017]

Model inference from event data

Conditional intensities

$$\{\lambda_d(t)\} \quad \{\mu_i(t)\}$$

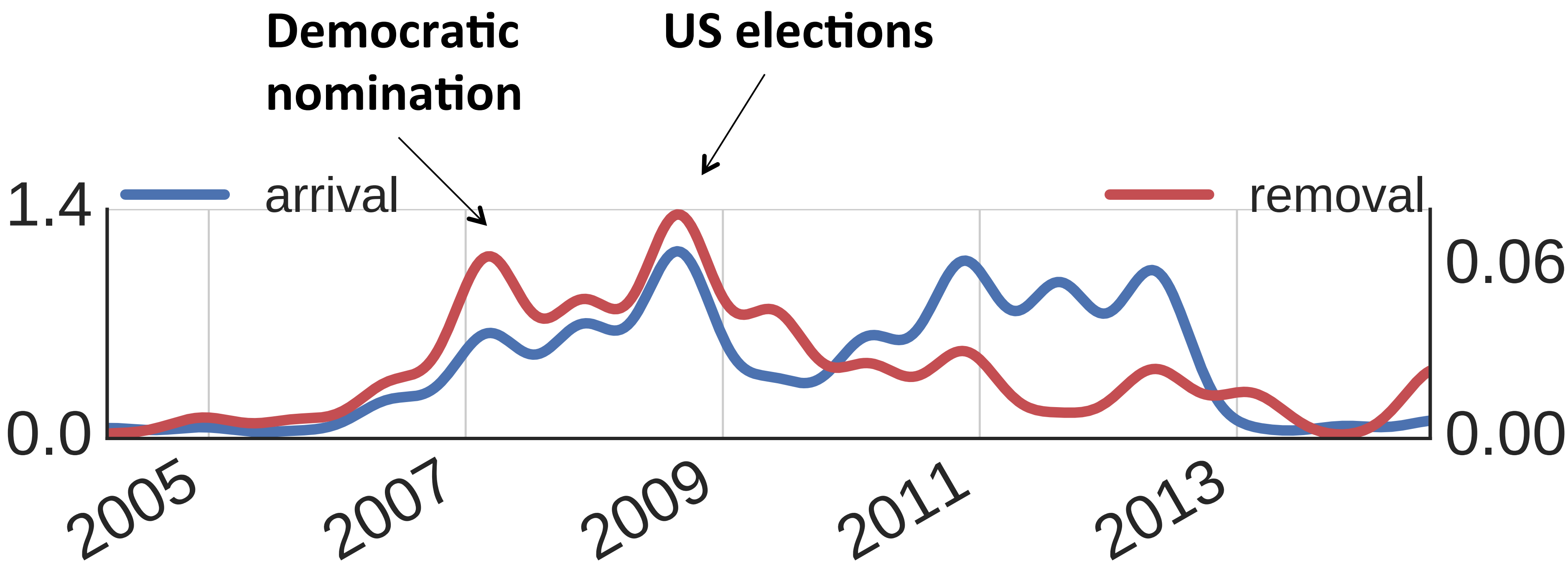
Events likelihood

$$\sum_{d=1}^{|\mathcal{D}|} \sum_{i:e_i \in \mathcal{H}_d(T)} \log \underbrace{p(t_i|\mathcal{H}_d(t_i), \boldsymbol{\phi}_d, \{\boldsymbol{\gamma}_s\}, \mathbf{w}_d)}_{\text{statements additions}} + \sum_{d=1}^{|\mathcal{D}|} \sum_{i:e_i \in \mathcal{H}_d(T)} \log \underbrace{p(\Delta_i|t_i, \boldsymbol{\beta}_d, \{\boldsymbol{\alpha}_s\}, \mathbf{w}_d)}_{\text{statements evaluations}}$$

Theorem. The **maximum likelihood** problem is **convex in the model parameters**.

[Tabibian et al., WWW 2017]

