Stochastic optimal control

of Marked Temporal Point Processes

HUMAN-CENTERED MACHINE LEARNING

http://courses.mpi-sws.org/hcml-ws18/

MAX PLANCK INSTITUTE
FOR SOFTWARE SYSTEMS

What is optimal control used for?

Optimal control aims to find an *optimal action* to solve a *task* in an *environment*

→ https://www.youtube.com/watch?v=Lphi7EeU37s Cart pole balancing

→ https://www.youtube.com/watch?v=vjSohj-Iclc Boston dynamics I

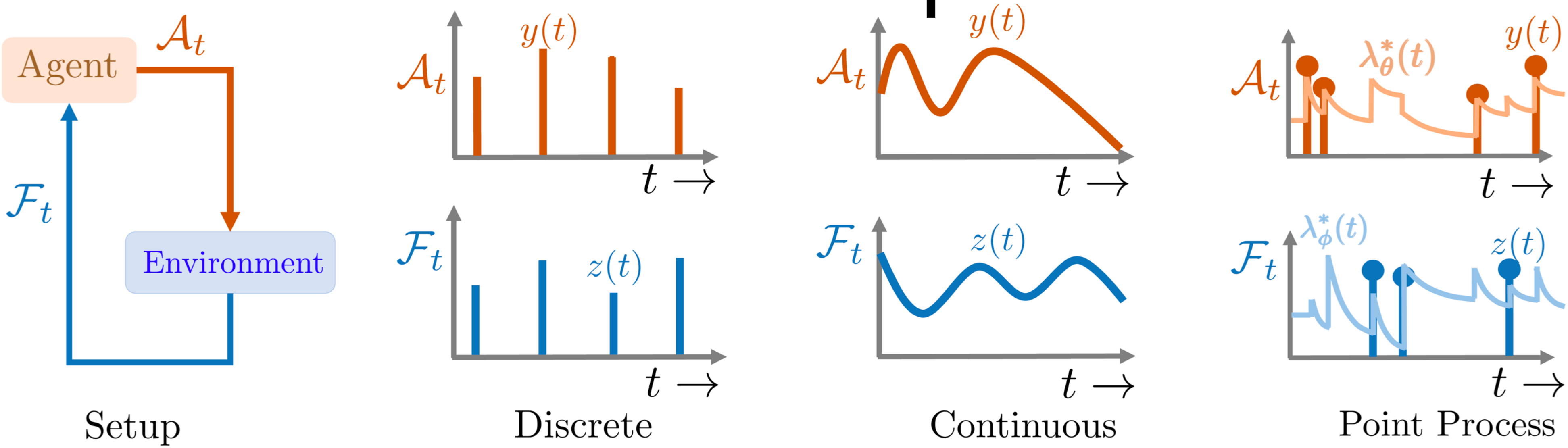
→ https://www.youtube.com/watch?v=fUyU3lKzoio Boston dynamics II

One needs to accurately model how the environment reacts to the actions via:

- **(Stochastic) differential equations**
- **(Stochastic) difference equations**

Optimal control on different problem settings

Actions and feedback are real-valued functions in continuous time



Actions and feedback occur in discrete time

Actions and feedback are asynchronous events localized in continuous time

Example I: Viral marketing

Agent

Social media user

Environment

Followers' Feed

Forbes

For Brands And PR: When Is The Best Time To Post On Social Media?

THE HUFFINGTON POST

The Best Times to Post on Social Media

When to post to maximize views or likes?

$$\mu_i(t) = u(t) \longrightarrow N_i(t)$$

Design (optimal) posting intensity

Marks (feedback) given by environment

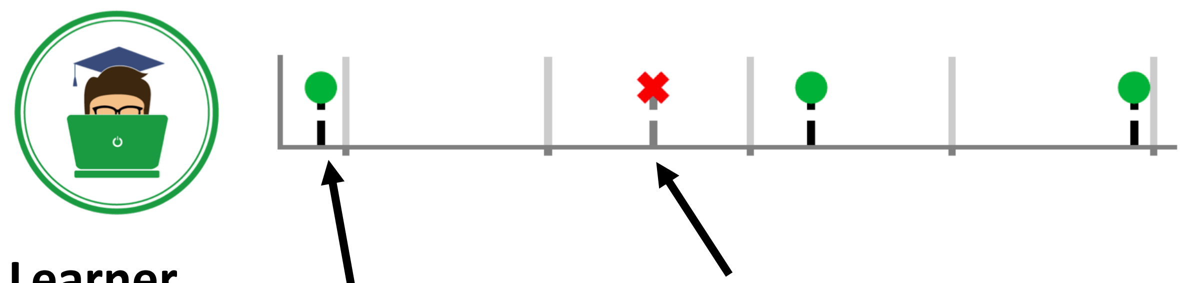
Example II: Spaced repetition

Agent

duolingo

Online learning platform

Environment



Learner

Review & successful recall

Review & unsuccessful recall

When to review to maximize recall probability?

$\lambda_i(t) \longrightarrow N_i(t)$

Design (optimal) reviewing intensities

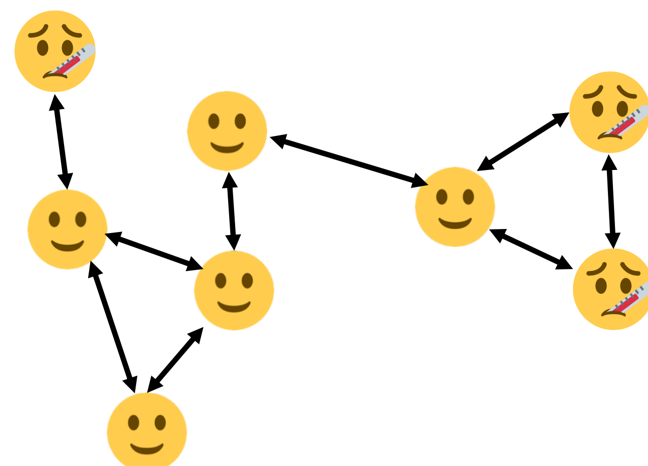
Marks

Example III: Suppressing epidemics

Agent

Health policy
(Resource allocation)

Environment



Population (social network)

Who to treat and when to reduce infections?

$\lambda_i(t) \longrightarrow N_i(t)$

Design (optimal)
treatment intensities

Marks

Stochastic optimal control of SDEs with jumps

If the problem dynamics can be expressed using SDEs with jumps:

➡ **Optimal control of marked temporal point processes**

- → **HJB equation** [Zarezade et al., 2017, 2018; Tabibian et al., 2017; Kim et al. 2018; Wang et al., 2018]
- → **Variational inference** [Wang et al., 2017]

Key idea:

Policy is characterized by an intensity function!

