

The unprecedented increase in social media use brings many opportunities and threats at the same time. Social media help people to connect and share their opinions and experiences with millions of others. We can consider social media as a microscope for online world which magnifies individual and group behaviors. Using social media as a tool, researchers can study online protests, political debates, and changes in user behaviors. In addition social media provide a valuable medium for creating, sharing, and consuming information. Nowadays the internet provides instantaneous reach to information, but it also enables the creation of misinformation. Malicious intentions can be observed in the form of orchestrated campaigns and promotion of content with the help of social bots.

My dissertation work focused on the detection of persuasion campaigns and social bots using the tools of machine learning, data mining, and network analysis. I focused on three problems to study *online discourse and its manipulation*: (i) analysis of information diffusion and characterizations of user roles during online discourse; (ii) how to distinguish orchestrated social media campaigns from grassroots discussions; (iii) how to identify social bots and analyze interaction between different entities in the Twitter ecosystem.

My research on online discourse explores information diffusion and censorship under geographical constraints. I analyzed trending topics on Twitter and how they compete for popularity at the local level to emerge as winners at the country level. I identified two distinct mechanisms governing diffusion: those that surface locally, coinciding with three different geographic clusters; and those that emerge globally from several metropolitan areas, coinciding with major air traffic hubs [5]. In another project, I studied censorship to understand the impact of external influence on diffusion of censored content and behavior change of users. Results of this analysis are surprising because I found that censorship correlates with increased attention towards censored users. Users with censorship are more likely to attract more followers and create more impressions. I also show that when user timezone and languages are considered as location proxies, retweeting users find alternative ways to reach censored content and help its dissemination [11].

Along a similar vein to the online discourse research, I analyzed a social upheaval occurred in Turkey. Our analysis first characterizes the spatio-temporal nature of the conversation about Gezi park demonstrations, showing that similarity in trends of discussions mirrors geographic cues. More importantly, we describe the characteristics of the users involved in this conversation and what roles they played. We study how roles and individual influence evolved during the period of the upheaval. This analysis reveals that the conversation becomes more democratic as events unfold, with a redistribution of influence over time in the user population. We conclude by observing how the online and offline worlds are tightly intertwined, showing that exogenous events, such as political speeches or police actions, affect social media conversations and trigger changes in individual behavior such as changing screen name as a way to reveal their opinions [14, 8].

Information spreading on social media substantially contributes shaping collective opinions. Most information campaigns can be benign, but some may be designed for terrorist propaganda, political astroturf, and manipulation of financial markets and individual emotional states. The detection of viral memes that are sustained by coordinated campaigns has important social implications and poses numerous technical challenges. So does the identification of campaigns that might reach critical popularity in the future. Our team built a system that analyzes social media data and extracts network, temporal, content, and user-based features to detect online campaigns. I worked on several modules of this framework: (i) a clustering procedure that uses metadata to compute similarity between memes [3, 7]; (ii) a classification system that determines whether a meme is potentially an orchestrated campaign or a genuine, grassroots conversation [6, 13]; (iii) a social bot detection framework described more in detail below.

Increasing evidence suggests that a growing amount of social media content is generated by autonomous entities known as social bots. While not all bots are harmful, there is a growing record of malicious applications of social bots [4]. To detect social bots we built a machine learning system that extracts more than a thousand features in six different classes: users and friends meta-data, tweet content and sentiment, network patterns, and activity time series. We trained our models on publicly available datasets of social bots and evaluated them with a manually curated dataset of active users. The resulting system performs with high accuracy in detecting both real users and different classes of social bots. Using this classification framework, I analyzed over 14 million active users on Twitter to characterize communication between different entities and their ecosystem [12].

As a demonstration of my work on the detection of social bots, we built a service called BotOrNot¹ that evaluates the bot-likeness of Twitter accounts [12, 2]. Leveraging the lessons learned from this project, we participated in the DARPA bot detection challenge and we finished this competition as the second fastest and the third most accurate team [10].

Besides my thesis work, I was fortunate to be able to explore different topics on modeling heterogeneous intents of users on information diffusion [1] and analysis of social media timelines to study individual experiences and their outcomes [9]. This work is the result of my two summer internships at Microsoft Research in Redmond.

In addition to my main research topics, I am interested in psychology and cognitive science. Specifically, I am curious about dreams and the unconscious mind. My excitement about dreams led to a leisure-time research project in which I analyzed dream interpretations to investigate cultural differences and universal archetypes [15]. I am currently analyzing individual dream journals as unconscious early-warning signals of mood changes and precursor of behavioral transitions.

My ongoing research activities focus on modeling user interactions, leveraging online data across multiple platforms to understand conscious and unconscious behaviors. My prior work in social bots can aid in this endeavor. However, users and automated accounts use social media differently. Social bot accounts follow simple instructions to

¹BotOrNot: truthy.indiana.edu/botornot/

increase visibility, spread content, or influence others. Humans have more complex patterns of interactions, creation of content, and information consumption. Properties such as sentiment of conversations or interactions with friends follow unique but observable patterns. Users with similar motives share similar temporal signatures of their behaviors. But users may have multiple motives leading to different behaviors. The identification of distinct patterns of behavior is crucial to the study of the social system at the level of users.

Research Agenda

I am excited about the opportunities to mine social signals for gaining new insights about human behavior and society. The world we have been experiencing is changing and we have data with higher temporal resolution, more accurate as well as reflecting a complete picture of individual lives. The ethical collection of multi-modal data about individuals will be instrumental to understand human behaviors. I want to be one of the pioneers in this area by developing new models and tools to study complexity in terms of analyzing behaviors of individuals.