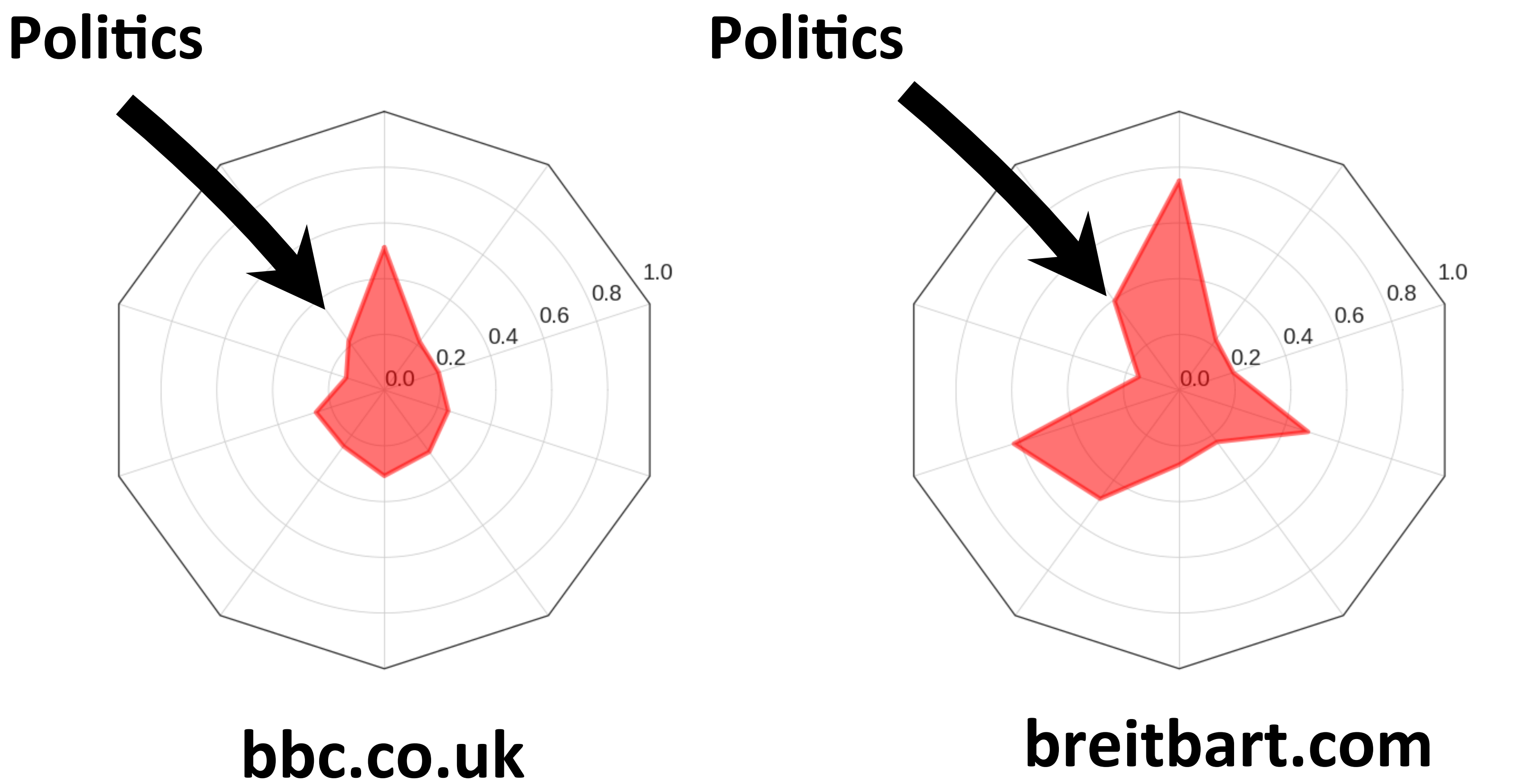
Wikipedia article reliability



Barack Obama's Wikipedia Article
(Arrival of information vs intrinsic unreliability)

[Tabibian et al., WWW 2017]

Source trustworthiness



Probability of refutation within 6 months in a *stable* Wikipedia article

[Tabibian et al., WWW 2017]