Stochastic optimal control of SDEs with jumps

If the problem dynamics can be expressed using SDEs with jumps:

➡ Optimal control of marked temporal po

→ m et al. 2018;

→

Next, details on one approach to the when to post problem

Key idea:

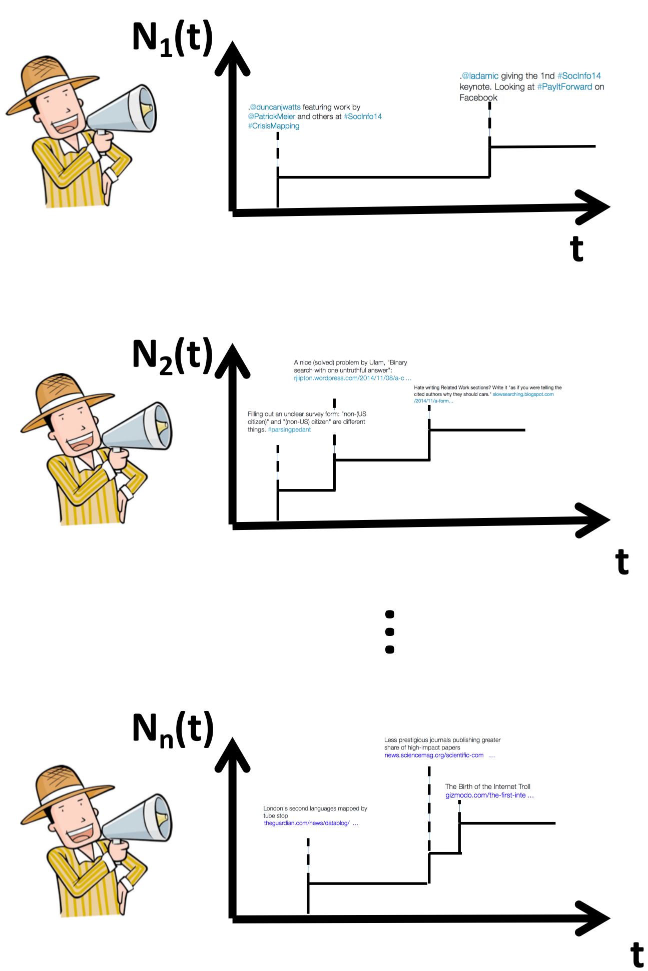
Policy is characterized by an intensity function!

Strategy to solve the when-to-post problem

Data representation	→	**Visibility and feed dynamics**	→	**Optimizing visibility**	→	**Experiments**
↕		↕		↕		↕
Temporal point processes		System of stochastic equations with jumps		Optimal control of jumps		Twitter

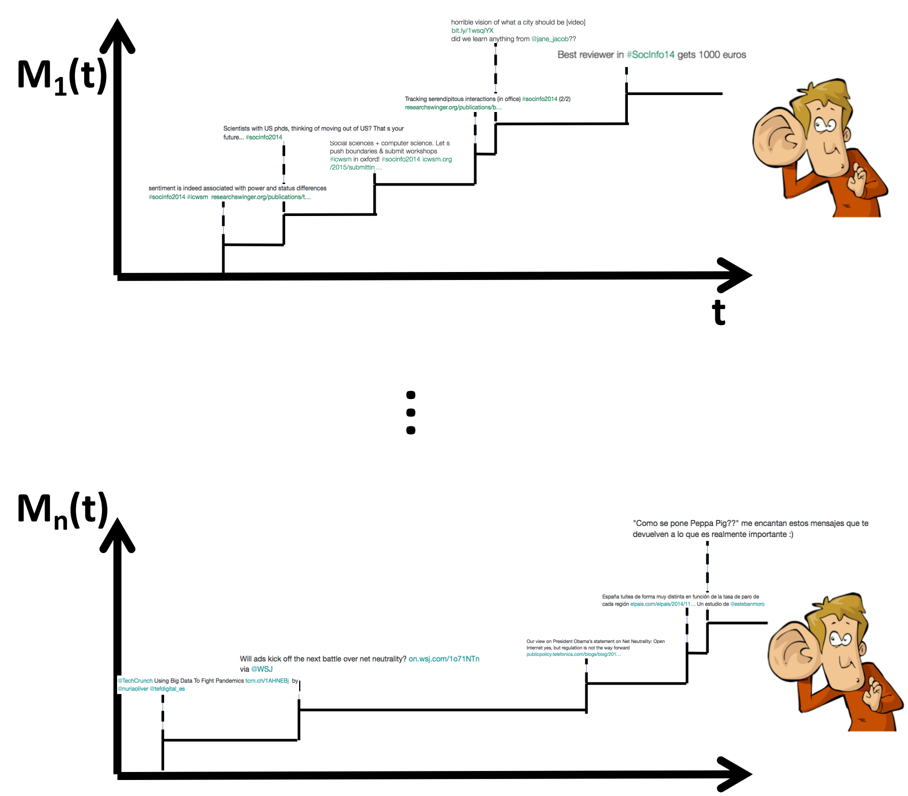
Representation of broadcasters and feeds

Broadcasters' posts as a counting process N(t)

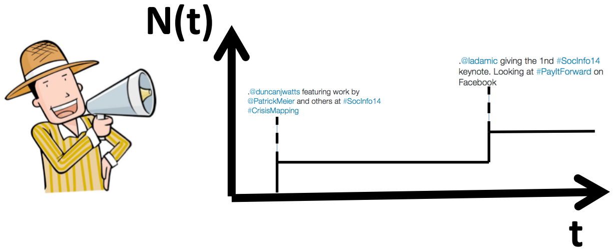


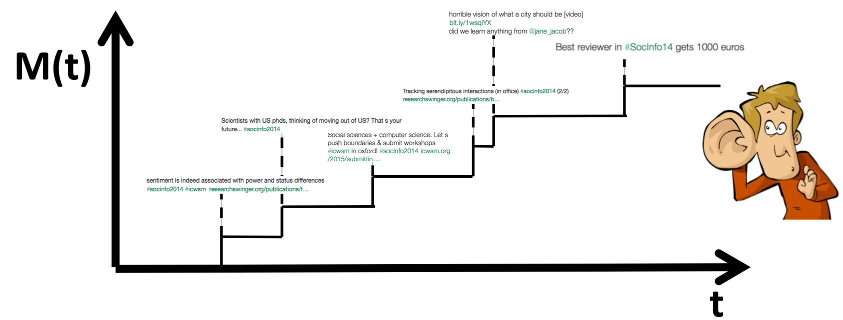
$$M(t) = A^T N(t)$$

Users' feeds as sum of counting processes M(t)



Broadcasters and feeds





$$\mathbb{E}[d\boldsymbol{N}(t)|\mathcal{H}(t)] = \underbrace{\boldsymbol{\mu}(t)}\,dt \qquad \Rightarrow \qquad \mathbb{E}[d\boldsymbol{M}(t)|\mathcal{H}(t)] = \underbrace{\boldsymbol{\gamma}(t)}\,dt$$

Policy → **Broadcaster intensity function (tweets / hour)**

$$A^T \boldsymbol{\mu}(t)$$

Feed intensity function (tweets / hour)

Given a broadcaster i and her followers

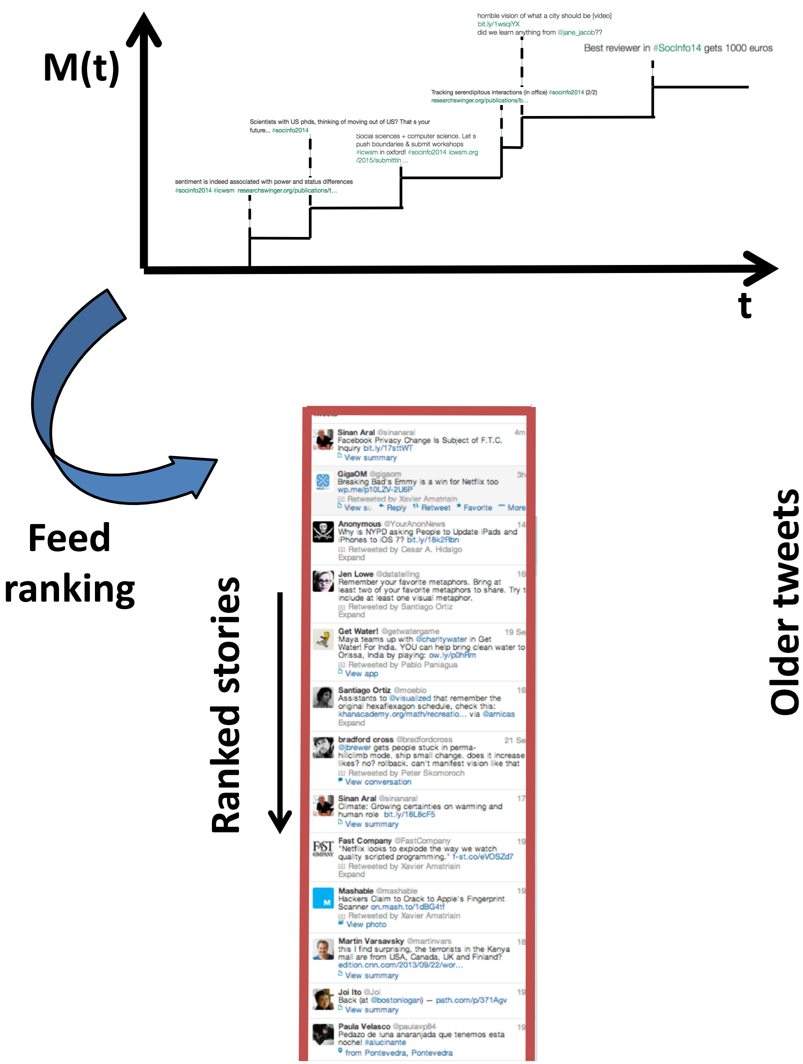
$$\boldsymbol{M}_{\backslash i}(t) = A^T \boldsymbol{N}(t) - A_i N_i(t)$$

