Understanding user behaviour in decentralized sharing networks

A.Gavaldà-Miralles[1,2], D.Choffness[1], J.Otto[1], M.Sanchez[1], F.Bustamante[1], L.Amaral[1], R.Guimerà[2], and J. Duch[2]

(1) Northwestern University, Evanston (US).
(2) Universitat Rovira Virgili, Tarragona (Spain).

The emergence of the Internet as the primary medium for information exchange has led to the development of many decentralized sharing systems. The most popular among them, BitTorrent, is used by tens of millions of people monthly and is responsible for more than one-third of the total Internet traffic. Despite its growing social, economic, and technological importance, there is little understanding of how users behave in this network. Due to the decentralized structure of peer-to-peer (P2P) networks, it is very difficult to gather large-scale data about interactions and behavioral patterns of the users without their explicit consent; this is in contrast to other forms of online exchange where all of the information is stored in a central system, such as Wikipedia, Twitter or Facebook.
In our works [1,2] we collected anonymized user activity data during 5 years from more than 1.4 million users of the Bittorrent P2P network. Using this dataset, we analized user activity patterns and the socioeconomic factors that could explain the behavior. In [1] we revealed that (i) the ecosystem is heterogeneous at several levels: content types are heterogeneous, users specialize in a few content types, and countries are heterogeneous in user profiles; and (ii) there is a strong correlation between socioeconomic indicators of a country and users behavior. In [2] we presented a bayesian model to describe the temporal patterns of activity of the users, and showed the robustness of these behaviors when there are exogenous events that perturbate the system.
These works demonstrate that despite the decentralized nature and privacy safeguarding intrinsic to P2P networks, they provide researchers with an extraordinary opportunity for investigating social and economic transactions on a large scale and to a level of detail not typically found for such large systems. The results presented have implications in the user's privacy, in the design of peer-to-peer algorithms and for policymaking.

[1] A. Gavaldà-Miralles et al. PNAS,11, (43) 15322-15327. (2014)

[2] A. Gavaldà-Miralles et al. Proc. of 10th CoNEXT, 319-324. (2014)