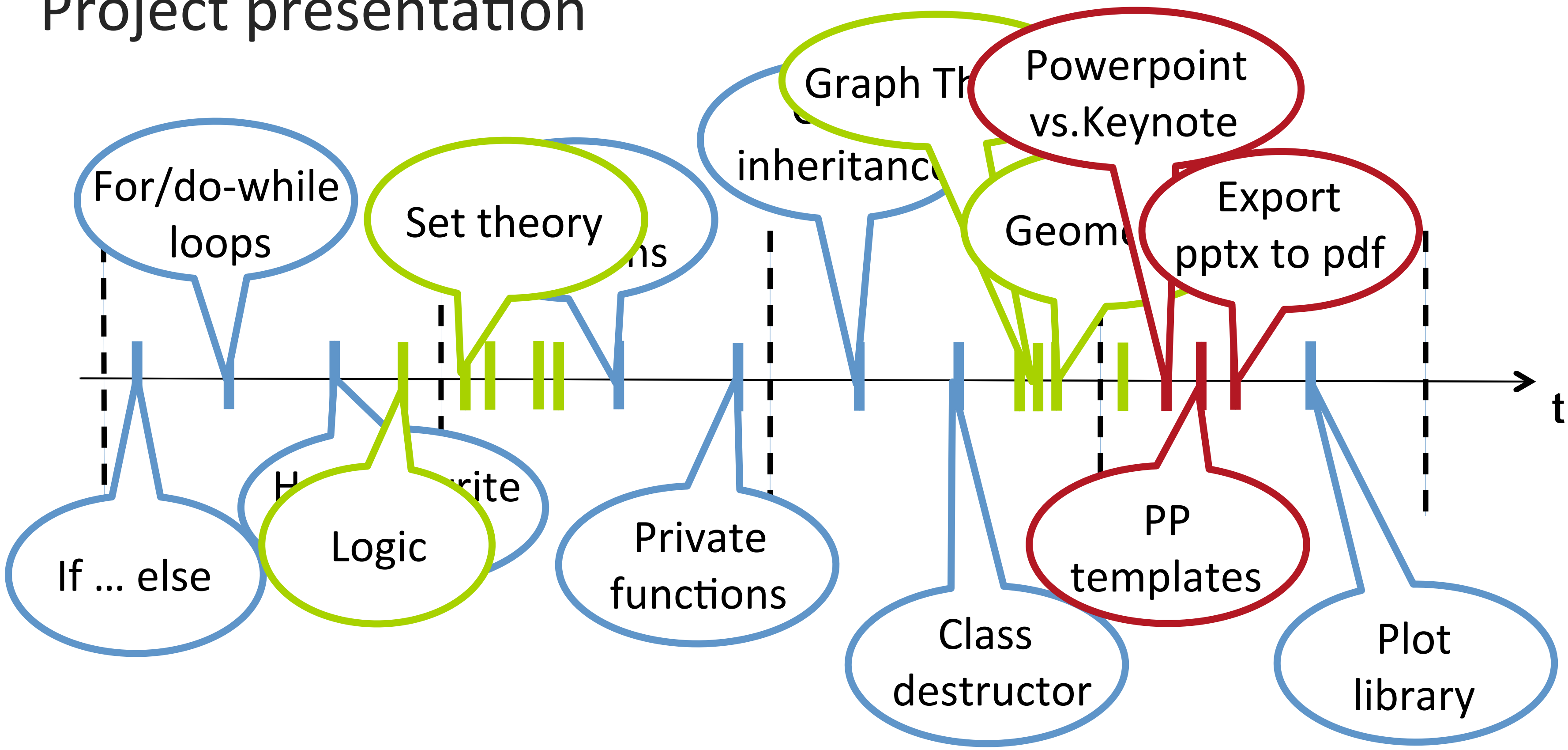
Applications: Models

1. Information propagation
2. Information reliability
3. Learning patterns

Learning patterns: An example

1st year computer science student

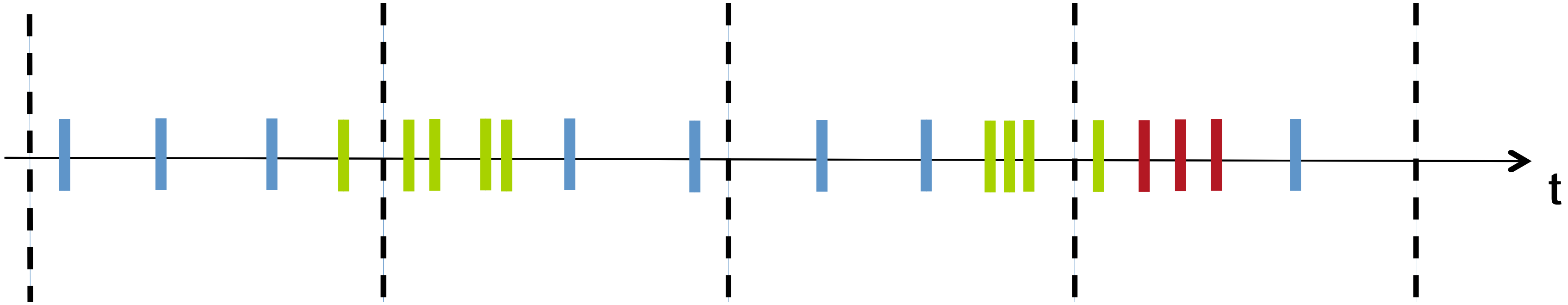
- Introduction to programming
- Discrete math
- Project presentation