$$\gamma_{j \backslash i}(t) = \gamma_j(t) - \mu_i(t)$$

Feed due to other broadcasters

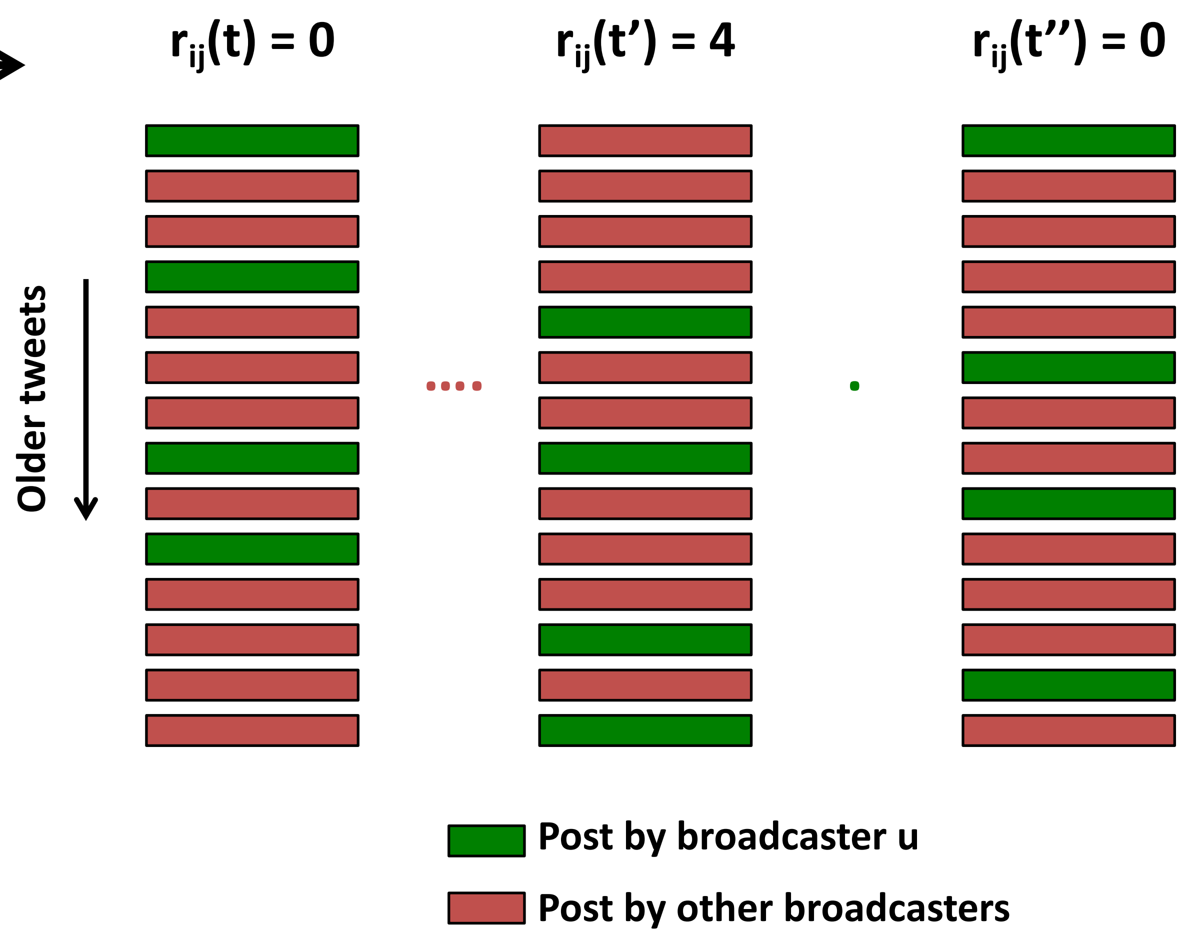
[Zarezade et al., 2017 & 2018]

Definition of visibility function



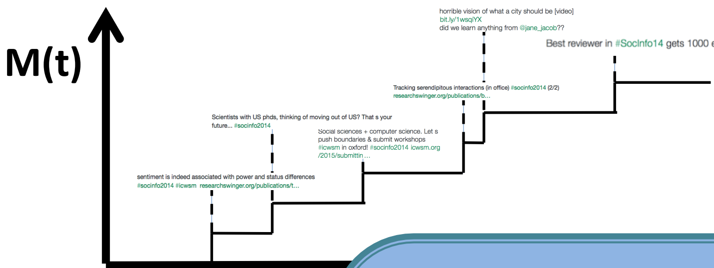
Visibility of broadcaster i at follower j

Position of the highest ranked tweet by broadcaster i in follower j's wall



[Zarezade et al., 2017 & 2018]

