Abstract

Using data from source network at Sourceforge.net, the largest repository of Open Source Software (OSS) projects and contributors on the Internet, we construct two related networks: A Project network and a Contributor network. We define a link between two projects if the projects have at least one contributor in common. Similarly two contributors are linked if they work together on at least one project. Interestingly, both the project network and the contributor network consist of one "giant" connected component and many smaller unconnected components. Knowledge spillovers may be closely related to the structure of such networks, since contributors who work several projects likely exchange information and knowledge. We thus examine the effect of the structure of these network on the success of OSS projects where success is defined as the number of downloads. Our main results are: (i) additional contributors are associated with an increase in output, but that additional contributors in projects in the giant component are associated with greater output gains than additional contributors in projects outside of the giant component; (ii) *Betweenness* centrality is highly associated with the number of downloads and this association is stronger than the relationship between other measures of centrality (*closeness* and *degree*) and the number of downloads. This result suggests that there are positive spillovers of knowledge for projects occupying critical junctures in the information flow. When we define projects as connected if and only if they had at least two contributors in common, we again find that additional contributors are associated with an increase in output, and again find that this increase is much higher for projects with strong ties than other projects in the giant component.