My long-term research goal is to develop models that describe intents and actions of individuals and groups. Detecting strategies employed by users is crucial for many reasons: preventing terrorist recruitment, identifying different classes of sophisticated social bots, and detecting orchestrated campaigns. Deviations from the regular patterns can also point to important events and pre-cursors of significant transitions. Understanding change in behavior helps to study mood changes and to identify significant life events. In the following, I describe several future directions I am excited to pursue.

Detecting strategies and orchestrated activities governed by malicious accounts. Intents and strategies of malicious entities such as social bots and orchestrated campaigns are either fully automated by software or directed by motivated human agents. Armies of social bots and misinformation campaigns are executed to promote ideas, advertise products, or sway public opinion. We have been observing social bots that attempt to persuade, influence, and deceive. My experience in the identification of social bots and early detection of campaigns helps to isolate those activities and study their strategies in-depth. I am interested in building detection systems that are evolving to lead in this arms-race by exploring behavioral signatures of users and characterizing their strategies.

Identifying individual intents and improving their well-being. One of the applications of ego-centric network research is to model mental health problems. In this domain, I would like to infer whether a user has issues like bipolar disorder and depression based on prior online interactions. To improve such inferences, I am studying the transfer of knowledge about users across platforms. My goal is to build models of interconnected data sources to highlight the relationships between user attributes and behavioral features. Once a particular group of people is selected on one platform, users with similar characteristics can be identified on other platforms. Additional features about the group can be extracted from these platforms to improve the inference model

and predict user behaviors. My goal of this research is to formulate new hypotheses about disease progression and developing mechanisms for support.

Studying dream reports to decipher the unconscious mind. I want to pursue my interest in dreams by building collaborations with clinical psychologists. Data driven research to understand the meaning of dreams and their implication on real life can be further improved by controlled experiments and data collection through mobile devices. Collaborative work in this area, in my opinion, will be greatly appreciated and rewarding to understand unconscious behaviors.

References

- [1] Abhimanyu Das, Sreenivas Gollapudi, Emre Kıcıman, and Onur Varol. Information dissemination in heterogeneous-intent networks. In *Proceedings of the 8th ACM Conference on Web Science*, pages 259–268. ACM, 2016.
- [2] Clayton Allen Davis[†], Onur Varol[†], Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. Botornot: A system to evaluate social bots. In *Proceedings of the 25th International Conference Companion on World Wide Web*, pages 273–274. International World Wide Web Conferences Steering Committee, 2016.
- [3] Emilio Ferrara, Mohsen JafariAsbagh, Onur Varol, Vahed Qazvinian, Filippo Menczer, and Alessandro Flammini. Clustering memes in social media. *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Network Analysis and Mining (ASONAM’13)*, 2013.
- [4] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. The rise of social bots. *Communications of the ACM*, 59(7):96–104, 2016.
- [5] Emilio Ferrara, Onur Varol, Filippo Menczer, and Alessandro Flammini. Traveling trends: social butterflies or frequent fliers? In *Proceedings of the first ACM conference on Online Social Networks (COSN)*, pages 213–222. ACM, 2013.
- [6] Emilio Ferrara, Onur Varol, Filippo Menczer, and Alessandro Flammini. Detection of promoted social media campaigns. In *Tenth International AAAI Conference on Web and Social Media*, 2016.
- [7] Mohsen JafariAsbagh, Emilio Ferrara, Onur Varol, Filippo Menczer, and Alessandro Flammini. Clustering memes in social media streams. *Social Network Analysis and Mining*, 4(1):1–13, 2014.
- [8] Christine Ogan and Onur Varol. What is gained and what is left to be done when content analysis is added to network analysis in the study of a social movement: Twitter use during gezi park. *Information, Communication & Society*, pages 1–19, 2016.
- [9] Alexandra Olteanu, Onur Varol, and Emre Kıcıman. Towards an open-domain framework for distilling the outcomes of personal experiences from social media timelines. In *Tenth International AAAI Conference on Web and Social Media*, 2016.
- [10] V.S. Subrahmanian, Amos Azaria, Skylar Durst, Vadim Kagan, Aram Galstyan, Kristina Lerman, Linhong Zhu, Emilio Ferrara, Alessandro Flammini, Filippo Menczer, Rand Waltzman, Andrew Stevens, Alexander Dekhtyar, Shuyang Gao, Tad Hogg, Farshad Kooti, Yan Liu, Onur Varol, Prashant Shiralkar, Vinod Vydiswaran, Qiaozhu Mei, and Tim Huang. The darpa twitter bot challenge. *Computer*, 49(6):38–46, 2016.
- [11] Onur Varol. Spatiotemporal analysis of censored content on twitter. In *Proceedings of the 8th ACM Conference on Web Science*, pages 372–373. ACM, 2016.
- [12] Onur Varol, Emilio Ferrara, Clayton Davis, Alessandro Flammini, and Filippo Menczer. How many twitter bots? large-scale evaluation of social bot detection. In *Under review*, 2017.
- [13] Onur Varol, Emilio Ferrara, Filippo Menczer, and Alessandro Flammini. Prediction and detection of social media campaigns. *EPJ-DS (Under review)*, 2017.
- [14] Onur Varol, Emilio Ferrara, Christine L Ogan, Filippo Menczer, and Alessandro Flammini. Evolution of online user behavior during a social upheaval. In *Proceedings of the 2014 ACM conference on Web science*, pages 81–90. ACM, 2014.
- [15] Onur Varol and Filippo Menczer. Connecting dream networks across cultures. In *Proceedings of the companion publication of the 23rd international conference on World Wide Web companion*, pages 1267–1272. International World Wide Web Conferences Steering Committee, 2014.