Information Interaction Profile of Choice Adoption

G. Poux-Médard, J. Velcin, S. Loudcher

Gaël Poux-Médard

Université de Lyon
Lyon 2, ERIC UR 3083

September 13, 2021

1. Introduction
2. Background
3. InterRate
4. Experiments
5. Conclusion
6. References

1 Introduction

2 Background

3 InterRate

4 Experiments

5 Conclusion

6 References

Problem and definitions

- “Information Interaction Profile of Choice Adoption”
  - Information : a tweet, a situation, an ad, etc.
  - Choice adoption : a retweet, a reaction, a click, etc.
  - Interaction profile : how the exposure to a first piece of information affect a user’s choice on a later one according to time.

Problematic

- We are looking for trends in information interaction : how does it evolve with time, and does it take place between every piece of information ?

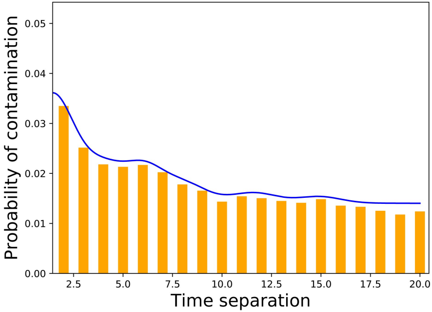


Figure 1 – Interaction profile – Probability of clicking a Twitter URL after seeing a Migre URL

1 Introduction

2 Background

3 InterRate

4 Experiments

5 Conclusion

6 References

Previous works

- Info-info interaction :
  - Clash of the Contagions [1] – Models interactions as modulations over a pre-computed virality ; considers each time interval separately ; SGD
  - IMMSBM [2] – Models symmetric interaction with no assumption on virality ; does not consider time ; EM
- Other types of interaction :
  - Information overload [3]
  - Users' attention [4]
  - Topic-dependent networks [5]

1 Introduction

2 Background

3 InterRate

4 Experiments

5 Conclusion

6 References

The data

- Users are exposed to pieces of information (light squares).
- Sometimes, they will act on one of them after being exposed (orange squares).
- The action takes place a time t_s after the exposure.
- The time separating exposures can be constant (order of exposition) or vary (absolute exposure times).

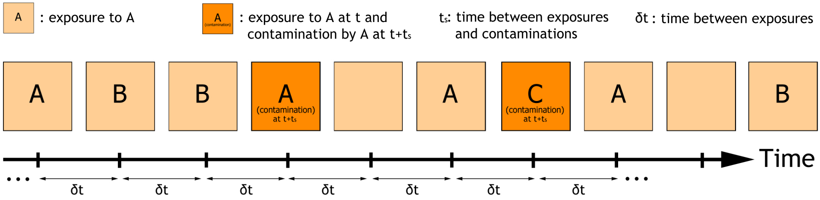


Figure 2 – Illustration of the interacting process

Assumptions

- This modeling comes with several assumptions :
  - Pieces of information appear independently from each other (unfit for survival or Hawkes modeling, in particular the system is not self-exciting)
  - A user's action is conditioned by previous exposures only (as in [1, 6]).
  - t_s is constant (i.e. the time between a read and a retweet in the case of Twitter). This hypothesis is a simplification for clarity of presentation and can be easily alleviated.

The model

- Data $\mathcal{D} \equiv \{(t_i^{(x)}, c_{t_i}^{(x)}, \mathcal{H}_i^{(x)})\}_{i,x}$ with :
  - ◇ $t_i^{(x)}$: i^{th} exposure to x at time $t_i + t_s$
  - ◇ $c_{t_i}^{(x)}$: whether the user act on the exposure
  - ◇ $\mathcal{H}_i^{(x)}$: history of exposures prior to $t_i^{(x)}$
- Log-likelihood :

$$\ell(\beta|\mathcal{D}, t_s) = \sum_{\mathcal{D}} \sum_{t_j^{(y)} \in \mathcal{H}_i^{(x)}} c_{t_i}^{(x)} \log\left(H(t_i^{(x)} + t_s | t_j^{(y)}, \beta_{xy})\right) + (1 - c_{t_j}^{(y)}) \log\left(1 - H(t_i^{(x)} + t_s | t_j^{(y)}, \beta_{xy})\right) \tag{1}$$

- $H(t_i^{(x)} + t_s | t_j^{(y)}, \beta_{xy})$ is a the instantaneous probability of action on x at time $t_i^{(x)} + t_s$ given an exposure to y at $t_j^{(y)}$.
- H is a time varying kernel function parametrized by the inforamtion interaction matrix β.

