

Research vision statement – Marian-Andrei RizoIU

Online social media is increasingly prevalent in shaping “offline” events, ranging from Twitter’s role in the 2016 US presidential campaign to the recent allegations that Facebook was used to convey hateful and racist messages towards the Rohingya minority in Myanmar. My research aims to tackle the grand challenge of assessing and mitigating the impact of online social media on the fundamental processes of our society. By successfully modelling and predicting the dynamics of online human attention, the benefit of my research is to understand novel societal and economic “offline” phenomena, like the spread of misinformation and the role of social bots in recent political elections, the link between the spread of hateful messages and violence towards minorities, or the emergence of disruptive business models like Uber or Airbnb.

Aims. My research project aims to link the dynamics of collective human attention to the individual actions of the users of online platforms. I intend to develop tools and methods to:

- **model** the links between a) user-level behaviour and their interactions and b) collective-level attention, which is currently not adequately understood;
- **explain** novel societal phenomena, such as the role and influence of socialbots in the democratic process, or the virality of malicious content;
- **enable addressing** some of these societal challenges, for example by constructing early detection systems for "fake news", assessing information reliability, or detecting and tracking the flow of radical messages in the online.

Tools, novelty and challenges. From a technical point of view, my main research interest lies in developing machine learning and large-scale data mining methods to model, understand and improve the functioning of large social and information systems. The above-stated objectives typically require conceptual innovations and technical breakthroughs along three different dimensions:

- **data acquisition:** Online social data is abundant, however, it is also noisy, and it comes in incompatible formats across different platforms or aggregated for privacy reasons. This introduces difficulties in scalability, modelling with missing or false data, as well as ethical concerns. My experience of more than 5 years of working with real data at the intersection of multiple social and information platforms [1, 2, 3, 4, 5, 6], places me in a privileged position to address this challenge;
- **interpretable modelling:** Accounting for individual-level actions and predicting collective-level measures can be seen as different problems, which pose difficulties for traditional machine learning tools. Stochastic point processes are the current state-of-the-art interpretable data-driven models which can account for large social and information systems, and for the dynamic processes that take place over them. In my previous work, I have developed event-level [1] and volumes formulations [6] for stochastic processes, and I have linked two classes of stochastic processes previously considered independent one from another [3];
- **aggregate measures:** Linking online user actions to social phenomena often entails uncovering previously unexplored connections between different scientific fields. For example, linking Twitter diffusions to the influence of users in political discussions [2] resulted in a multi-disciplinary collaboration with social scientists. Modelling dynamic social processes using stochastic processes allows tapping into the rich stochastic literature to build such aggregate measures. I have previously used such techniques to link diffusions to content popularity [6], content promotability [4] and to explain popularity unpredictability [2].

REFERENCES

- [1] Swapnil Mishra, **Marian-Andrei RizoIU**, and Lexing Xie. Feature Driven and Point Process Approaches for Popularity Prediction. In *CIKM '16*, pages 1069–1078, 2016.
- [2] **Marian-Andrei RizoIU**, Timothy Graham, Rui Zhang, Yifei Zhang, Robert Ackland, and Lexing Xie. #DebateNight: The Role and Influence of Socialbots on Twitter During the 1st 2016 U.S. Presidential Debate. In *International AAI Conference on Web and Social Media (ICWSM '18)*, pages 1–10, Stanford, CA, USA, 2018.
- [3] **Marian-Andrei RizoIU**, Swapnil Mishra, Quyu Kong, Mark Carman, and Lexing Xie. SIR-Hawkes: Linking Epidemic Models and Hawkes Processes to Model Diffusions in Finite Populations. In *WWW '18*, pages 1–16, Lyon, France, nov 2018.
- [4] **Marian-Andrei RizoIU** and Lexing Xie. Online Popularity under Promotion: Viral Potential, Forecasting, and the Economics of Time. In *ICWSM '17*, pages 182–191, Montréal, Québec, Canada, 2017.
- [5] **Marian-Andrei RizoIU**, Lexing Xie, Tiberio Caetano, and Manuel Cebrian. Evolution of Privacy Loss in Wikipedia. In *WSDM '16*, pages 215–224, dec 2016.
- [6] **Marian-Andrei RizoIU**, Lexing Xie, Scott Sanner, Manuel Cebrian, Honglin Yu, and Pascal Van Hentenryck. Expecting to be HIP: Hawkes Intensity Processes for Social Media Popularity. In *WWW '17*, pages 735–744, Perth, Australia, 2017.