An Evolutionary Study on The Popularity of Packages in NPM JavaScript Package Ecosystem

Shi Qiu, Raula Gaikovina Kula and Katsuro Inoue
Osaka University, Japan
{qiujitsu, raula-k, inoue}@ist.osaka-u.ac.jp

User contributed open source library ecosystems like the NPM JavaScript libraries for the node.js platform has had a tremendous impact on society and software development, capturing the attention of businesses, researchers, and policy makers. Millions of participants, from independent volunteers to those representing companies or government organizations, have created and maintain massive numbers of popular JavaScript libraries suchlike JQuery, Bootstrap, React, and Angular, and components from frameworks including Polymer. Yet, despite the substantial amount of research on these self-organizing ecosystems, it remains unclear how and why these ecosystems form, how they achieve their impact, or how they sustain themselves. Related, studies investigate how do the different factors come together to form a community that develop and maintain valuable and freely available software, and how does an ecosystem with millions of repositories and developers operate given the lack of centralized planning. Additionally, how does a library emerge to become prominent and gain popularity and sustainability and do the libraries eventually decay and get replaced by other libraries after they have reached their end-of-life. These are just the many research questions that researchers and stakeholders ponder when using software from these ecosystems.

In this exploratory study, our main goal is towards the emergence of popular libraries, which we believe is a key factor in the sustainability of an ecosystem. Concretely, we empirically conduct an evolutionary study of popular libraries, with an attempt to find the relationship between popularity and evolution of a library over time. Firstly, we mine and extract 151,276 library packages from the NPM library ecosystem, understanding the relationship between the popularity over time, which is measured by the download times. Preliminary results show that the popularity of packages in NPM range on a scale of being extremely popular, to being only used by a small group of systems. Then, to better understand the growth of popular libraries, we model popularity over time to find the most common popularity growth pattern. First results show that the most common patterns are linear growth, then followed by the

logarithmic growth and quadratic growth patterns. These early results are promising and indicate that popular libraries are only for an elite subset of libraries, and once they gain first popularity, then the growth is quadratic. We envision that achieving understanding of how popular libraries emerge in the software ecosystem would inform approaches to structuring the success of future open source communities, and lead to greater sustainability of the ecosystem.