Optimization

- $\ell(\beta|\mathcal{D}, t_s)$ can be split in as many subproblems as there are pieces of information. Input data for each of them : $\mathcal{D}^{(x)} \equiv \{(\mathcal{H}_i^{(x)}, t_i^{(x)}, c_{t_i}^{(x)})\}_i$
- $\ell(\beta|\mathcal{D}, t_s)$ is convex if

$$\begin{cases} H'^2 \geq H''H \\ H'^2 \geq -H''(1-H) \\ H \in \,]0;1[\end{cases} \tag{2}$$

- A number of kernel functions satisfy convexity : exponential, Rayleigh, power-law, gaussian, linear combination of those.
- Final problem : $\min_{\beta \geq 0} -\ell(\beta|\mathcal{D}, t_s)$

1 Introduction

2 Background

3 InterRate

4 Experiments

5 Conclusion

6 References

Datasets

- Synthetic (20.000 sequences of length 50 steps)
- Twitter dataset (104,349 sequences of average length 53.5 steps, exposure are tweets and contagions are retweets) [7]
- Iterated Prisoner's dilemma dataset (2,337 sequences of average length 9.0 steps, exposures are situations, contagions are players choices) [8, 9]
- Ads dataset (87,500 sequences of average length 23.9 steps, exposures are ads, contagions are clicks on ads) [10].

Background noise

- A background noise is induced due to the data gathering process : random correlations.

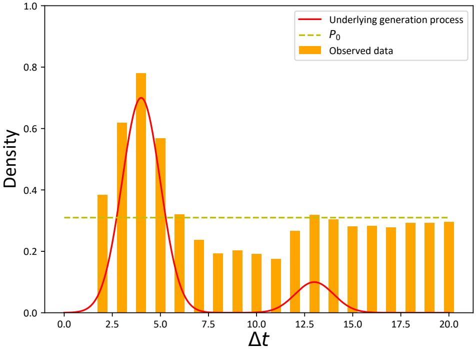


Figure 3 – Illustration of the background noise on synthetic data

Kernel choice

- We choose 2 different kernels :
  - InterRate-EXP (exponential kernel) :
    $\log H(t_i^{(x)} + t_s | t_j^{(y)}, \beta_{xy}) = -\beta_{ij}^{(bg)} - \beta_{ij}(t_i + t_s - t_j)$
  - InterRate-RBF (Gaussian kernel) :
    $\log H(t_i^{(x)} + t_s | t_j^{(y)}, \beta_{ij}) = -\beta_{ij}^{(bg)} - \sum_{s=0}^{S} \frac{\beta_{ij}^{(s)}}{2}(t_i + t_s - t_j - s)^2$
- The time-independent component of both kernels $\beta_{ij}^{(bg)}$ fits the background noise, while other parameters fit temporal dynamics.

Baselines

- ICIR – Version of InterRate without interactions : off-diagonal entries of β are null. We use a RBF kernel.
- Clash of the contagions (CoC) [1] – considers each time-bin separately
- IMMSBM [2] – does not consider temporal aspect ; enforces symmetry $\beta_{ab} = \beta_{ba}$
- Naive – Time-independent frequency of action for each pair of pieces of information.

- We split data into training and test sets (80%-20%) and perform a 5-fold cross validation. We compare the inferred and true interaction profiles using the residual sum of squares (RSS), Jensen-Shanon (JS) divergence and the best-case F1-score (BCF1).

Comparison on synthetic data

- Note that InterRate as such aims to describe the process, not to make predictions (background noise included in evaluation)

		RSS	JS div.	BCF1	MSE β
Synth-20	IR-RBF	18.4151	0.002 28	0.919	0.001
	ICIR	139.5926	0.009 98	0.827	0.016
	Naive	145.5132	0.010 38	0.822	
	CoC	123.0583	0.009 38	0.822	
	IMMSBM	222.0555	0.017 29	0.727	
Synth-5	IR-RBF	0.1169	0.000 22	0.974	0.005
	ICIR	8.2661	0.008 12	0.850	0.019
	Naive	10.0264	0.009 96	0.821	
	CoC	0.1154	0.000 20	0.976	
	IMMSBM	11.6936	0.013 62	0.769	