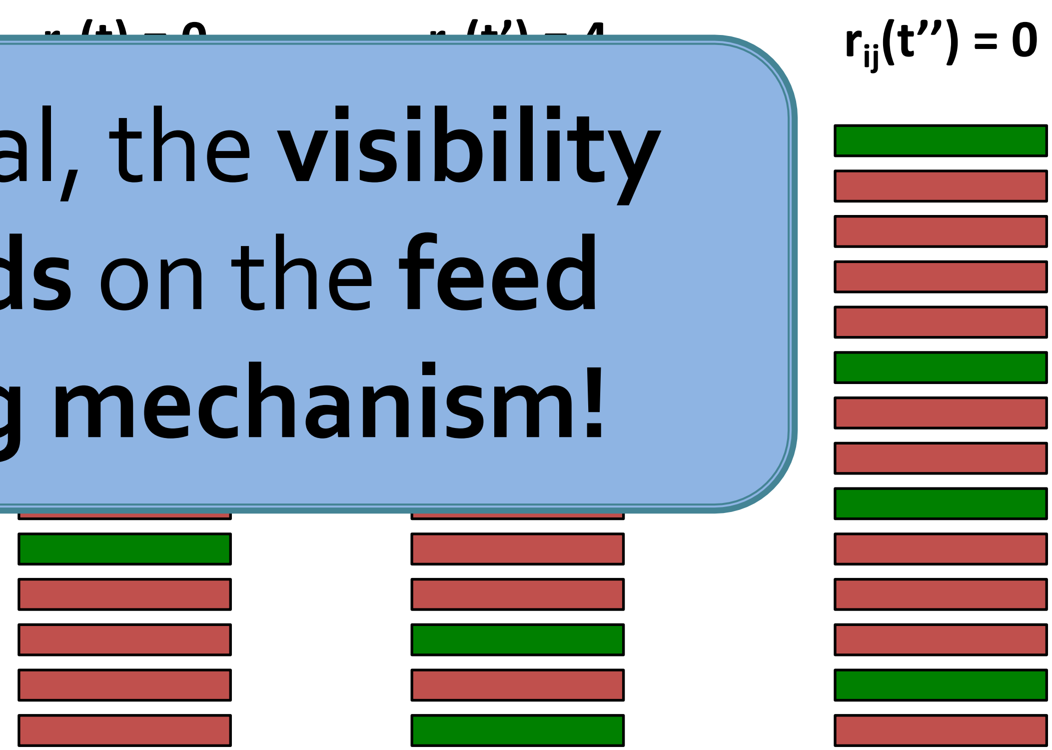
Definition of visibility function



Feed ranking

Visibility of broadcaster i at follower j

Position of the highest ranked tweet by broadcaster i in follower j's wall



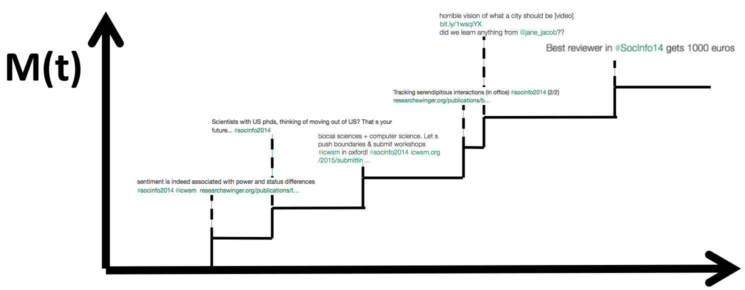
In general, the **visibility depends** on the **feed ranking mechanism!**

Post by broadcaster u

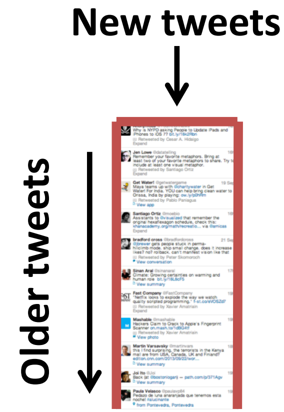
Post by other broadcasters

[Zarezade et al., 2017 & 2018]

Visibility dynamics in a FIFO feed (I)

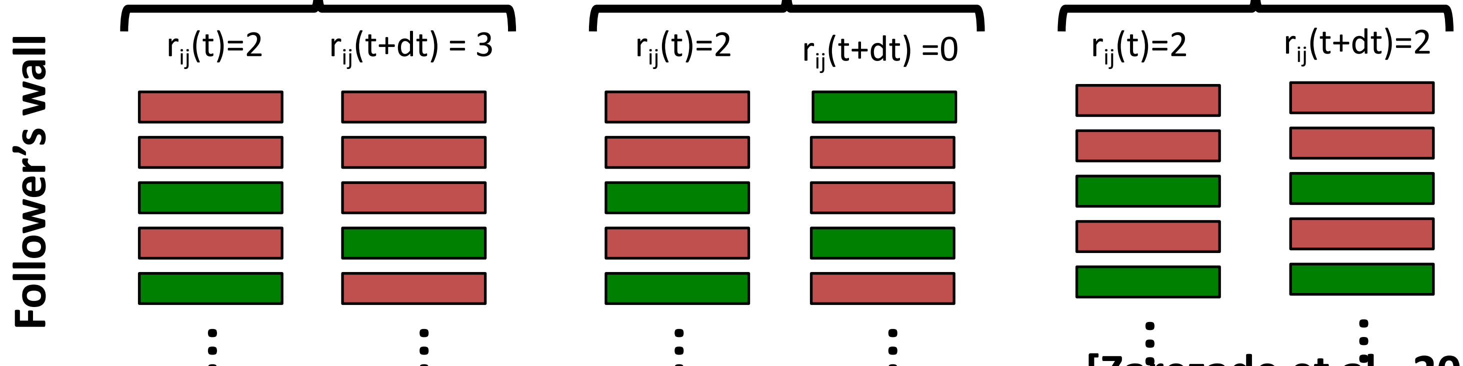


Reverse chronological order



$$\underbrace{r_{ij}(t+dt) =}_{\text{Rank at t+dt}} \underbrace{(r_{ij}(t)+1)dM_{j\setminus i}(t)(1-dN_i(t))+}_{\text{Other broadcasters post a story and broadcaster i does not post}} \; 0 + \underbrace{r_{ij}(t)(1-dM_{j\setminus i}(t))(1-dN_i(t))}_{\text{Nobody posts a story}}$$

Broadcaster i posts a story and other broadcasters do not post



[Zarezade et al., 2017 & 2018]

Visibility dynamics in a FIFO feed (II)

$$r_{ij}(t+dt) = \quad (r_{ij}(t)+1)dM_{j\setminus i}(t)(1-dN_i(t)) + \; 0 + \; r_{ij}(t)(1-dM_{j\setminus i}(t))(1-dN_i(t))$$

Zero-one law $dN_i(t)dM_{j\setminus i}(t) = 0$

$$dr_{ij}(t) = -r_{ij}(t)\,dN_i(t) + dM_{j\setminus i}(t)$$

$r_{ij}(t+dt) - r_{ij}(t)$

Broadcaster i posts a story

Other broadcasters posts a story

Stochastic differential equation (SDE) with jumps

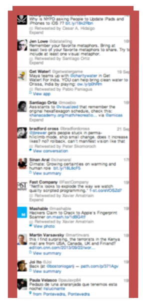
OUR GOAL:

Optimize $r_{ij}(t)$ over time, so that it is small, by controlling $dN_i(t)$ through the intensity $\mu_i(t)$

[Zarezade et al., 2017 & 2018]

Feed dynamics

$N_i(t)$ $M_{j\setminus i}(t) = M(t) \rightarrow \gamma_{j\setminus i}(t) = \lambda(t)$



We consider a **general intensity:**

(e.g. Hawkes, inhomogeneous Poisson)

$$\lambda^*(t) = \underbrace{\lambda_0(t)}_{\substack{\textbf{Deterministic}\\ \textbf{arbitrary intensity}}} + \underbrace{\alpha \int_0^t g(t-s)dN(s)}_{\substack{\textbf{Stochastic}\\ \textbf{self-excitation}}}$$

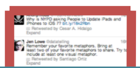
Jump stochastic differential equation (SDE)

$$\left\{ d\lambda^*(t) = \left[\lambda_0'(t) + w\lambda_0(t) - w\lambda^*(t)\right] dt + \alpha dN_i(t) \right.$$

[Zarezade et al., 2017 & 2018]

Feed dynamics

$N_i(t)$ $\quad M_{j\backslash i}(t) = M(t) \rightarrow \gamma_{j\backslash i}(t) = \lambda(t)$



Surprisingly, we will **not** have to estimate the intensity to optimize visibility!

We con

general

$N(s)$

(e.g. Hawkes, inhomogeneous Poisson)

Deterministic arbitrary intensity

Stochastic self-excitation

Jump stochastic differential equation (SDE)

$$\left\{ d\lambda^*(t) = [\lambda_0'(t) + w\lambda_0(t) - w\lambda^*(t)]\, dt + \alpha dN_i(t) \right.$$

[Zarezade et al., 2017 & 2018]

The when-to-post problem

$\mu_i(t) = u(t) \rightarrow N_i(t)$

$N_i(t)$ $M_{j\setminus i}(t)$

$dr_{ij}(t) = -r_{ij}(t)\, dN_i(t) + dM_{j\setminus i}(t)$

$N_i(t)$ $M_{j\setminus i}(t)$

$dr_{ij}(t) = -r_{ij}(t)\, dN_i(t) + dM_{j\setminus i}(t)$

...