Learning patterns: content + dynamics

1st year computer science student

- Introduction to programming
- Discrete math
- Project presentation



Content + Dynamics = ***Learning pattern***

programming + semester

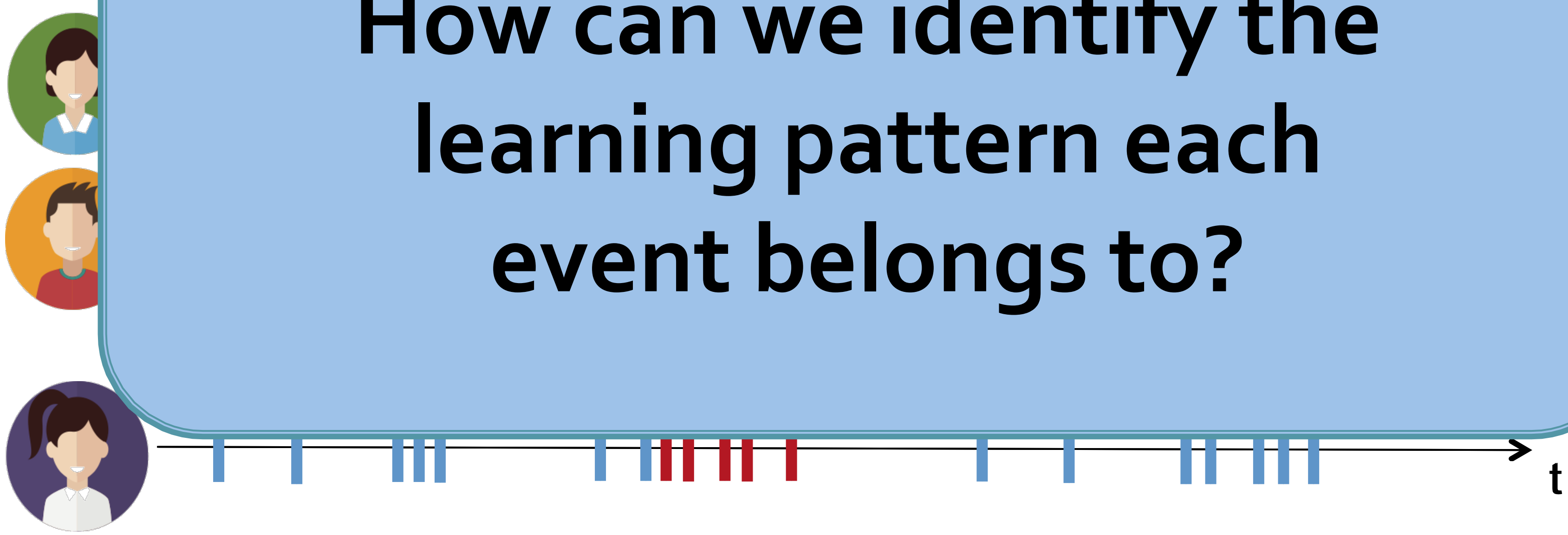
math + semester

presentation + week

People share same learning patterns

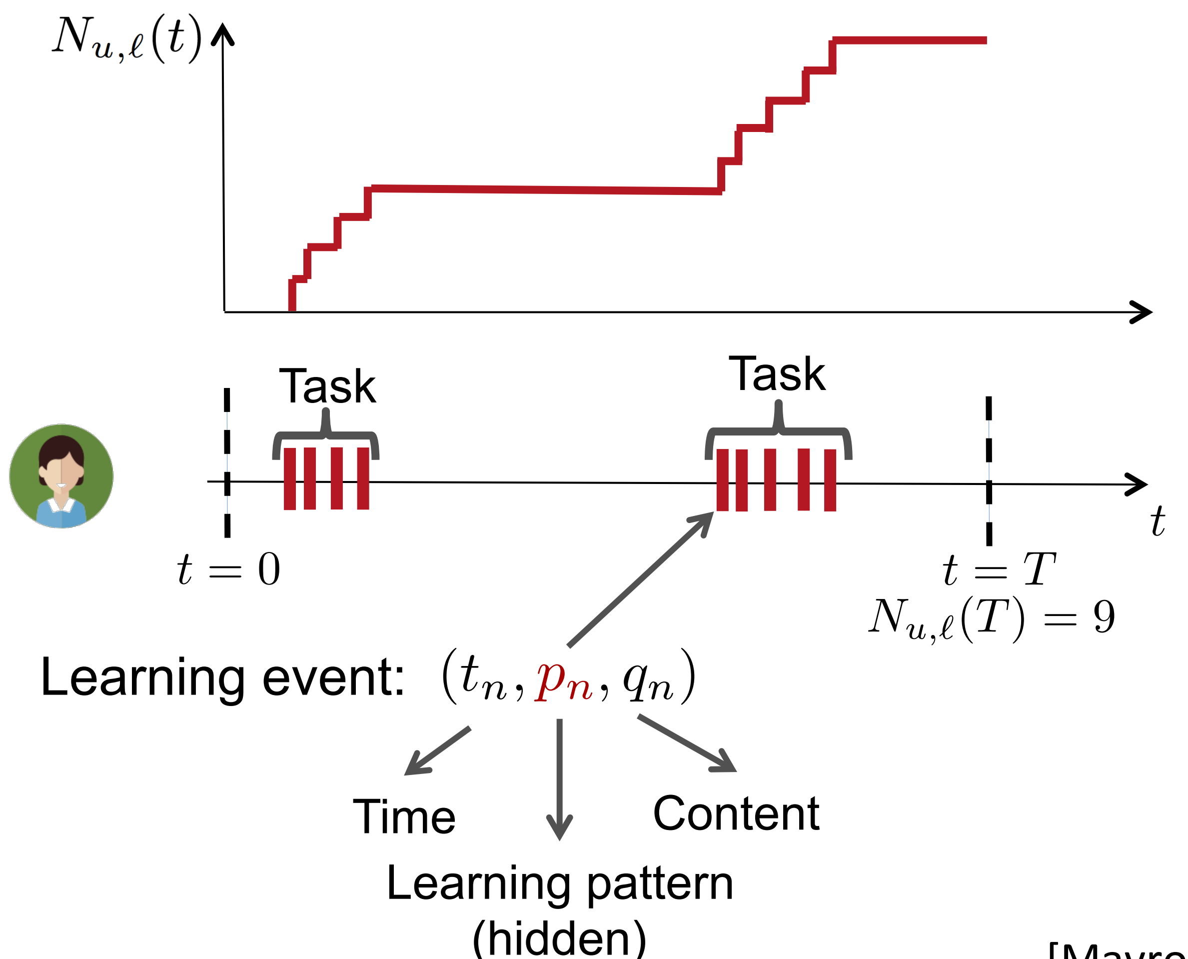
- Introduction to programming
- Discrete math
- Pr

