Socially-enhanced discovery processes

Gabriele Di Bona[1], Iacopo Iacopini[2], Enrico Ubaldi[3], Vito Latora[1,4,5], and Vittorio Loreto[3,5,6]
[1]School of Mathematical Science, Queen Mary University of London, London E1 4NS, United Kingdom
[2]Department of Network and Data Science, Central European University, 1100 Vienna, Austria
[3]SONY Computer Science Laboratories, Paris, 6, rue Amyot, 75005, Paris, France
[4]Dipartimento di Fisica ed Astronomia, Università di Catania and INFN, I-95123 Catania, Italy
[5]Complexity Science Hub, Josefst äadter Strasse 39, A 1080 Vienna, Austria
[6]Sapienza Univ. of Rome, Physics Dept., Piazzale Aldo Moro 2, 00185 Rome, Italy

Discoveries are essential milestones for the progress of our society. Therefore, unveiling the hidden mechanisms behind the emergence of new ideas is not only interesting from a scientific point of view, but also has a tangible sociological and economical impact. Recently, different mathematical approaches have been proposed to investigate and model the dynamics leading to the emergence of the new. Among these, of particular interest are random processes with reinforcement, such as urn models and biased random walks [1, 2]. These models successfully replicate the basic signatures of real-world discovery and innovation processes. However, they neglect the effects of social interactions. In particular, by considering the exploration dynamics as the one of a single entity, these models (i) do not capture the heterogeneity of the pace of the individual explorers; (ii) do not include the benefits brought by collaborations and, more broadly, social interactions. Indeed, empirical evidences of these mechanisms have been found in various contexts, from music-listening and language to politics and voting.

In this talk I give insights on how our peers can influence our experience of the new, using theoretical and data-driven mathematical models. In the first part of the talk, I introduce the model proposed in [3], where each explorer is associated with an urn model with triggering [1] (UMT) that governs its discovery dynamics. Urns are coupled through the links of a complex network, so that explorers can exploit opportunities (possible discoveries) coming from their social contacts, in a cooperative manner. We study the impact of the network topology on the exploration dynamics and we find that the pace of discovery β_i of an explorer i strongly depends on its position in the social network, as shown in Fig. 1(A) for the Zachary Karate Club network. Notice the higher pace of discovery displayed by the notoriously central nodes, suggesting that a strategic location on the social network correlates with the discovery potential of an individual. In particular, we show that the ranking of the nodes that distinguishes the fastest explorers can be predicted analytically by using the eigenvector centrality. This highlights that the structural—not just local—properties of the network can strongly affect the agents' ability to discover novelties.

In the second part of the talk I investigate a data set containing the whole listening histories of a large, socially connected sample of users from the online music platform *Last.fm*. We demonstrate in [4] that users exhibit highly heterogeneous discovery rates of new artists and that their social neighborhood significantly influences their behavior. In particular, we find that more explorative users tend to interact with peers more prone to explore new content. We capture this phenomenology in a data-driven modeling scheme where users are represented by random walkers exploring a graph of artists, shown in Fig. 1(B), and interacting with each other through their social links. Differently from the previous model, they interact with their peers by checking the current artist listened by one of their friends. Interestingly, even starting from a uniform population of agents, our model predicts the emergence of strong heterogeneous exploration patterns, with users clustered according to their musical tastes and propensity to explore.

We hope our work can represent a significant step forward to develop a general framework to understand how social interactions shape discovery and innovation processes.

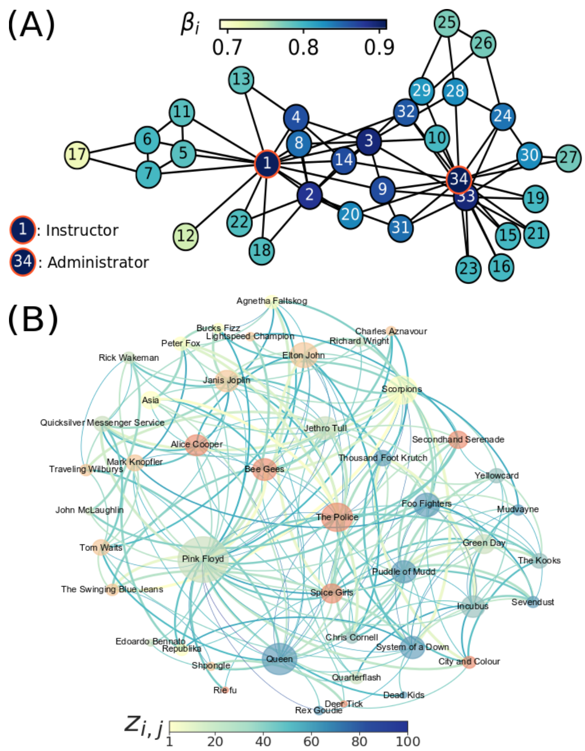

Fig. 1. **(A)** Zachary Karate Club Network, whose nodes are colored according to the discovery rates resulting from the model proposed in [3]. The discovery rate is an approximation of the Heaps' exponent, appearing in the famous empirical Heaps' law, a sub-linear growth of the number of distinct elements $D(t) \sim t^{\beta}$ with the number of elements t, which governs the rate at which novelties grow. Notice the higher pace of discovery displayed by the notoriously central nodes corresponding to the instructor (node 1) and the administrator of the club (node 34). **(B)** Snowball-sampled snapshot of the neighborhood of *Pink Floyd* in the space of artists in the *Last.fm* dataset [4]. Node sizes are proportional to their degree, while their color depends on the community they belong to. The color of the edges denotes the significance of the proximity of the two artists, the bluer, the larger.

[1] F. Tria et al. In: *Sci. Rep.* 4 (2014), p. 5890.
[2] I. Iacopini et al. In: *Phys. Rev. Lett.* 120 (2018), p. 048301.
[3] I. Iacopini et al. In: *Phys. Rev. Lett.* 125.24 (2020), p. 248301.
[4] G. Di Bona et al. In: *arXiv:2202.05099* (2022).