$N_i(t)$ $M_{j\setminus i}(t)$

$dr_{ij}(t) = -r_{ij}(t)\, dN_i(t) + dM_{j\setminus i}(t)$

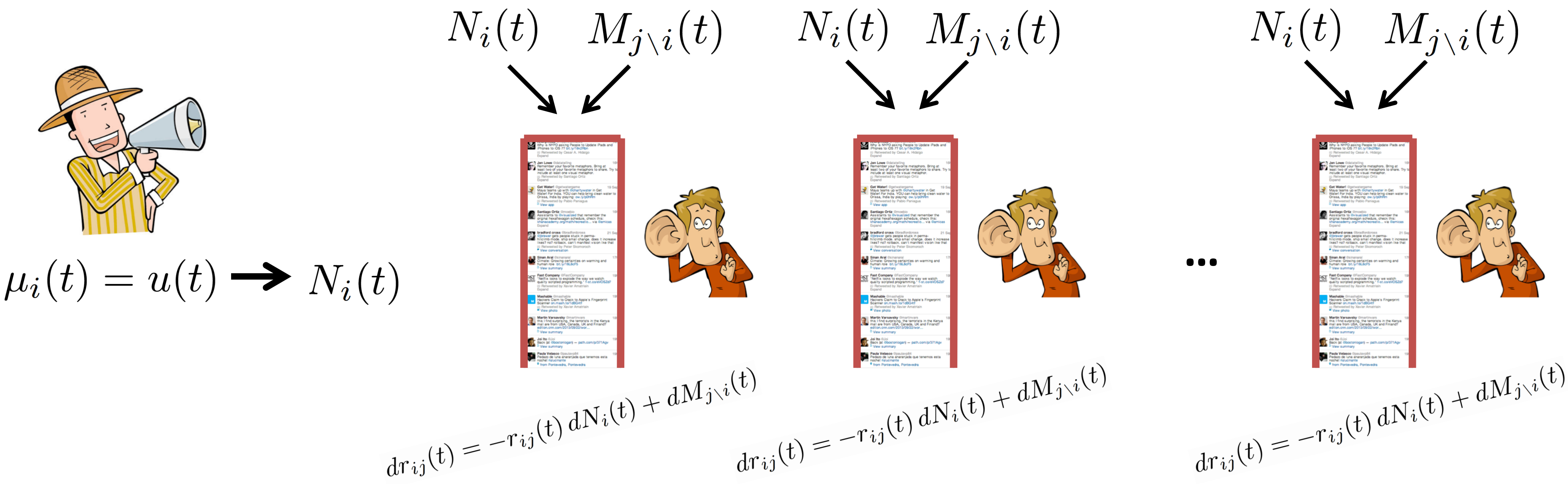
Terminal penalty

$$\underset{u(t_0,t_f]}{\text{minimize}} \quad \mathbb{E}_{(N_i,\boldsymbol{M}_{\setminus i})(t_0,t_f]}\left[\overbrace{\phi(\boldsymbol{r}(t_f))}^{\text{Terminal penalty}} + \int_{t_0}^{t_f} \underbrace{\ell(\boldsymbol{r}(\tau), u(\tau))}_{\text{Nondecreasing loss}} d\tau\right]$$

$$\text{subject to} \quad u(t) \geq 0 \quad \forall t \in (t_0, t_f],$$

Nondecreasing loss
on the visibility and the
broadcaster's intensity

The when-to-post problem



Optimization problem

$$\begin{aligned} &\underset{u(t_0,t_f]}{\text{minimize}} \quad \mathbb{E}_{(N_i,\boldsymbol{M}_{\setminus i})(t_0,t_f]}\left[\overbrace{\phi(\boldsymbol{r}(t_f))}^{\text{Terminal penalty}} + \int_{t_0}^{t_f} \overbrace{\ell(\boldsymbol{r}(\tau), u(\tau))}^{\text{Nondecreasing loss}} d\tau\right] \\ &\text{subject to} \quad u(t) \geq 0 \quad \forall t \in (t_0, t_f], \end{aligned}$$

Dynamics defined by Jump SDEs

$$dr(t) = -r(t)\, dN(t) + dM(t)$$

$$d\lambda(t) = [\lambda_0'(t) + w\lambda_0(t) - w\lambda(t)]\, dt + \alpha\, dM(t)$$

[Zarezade et al., 2017 & 2018]

When-to-post for a single follower

Optimization problem

$$\begin{cases} \underset{u(t_0,t_f]}{\text{minimize}} \quad \mathbb{E}_{(N,M)(t_0,t_f]}\left[\phi(r(t_f)) + \int_{t_0}^{t_f} \ell(r(\tau),u(\tau))\,d\tau\right] \\ \text{subject to} \quad u(t) \geq 0 \quad \forall t \in (t_0,t_f], \end{cases}$$

Dynamics defined by Jump SDEs

$$\begin{cases} dr(t) = -r(t)\,dN(t) + dM(t) \\ d\lambda(t) = [\lambda_0'(t) + w\lambda_0(t) - w\lambda(t)]\,dt + \alpha\,dM(t) \end{cases}$$

To solve the optimization problem, we first define the **optimal cost-to-go:**

$$J(r(t),\lambda(t),t) = \min_{u(t,t_f]} \mathbb{E}_{(N,M)(t,t_f]}\left[\phi(r(t_f)) + \int_{t}^{t_f} \ell(r(\tau),u(\tau))\,d\tau\right]$$

The cost-to-go, evaluated at t_0, recovers the optimization problem!

Bellman's Principle of Optimality

Lemma. The **optimal cost-to-go** satisfies **Bellman's Principle of Optimality**

$$J(r(t), \lambda(t), t) = \min_{u(t,t+dt]} \mathbb{E}\left[J(r(t+dt), \lambda(t+dt), t+dt)\right] + \ell(r(t), u(t))\, dt$$

Proof sketch

$$\begin{aligned}
J(\gamma(t), r(t), t) &= \min_{u(t,t_f]} \mathbb{E}_{(N,M)(t,t_f]}\left[\phi(r(t_f)) + \int_t^{t_f} \ell(r(\tau), u(\tau))\, d\tau\right] \\
&= \min_{u(t,t_f]} \mathbb{E}_{(N,M)(t,t_f]}\left[\phi(r(t_f)) + \int_t^{t+dt} \ell(r(\tau), u(\tau))\, d\tau + \int_{t+dt}^{t_f} \ell(r(\tau), u(\tau))\, d\tau\right] \\
&= \min_{u(t,t_f]} \mathbb{E}_{(N,M)(t,t+dt]}\left[\mathbb{E}_{(N,M)(t+dt,t_f]}\left[\phi(r(t_f)) + \ell(t,r,u)\, dt + \int_{t+dt}^{t_f} \ell(r(\tau), u(\tau))\, d\tau\right]\right] \\
&= \min_{u(t,t+dt]} \min_{u(t+dt,t_f]} \mathbb{E}_{(N,M)(t,t+dt]}\left[\ell(r(t), \gamma(t), t)\, dt + \mathbb{E}_{(N,M)(t+dt,t_f]}\left[\phi(r(t_f)) + \int_{t+dt}^{t_f} \ell(r(\tau), u(\tau))\, d\tau\right]\right] \\
&= \min_{u(t,t+dt]} \mathbb{E}_{(N,M)(t,t+dt]}\left[J(\gamma(t+dt), r(t+dt), t+dt)\right] + \ell(r(t), u(t))\, dt.
\end{aligned}$$

[Zarezade et al., 2017 & 2018]

The Hamilton-Jacobi-Bellman (HJB) equation (I)

Bellman's Principle of Optimality

$$J(r(t), \lambda(t), t) = \min_{u(t,t+dt]} \mathbb{E}\left[J(r(t+dt), \lambda(t+dt), t+dt)\right] + \ell(r(t), u(t))\, dt$$

$J(r(t+dt), \lambda(t+dt), t+dt) = J(r(t), \lambda(t), t) + dJ(r(t), \lambda(t), t)$

$$0 = \min_{u(t,t+dt]} \mathbb{E}\left[dJ(r(t), \lambda(t), t)\right] + \ell(r(t), u(t))\, dt$$

$$dr(t) = -r(t)\, dN(t) + dM(t)$$
$$d\lambda(t) = [\lambda_0'(t) + w\lambda_0(t) - w\lambda(t)]\, dt + \alpha\, dM(t)$$