How can we identify the learning pattern each event belongs to?



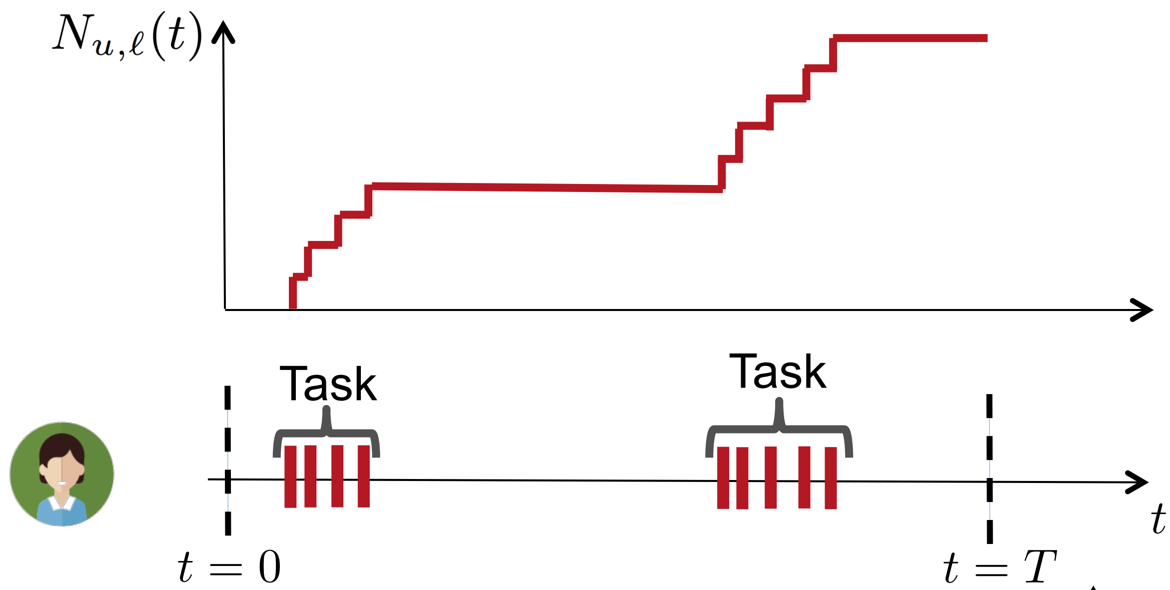
Learning events representation

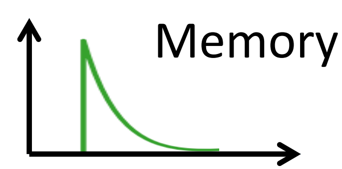
We represent the learning events using **marked temporal point processes**:



[Mavroforakis et al., WWW 2017]

Learning pattern intensity





New task rate

Learning pattern popularity

$$\underbrace{\lambda^*_{u,\ell}(t)}_{\text{Intensity or rate (events / hour)}} = \underbrace{\mu_u \pi_\ell}_{\text{New task}}$$

Intensity or rate
(events / hour)

New task

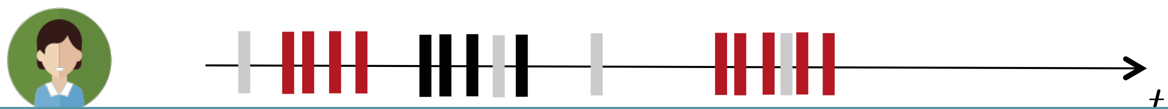
Follow-up

Hawkes process

[Mavroforakis et al., WWW 2017]

User learning events intensity

Users adopt more than one learning pattern:



# of learning patterns is **infinite**.

Efficient model inference using

Sequential Montecarlo!