Figure 4 – Numerical results for synthetic data

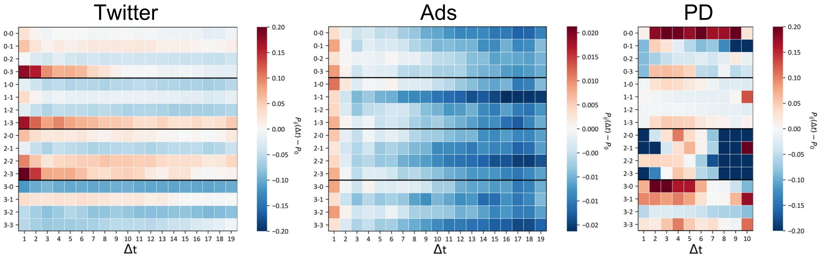
Comparison on real-world data

		RSS	JS div.	BCF1
Twitter	IR-RBF	0.0015	0.000 06	0.983
	IR-EXP	0.0011	0.000 05	0.986
	ICIR	0.0137	0.000 63	0.961
	Naive	0.0161	0.000 73	0.938
	CoC	0.0017	0.000 07	0.957
	IMMSBM	0.0147	0.000 68	0.954
PD	IR-RBF	1.1268	0.007 58	0.979
	IR-EXP	1.5526	0.008 67	0.966
	ICIR	3.5359	0.018 23	0.938
	Naive	3.6527	0.019 15	0.945
	CoC	1.2409	0.008 09	0.974
	IMMSBM	20.3773	0.087 01	0.767
Ads	IR-RBF	0.0043	0.000 04	0.981
	IR-EXP	0.0030	0.000 03	0.985
	ICIR	0.0983	0.000 85	0.966
	Naive	0.1453	0.001 26	0.913
	CoC	0.0045	0.000 05	0.974
	IMMSBM	0.0155	0.000 15	0.954

Figure 5 – Numerical results on real-world data

Visualization

- We plot the intensity of interactions in time (probability of action minus inferred background noise) for every dataset.
- We recover SotA conclusions :
    - ◇ Twitter : most informative URL interactions occur within the 3 time steps [1], and interactions are weak [1, 2]. Interactions are sparse [2].
    - ◇ Ads : globally decreasing probability of click when two exposures are distant in time [10].



1 Introduction

2 Background

3 InterRate

4 Experiments

5 Conclusion

6 References

Conclusion

- Efficient convex model to investigate the way interactions happen within various datasets.
- Interactions modeling has been little explored in data science. Unlike previous models, our method accounts for both the interaction effects and their influence over time (the interaction profile).
- Better results on synthetic and various real-world datasets
- Possible applications in recommender systems, spreading processes, human choice behavior, etc.,

1 Introduction

2 Background

3 InterRate

4 Experiments

5 Conclusion

6 References

[1] Seth Myers and Jure Leskovec.
Clash of the contagions : Cooperation and competition in information diffusion.
IEEE, pages 539–548, 2012.

[2] Gael Poux-Médard, Julien Velcin, and Sabine Loudcher.
Interactions in information spread : quantification and interpretation using stochastic block models.
RecSys, 2021.

[3] Manuel Rodriguez, Krishna P. Gummadi, and Bernhard SchÃűlkopf.
Quantifying information overload in social media and its impact on social contagions.
ICWSM, 03 2014.

[4] Weng L., Flammini A., and al.
Competition among memes in a world with limited attention.
Nature Sci Rep 2, 335, 2012.

[5] Nan Du, Le Song, Hyenkyun Woo, and Hongyuan Zha.
Uncover topic-sensitive information diffusion networks.
In *AISTATS*, volume 31 of *JMLR Workshop and Conference Proceedings*, pages 229–237. JMLR.org, 2013.

[6] Ali Zarezade, Ali Khodadadi, Mehrdad Farajtabar, Hamid R. Rabiee, and Hongyuan Zha.
Correlated cascades : Compete or cooperate.
AAAI, 2017.

[7] Nathan O. Hodas and Kristina Lerman.
The simple rules of social contagion.
Scientific Reports, 4(4343), 2014.

[8] John J. Nay and Yevgeniy Vorobeychik.
Predicting human cooperation.
PLoS One, 11(5), 2016.

[9] Y. Bereby-Meyer and A. E. Roth.
The speed of learning in noisy games : Partial reinforcement and the sustainability of cooperation.
American Economic Review, 96(4) :1029–1042, 2006.

[10] Junyu Cao and Wei Sun.
Sequential choice bandits : learning with marketing fatigue.
AAAI-19, 2019.

Thanks for your attention !