Hamilton-Jacobi-Bellman (HJB) equation

Partial differential equation in J (with respect to r, λ and t)

[Zarezade et al., 2017 & 2018]

The Hamilton-Jacobi-Bellman (HJB) equation

$$0 = \min_{u(t,t+dt]} \mathbb{E}\left[dJ(r(t), \lambda(t), t)\right] + \ell(r(t), u(t))\, dt$$

$$dr(t) = -r(t)\, dN(t) + dM(t)$$
$$d\lambda(t) = [\lambda_0'(t) + w\lambda_0(t) - w\lambda(t)]\, dt + \alpha\, dM(t)$$

$$\begin{aligned}0 = {} & J_t(r(t), \lambda(t), t) + [\lambda_0'(t) + w\lambda_0(t) - w\lambda(t)]\, J_\lambda(r(t), \lambda(t), t) + [J(r(t)+1, \lambda(t)+\alpha, t) - J(r(t), \lambda(t), t)]\lambda(t) \\ & + \underbrace{\min_{u(t,t+dt]} \ell(r(t), u(t)) + [J(0, \lambda(t), t) - J(r(t), \lambda(t), t)]u(t).}_{u^*(t) = q^{-1}\left[J(r(t), \lambda(t), t) - J(0, \lambda(t), t)\right]}\end{aligned}$$

$$\begin{aligned}0 = {} & J_t(r(t), \lambda(t), t) + [\lambda_0'(t) + w\lambda_0(t) - w\lambda(t)]\, J_\lambda(r(t), \lambda(t), t) + [J(r(t)+1, \lambda(t)+\alpha, t) - J(r(t), \lambda(t), t)]\lambda(t) \\ & + \frac{1}{2}s(t)\, r^2(t) - \frac{1}{2}q^{-1}\left[J(r(t), \lambda(t), t) - J(0, \lambda(t), t)\right]^2\end{aligned}$$

[Zarezade et al., 2017 & 2018]

Solving the HJB equation

Consider a quadratic loss

$$\ell(r(t), u(t)) = \frac{1}{2} s(t)\, r^2(t) + \frac{1}{2} q\, u^2(t)$$

Favors some periods of times (e.g., times in which the follower is online)

Trade-offs visibility and number of broadcasted posts

Then, it can be shown that the **optimal cost-to-go** is given by:

$$J(r(t), \lambda(t), t) = f(t) + \sqrt{s(t)/q}\, r(t) + \sum_{j=1}^{m} g_j(t) \lambda^j(t)$$

[Zarezade et al., 2017 & 2018]

Solving the HJB equation

Given the cost

$$J(r(t), \lambda(t), t) = f(t) + \sqrt{s(t)/q}\, r(t) + \sum_{j=1}^{m} g_j(t) \lambda^j(t)$$

Then, we can readily compute the optimal intensity:

$$u^*(t) = q^{-1} \left[J(r(t), \lambda(t), t) - J(0, \lambda(t), t) \right]$$
$$= \sqrt{s(t)/q}\, r(t)$$

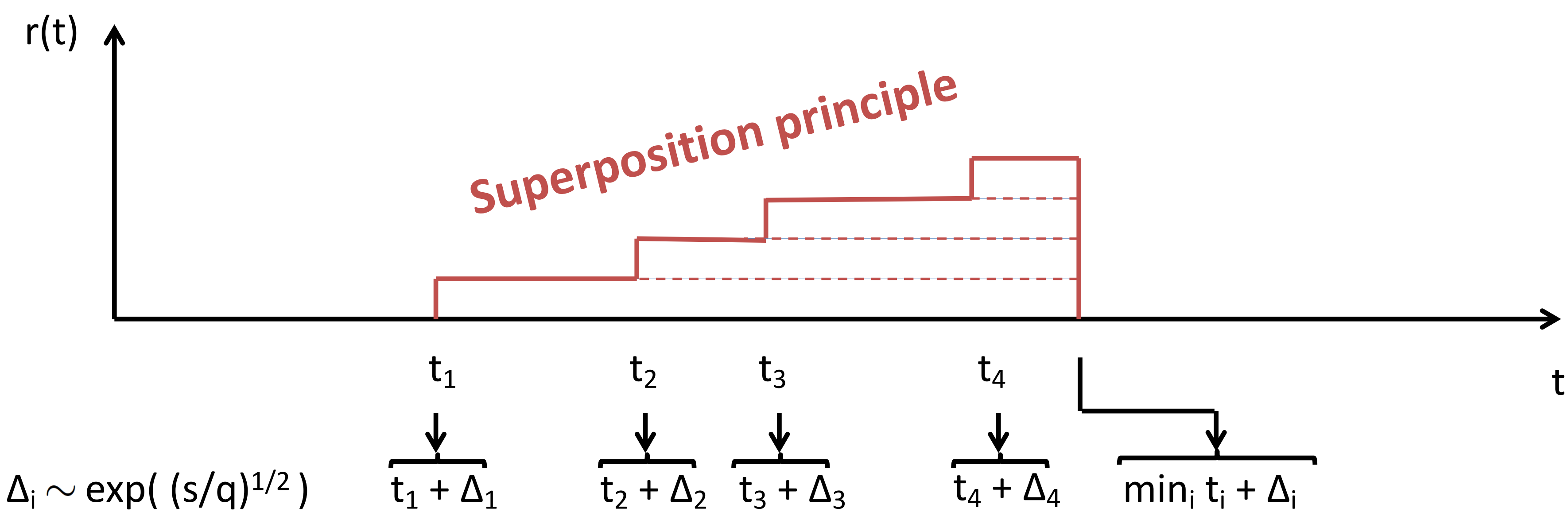
It only depends on the current visibility!

[Zarezade et al., 2017 & 2018]

The RedQueen algorithm

Consider $s(t) = s \longrightarrow u^*(t) = (s/q)^{1/2}\, r(t)$

How do we sample the next time?



It only requires sampling $M(t_f)$ times!

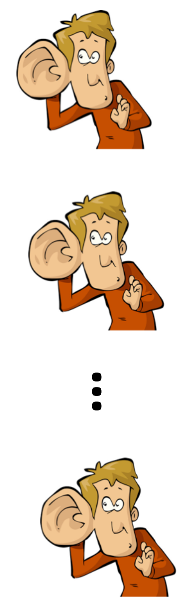
[Zarezade et al., 2017 & 2018]

Experiments on real data

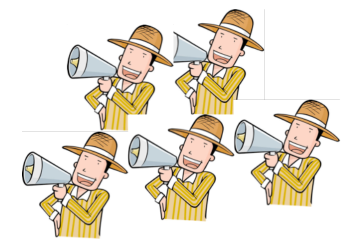
Consider 2,000 broadcasters (users) from **Twitter**

For each broadcaster:

Collect all his posted tweets during 2 months



Track down his followers

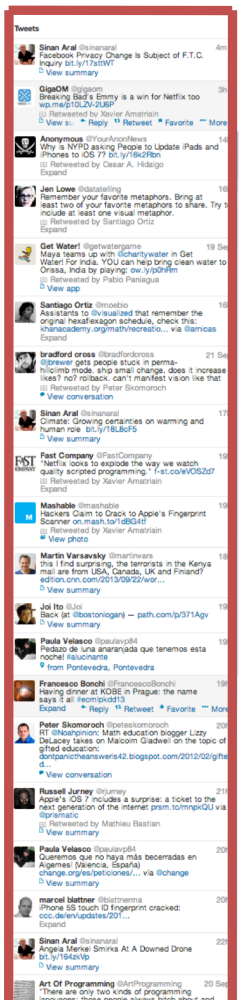


Collect the other broadcasters' posted tweets during 2 months

Millions of users!

