ORIGINAL RESEARCH

Detection of seed users vis-à-vis social synchrony in online social networks using graph analysis

Shabana Nargis Rasool[1] · Sarika Jain[1] · Ayaz Hassan Moon[2]

Abstract The dominance and prevalence of social media in the present world are significant because the role supplied by social networks is gradually growing with the passage of time. These social networks are often complicated networks in which each user is designated by a node and interactions between two users are symbolized by edges. People often express their opinions on any event via social media platforms. The interaction between users at a specific event, such as COVID-19, may constitute social synchrony, defined as a large population of users performing a specific action in unison. Identifying the seed users (influential users) from that event can be vital for a range of applications. Therefore, the current study proposes a framework to identify the seed users that works on the principles of graph analysis, viz. clustering, transitivity, and network centrality. Extensive experimentation is carried out using a self-collected dataset of COVID-19 tweets. Our dataset shows encouraging results in finding seed users in complicated networks.

Keywords Social Synchrony · Graph Analysis · Online Social Networks · COVID-19 · Real Dataset

✉ Shabana Nargis Rasool
   shabs.nargis@gmail.com

   Sarika Jain
   sjain@amity.edu

   Ayaz Hassan Moon
   moonah@rediffmail.com

[1] Amity Institute of Information Technology, Amity University, Noida, Uttar Pradesh 201313, India

[2] Islamic University of Science and Technology, Awantipora, Jammu & Kashmir 192122, India

1 Introduction

Online social media platforms facilitate users' ability to share and create content by offering greater permanence, frequency, and reach. Figure 1 illustrates that such online social media networks are widely prevalent among users [1]. Users numbering in the millions share content relatively faster through popular Online Social Networks (OSNs) such as Twitter, Facebook, WhatsApp, etc. The most popular social networks are typically available in various languages, allowing users to communicate with friends or acquaintances across economic, geographic, and political boundaries. Social networking sites had 3.96 billion users in 2022, and these statistics are anticipated to increase as smartphones and mobile social networking sites gain popularity in earlier underrepresented markets. Such profuse information has attracted researchers from varied domains to investigate and attain significant insights from the OSN data, with the presence of heterogeneity, noise, and unstructured data as open challenges. Detection of seed users in OSNs has a wide range of applications, such as assisting in creating successful product advertisements, supporting and disseminating information, removing a certain critical node, and preventing the spread of epidemics and rumors. [2–5].

The COVID-19 pandemic has sparked widespread concern in the community. Each nation made significant efforts to keep the virus's spread at bay. It would be simpler to prevent the dissemination of pandemics if we could analyze social networks. Propagation dynamics has been a significant research focus in recent generations. Many pathways, including epidemic spread, rumor dissemination, social upheaval spreading, and e-commercial ad campaigns, are all intimately associated with complicated network dynamics, and all of these pathways rely on the identification of seed users in some way or another [6–8]. Twitter is among

13