Application of evolutionary computation techniques for the identification of innovators in open innovation communities

M.R. Martínez-Torres *

Facultad de Turismo y Finanzas, University of Seville, Avda. San Francisco Javier s/n, 41018 Sevilla, Spain

ARTICLE INFO

ABSTRACT

Open innovation represents an emergent paradigm by which organizations make use of internal and external resources to drive their innovation processes. The growth of information and communication technologies has facilitated a direct contact with customers and users, which can be organized as open innovation communities through Internet. The main drawback of this scheme is the huge amount of information generated by users, which can negatively affect the correct identification of potentially applicable ideas. This paper proposes the use of evolutionary computation techniques for the identification of innovators, that is, those users with the ability of generating attractive and applicable ideas for the organization. For this purpose, several characteristics related to the participation activity of users though open innovation communities have been collected and combined in the form of discriminant functions to maximize their correct classification. The right classification of innovators can be used to improve the ideas evaluation process carried out by the organization innovation team. Besides, obtained results can also be used to test lead user theory and to measure to what extent lead users are aligned with the organization strategic innovation policies.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

The concept of open innovation, launched by Chesbrough (2003) and others, has become increasingly popular among scholars and industry practitioners since the term was coined. Open innovation refers to the use of external sources and actors to achieve innovation, and it is based on the idea that companies should not just rely on internally developed ideas and knowledge, but increasingly also on ideas and knowledge developed externally (Chesbrough, Vanhaverbeke, & West, 2006; Tödtling, Prud'homme van Reine, & Dörhöfer, 2011). It assumes that useful knowledge is widely diffused and abundant. There is, for example, a growing availability of knowledge from multiple innovation actors, including universities, specialized suppliers, inventors and knowledge brokers. In this conditions, the "old" model of closed innovation where innovation processes are controlled by the company needs to be changed in favor of the detection and assimilation of externally developed knowledge (Barge-Gil, 2010; De Jong, Kalvet, & Vanhaverbeke, 2010; Poetz & Schreier, 2012). Previous studies agree that open innovation is not a general phenomena and depends on certain company characteristics as well as external conditions. Chesbrough (2003) identified various external factors that explain why enterprises increasingly adopt the open paradigm. The

availability of a strong public knowledge base, a mobile and educated working population or the availability of ample external finance for innovation are the three conditions enabling open innovation to emerge.

From the viewpoint of the organization, there are various mechanisms and channels used for sourcing and acquiring external knowledge such as the absorption of local knowledge spillovers, collaboration in R&D and innovation with firms and universities, relations to spin-off companies, informal knowledge interactions, customer contributions through design toolkits or idea competitions (Keeble & Wilkinson, 2000; Schwab, Koch, Flachskampf, & Isenhardt, 2011; Tödtling, Lehner, & Trippl, 2006; von Hippel & Katz, 2002). The strategic challenge is how firms can best organize the sourcing, codification and exploitation of the internal and external knowledge and informational resources to maximize and sustain innovation (Love & Roper, 2009). One of the most popular mechanism for open innovation implementation is user innovation communities (Dahlander, Frederiksen, & Rullani, 2008). Firms such as Microsoft, Dell, IBM, BMW, and Nokia increasingly invest in virtual communities to solicit user contributions as part of their innovation processes. This trend is explained by the increase in digitalization and the decrease in the costs of communication that have lead to an exponential growth of user innovation platforms (Mahr & Lievens, 2012). Internet have facilitated the accessibility of these platforms by users geographically distributed all over the world. However, this accessibility is also causing the

* Tel.: +34 954 55 43 10; fax: +34 954 55 69 89.
  E-mail address: rmtorres@us.es

generation of a huge amount of information which it is difficult to process and evaluate by the innovation departments or experts within organizations. Posted ideas must be evaluated one by one by the innovation department or even some specific experts of the organization, and this evaluation consists of reading the idea, assessing its applicability attending to the strategic innovation policies of the organization and planning their possible implementation in case they are finally accepted. The problem is that online user innovation communities can generate hundreds or even thousands of solutions in a short period of time, saturating the capacity of internal evaluators and hiding the really attractive innovations. That is the reason why online user innovation communities typically include some type of scoring systems so that the community can evaluate potential solutions. This scoring scheme is based on the idea that the crowd can do a better evaluation than individuals since they own a group-based intelligence which can outperform individual knowledge (Surowiecki, 2004).