Details in the reference below!

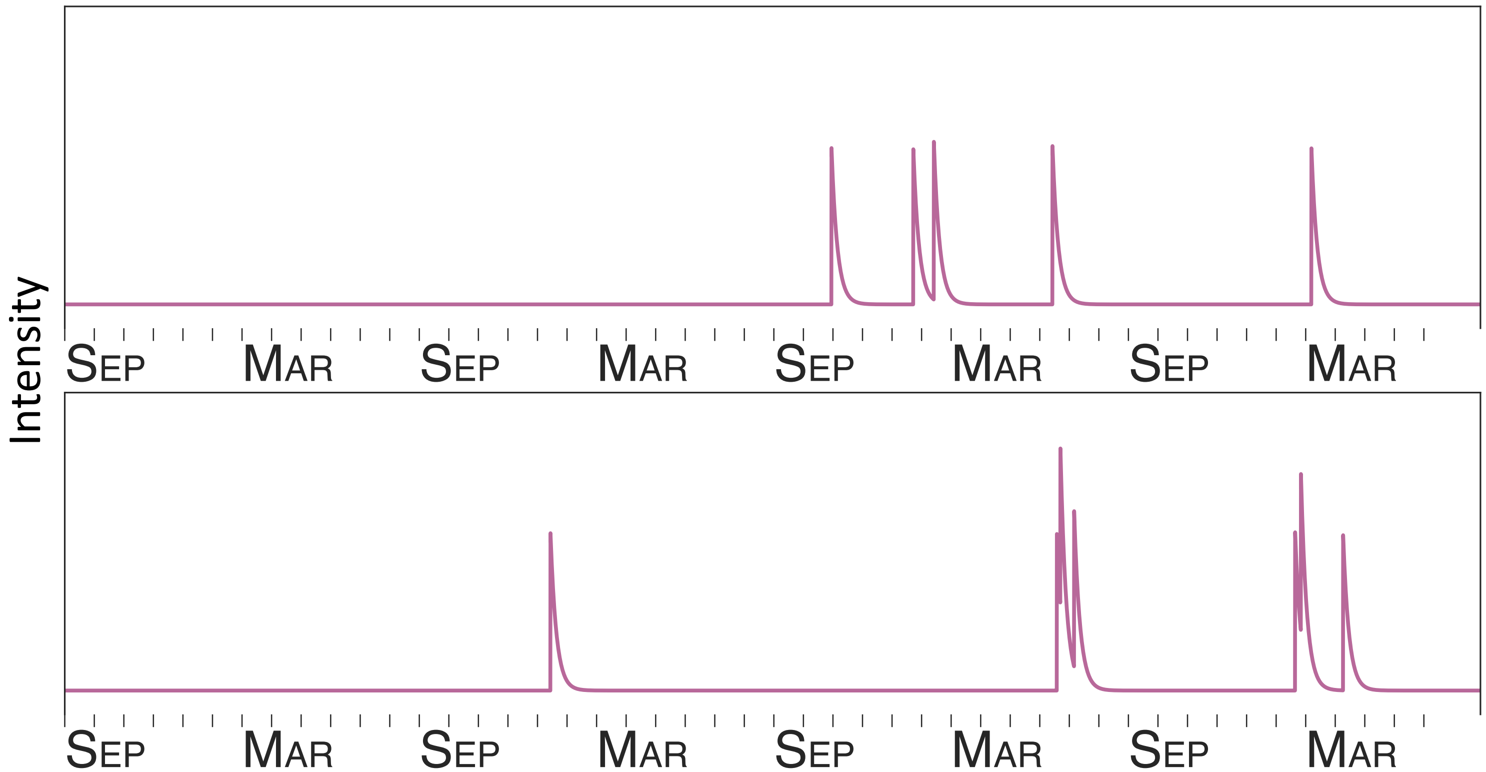
Content $\rightarrow$ $q_n = \boldsymbol{\omega}$ $\quad \omega_j \sim Multinomial(\boldsymbol{\theta}_p)$

[Mavroforakis et al., WWW 2017]