Experimental setup on real data

Experimental setup allows for a *truthful* what-if evaluation:



Playback other broadcasters' tweets on a held-out set

Tweet according to *optimal* intensity and compute visibility over time

Fit other broadcasters' intensities

$\gamma_{v \backslash u}(t)$

Find intensity

$\lambda_u(t) = \lambda(t)$

Needed for state-of-the-art Karimi's method

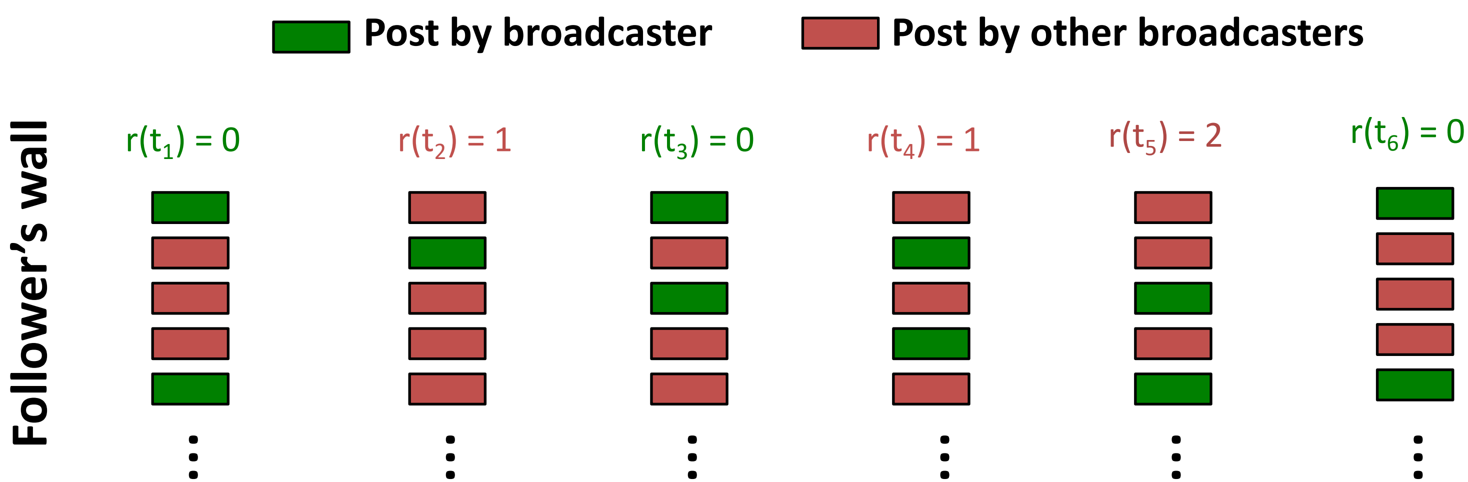
Evaluation metrics

Visibility over time

$$\int_0^T r(t)dt$$

Time at the top

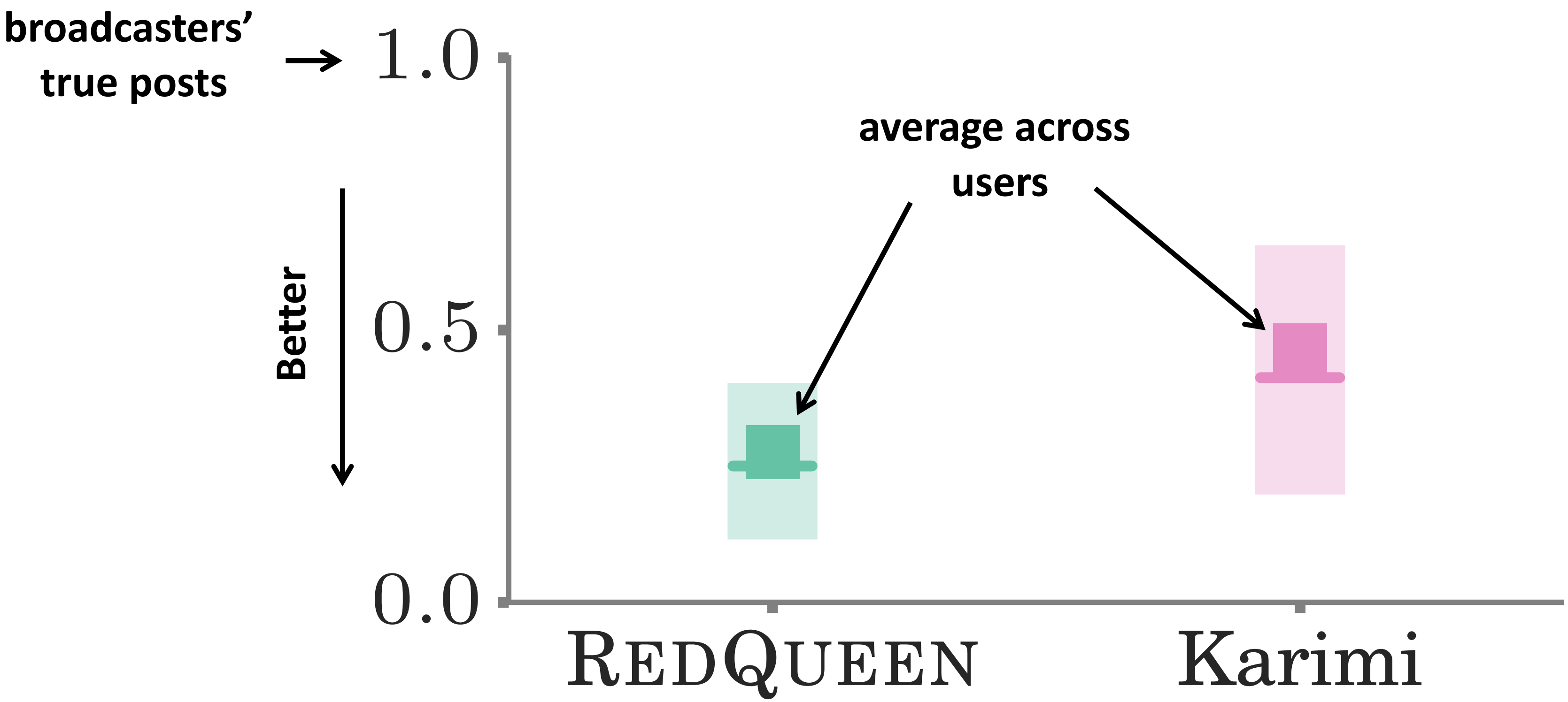
$$\int_0^T \mathbb{I}(r(t)<1)dt$$



Position over time = $0 \times (t_2 - t_1) + 1 \times (t_3 - t_2) + 0 \times (t_4 - t_3) + 1 \times (t_5 - t_4) + 2 \times (t_6 - t_5)$