However, strategic innovation policies of organizations are not always aligned with users' desires. Some non affordable ideas can be excellent for users but prohibitive for the company, and these ideas would probably receive a high score by other community users. In this sense, it is much more useful for the company the identification of users posting ideas that will be finally adopted. This information can be easily collected from innovation communities websites as they usually inform users about the status of their posted ideas. The purpose of this paper consists of the identification of these innovators, defined as those users generating ideas that will be finally adopted by the company. The condition of being innovator or non innovator is a dichotomic property of each user. Therefore, the identification of innovators is a classification problem that can be solved using a discriminant function over a set of variables characterizing the activity and behavior of users within the community, which on the other hand is the main available information. The main problem associated to this identification is that the considered dependent variable contains a high number of zeros (non-innovators), leading to the so called zero inflated problem. To solve this issue, a optimization procedure consisting of finding the values of the variables coefficients so that the discriminant function can maximize the percentage of correct classification of innovators and non innovators is formulated. Three different evolutionary computation techniques are used to solve the problem for evaluating the reliability of results. Additionally, a bootstrapping technique has also been implemented to obtain the confidence intervals of the resulting coefficients.

The rest of the paper is structured as follows. Section 2 details previous works related to the open innovation paradigm and the identification of users with special profiles. Section 3 describes the formulation of the problem in the form of an optimization problem and presents the three proposed evolutionary computation techniques: simulated annealing, particle swarm optimization and genetic algorithms. The three algorithms are then applied to the case study of IdeaStorm website, which is introduced in Section 4, as well as those variables measuring the activity and behavior of users within this innovation community. Obtained results are discussed in Section 5. Finally, conclusions are provided in Section 6.

2. Related work

Online user innovation communities make use of Internet as the prime communication channel, allowing company-to-customer as well as customer-to-customer communications (Di Gangi & Wasko, 2009; Rohrbeck, Steinhoff, & Perder, 2008). This strategy assumes that new product developments require interactions among like-minded customers who talk about their usage experiences, raise questions, present solutions, and offer answers (Fueller & Matzler, 2007). These interactions enable users to build on one another's knowledge and experiences, which plays a critical role in developing ideas (Rowley, Kupiec-Teahan, & Leeming, 2007). Previous research on open innovation communities have been mainly focused on their operational level. The first studies in this line discussed the characters of user innovation community based on the example of open-source software projects, which was a relatively well-developed and very successful form of internet-based innovation community (Martínez-Torres, 2012; Von Hippel, 2001; West & O'mahony, 2008). Later, Von Hippel and Von Krogh (2003) proved that user innovation communities illustrate a "private-collective" model of innovation incentive. However, there are some differences between open source software communities and open innovation communities. The most important one is that open source communities works based on user requests for information or help, while open innovation communities are more impersonal and users share information with others but not responding to any specific request of information. The effectiveness of open innovation is another interesting issue treated in the previous literature. Laursen and Salter (2006) found a non linear relationship between open innovation and performance, concluding that too much open innovation hurts organization performance. More specifically, they found that innovative performance is curvilinearly related through an inverted U-shape to the number of sources. The size is explained because firms gain innovative opportunities as they implement a wider and deeper search over a huge number of sources. However, innovation search is not costless and can be time consuming, expensive, and laborious. In the case of online innovation communities, the most important cost is the one associated to evaluation of posted ideas. Although the marginal evaluation cost of each idea is low, the cumulative evaluation cost when thousands of ideas are posted can be tremendous. Community based voting methodologies like simple discussion forums, community ratings (Carbone, Contreras, Hernández, & Gomez-Perez, 2011; Frey, Lüthje, & Haag, 2011) or more complex methodologies like prediction markets (Blohm, Riedl, Leimeister, & Krcmar, 2011; Spann & Skiera, 2009) that are based on stock-market trading algorithms can help to solve this challenge. Any assessment system based on a community scoring model, in which ideas can be awarded with a specific number of points, can help selecting the best idea (Hüsig & Kohn, 2011). However, obtained results through these procedures may be contradictory with the innovation strategic policies of organizations. Some of the top ranked ideas may be in the opposite direction of organization priorities or their implementation costs can be prohibitive. An alternative to the scoring model is the identification of best ideas through the identification of a particular subset of users called lead users. Lead users are characterized because they anticipate early on innovative characteristics, which are relevant only much later for other customers (Von Hippel, 1986, 1988). Additionally, lead users have the ability to develop a fully functional solution for their needs (Mahr & Lievens, 2012; Morrison, Roberts, & Midgley, 2004). They hence possess not only need information, but equally also solution information. Previous research about lead user have been focused on issues like their identification (Urban & von Hippel, 1988). The behavior of lead users have been described in the literature by several characteristics. For instance, their ability to bear innovative solutions is fundamentally linked to a person's individual creativity (Amabile, Barsade, Mueller, & Staw, 2005). They also make regular contributions to the community as they are actively engaged in problem-solving. However, some authors criticizes that their behavior is biased by their interest in obtaining a benefit from their proposed solutions. For instance, Berthon, Pitt, McCarthy, and Kates (2007) consider the idea of creative users as opposed to lead users. Creative users do not necessarily face needs that will become general as lead