Learning pattern (I): Version Control

Content

Intensities

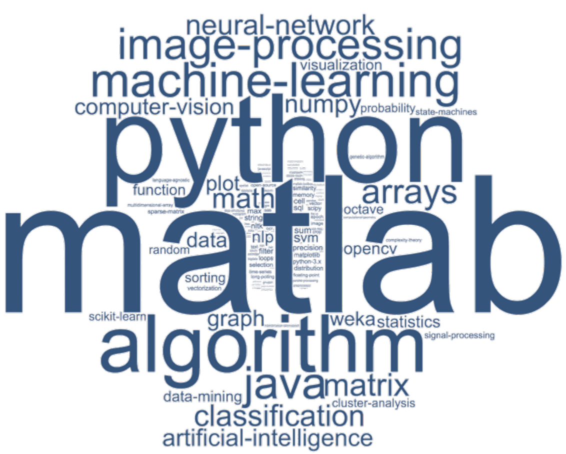


Version control tasks tend to be specific, quickly solved after performing few questions

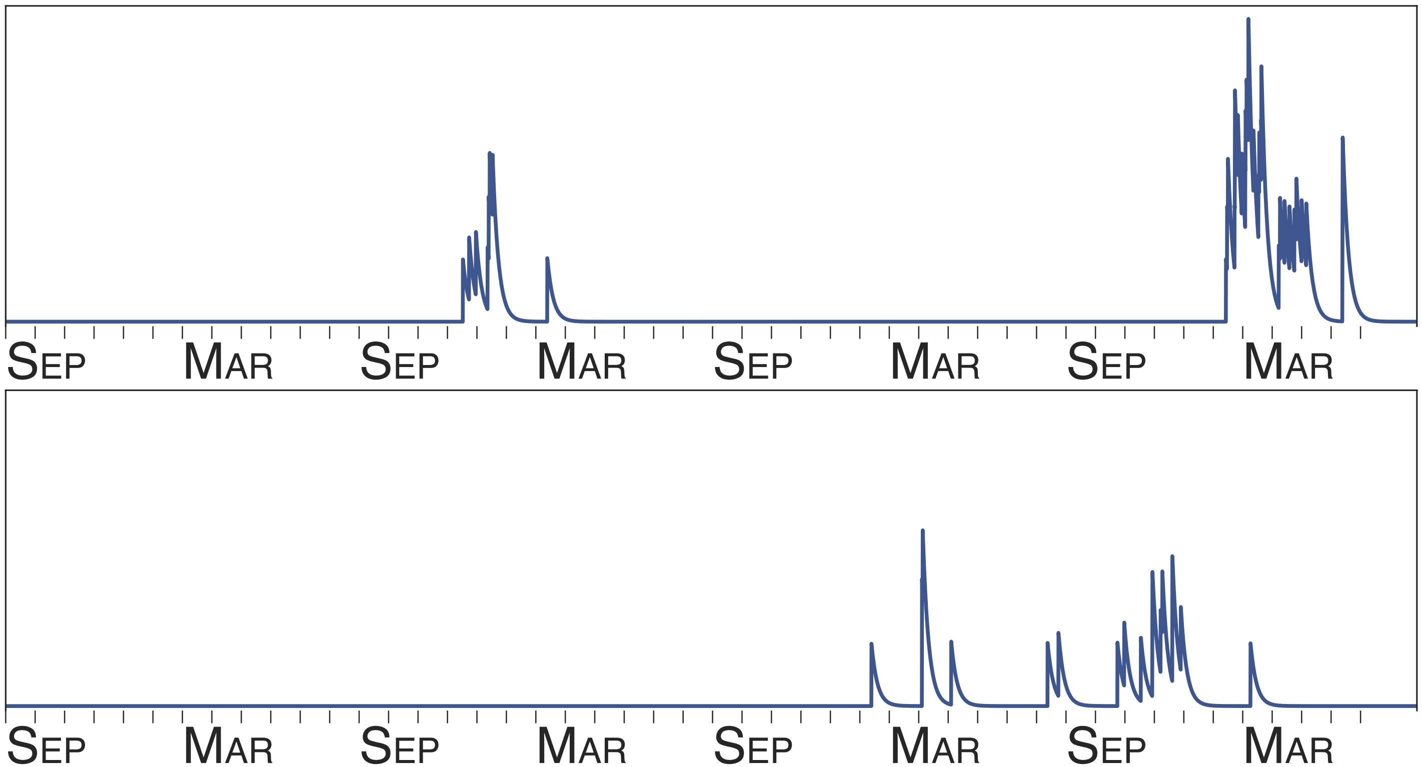
[Mavroforakis et al., WWW 2017]