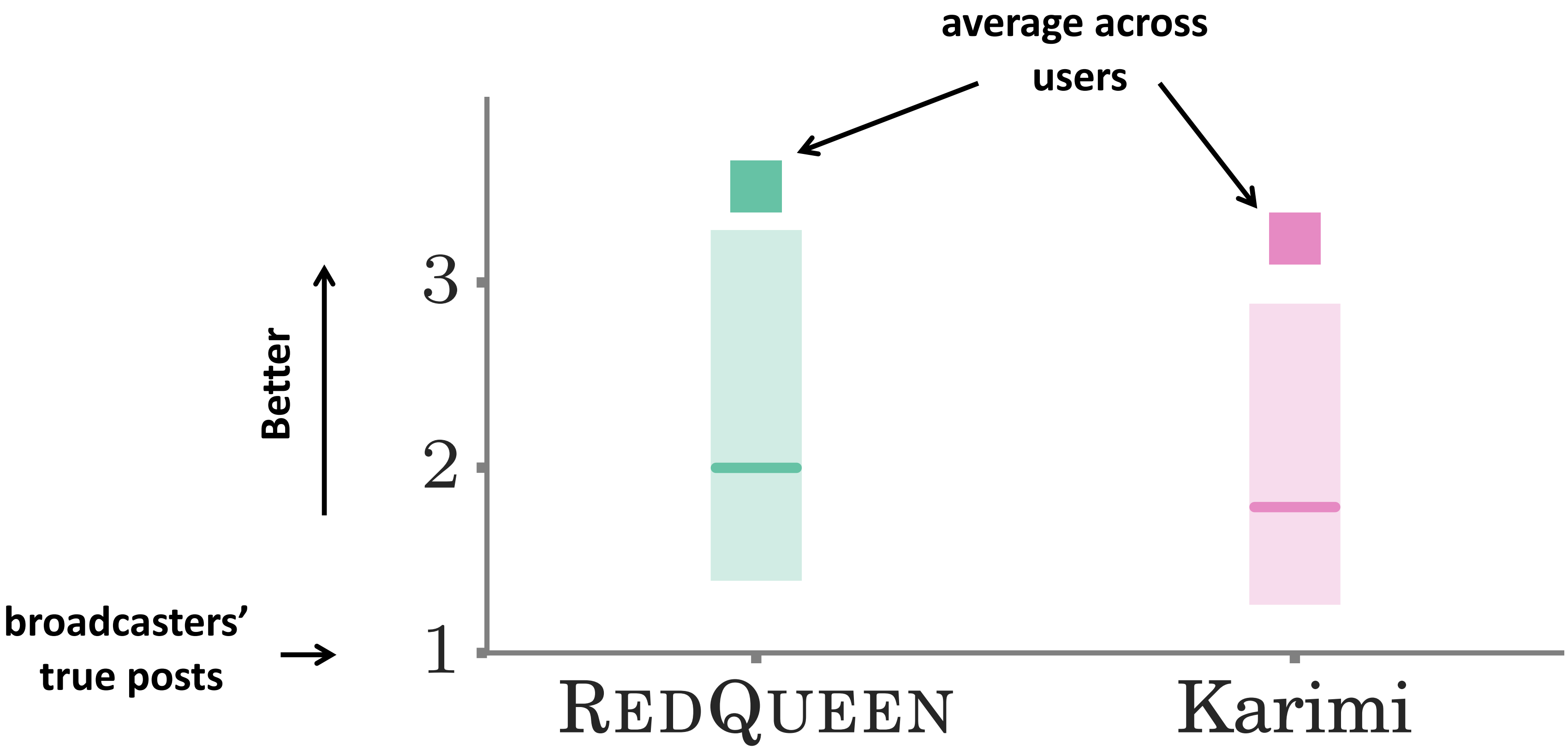
Time at the top = $(t_2 - t_1) + 0 + (t_4 - t_3) + 0 + 0$

Position over time



It achieves **(i) 0.28x lower average position, in average, than the broadcasters' true posts** and **(ii) lower average position for 100% of the users.**

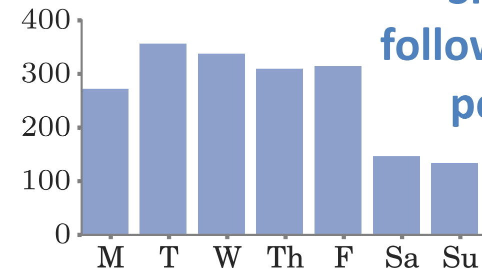
Time at the top



It achieves **(i) 3.5x higher time at the top, in average, than the broadcasters' true posts** and **(ii) higher time at the top for 99.1% of the users.**

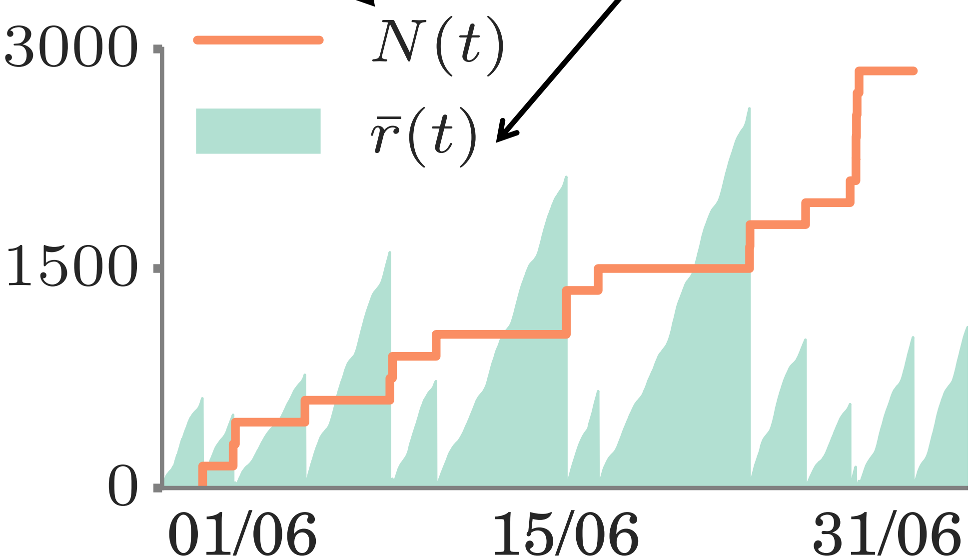
Example: a broadcaster in Twitter

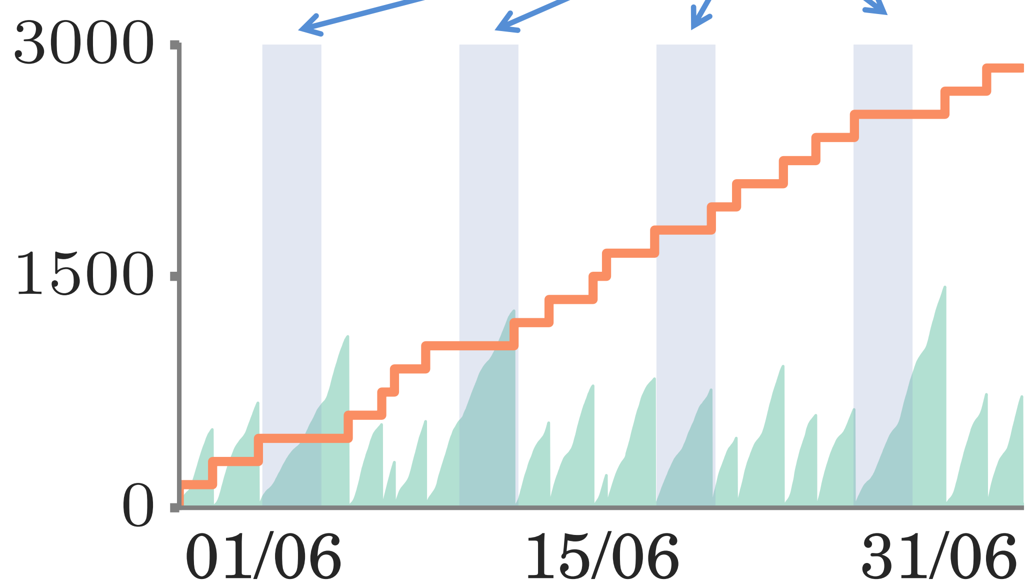
Significance: followers' retweets per weekday



Broadcaster's posts

Average position over time





True posts

$$\frac{1}{T}\int_0^T \bar{r}(t)dt = 698.04$$

40% lower!

REDQUEEN

$$\frac{1}{T}\int_0^T \bar{r}(t)dt = 425.25$$

[Zarezade et al., 2017 & 2018]

Why RedQueen?

“Now, here, you see, it takes all the running you can do, to keep in the same place”

Through the Looking-Glass, Lewis Carroll

more at

learning.mpi-sws.org