Learning pattern (II): Machine learning

Content



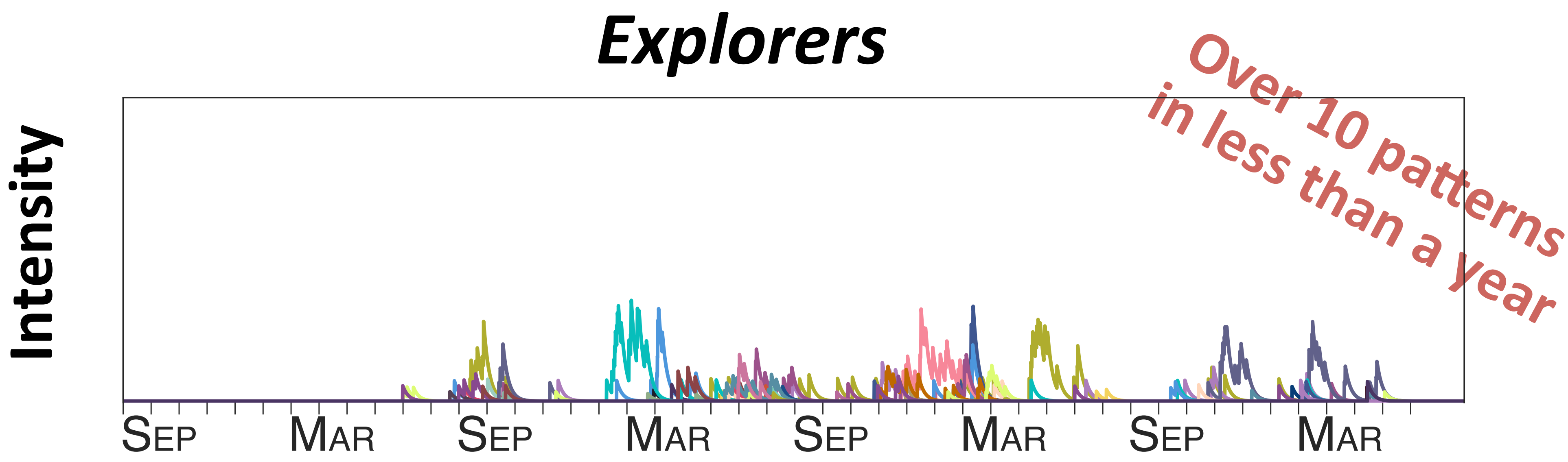
Intensities

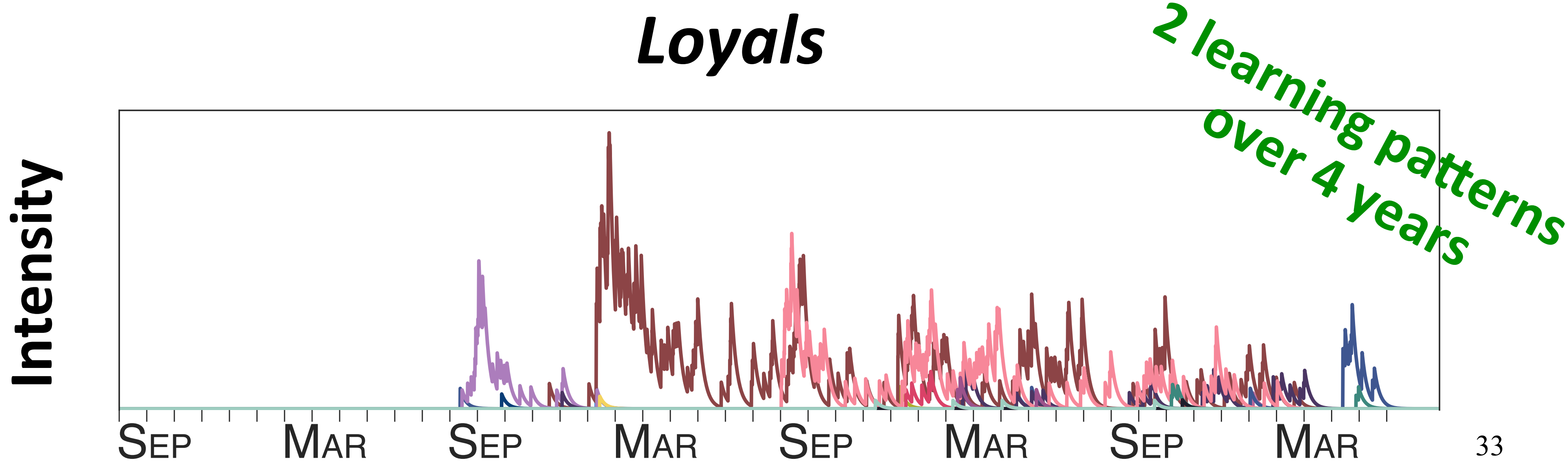


Machine learning tasks tend to be more complex and require asking more questions

[Mavroforakis et al., WWW 2017]

Types of users





[Mavroforakis et al., WWW 2017]

Representation: Temporal Point processes

1. Intensity function
2. Basic building blocks
3. Superposition
4. Marks and SDEs with jumps

Applications: Models

1. Information propagation
2. Information reliability
3. Knowledge acquisition

Applications: Control

1. Activity shaping
2. When-to-post

Next

Slides/references: learning.mpi-sws.org/ijcai-2017-tutorial