World Patent Information 51 (2017) 46–56

Contents lists available at ScienceDirect

World Patent Information

journal homepage: www.elsevier.com/locate/worpatin

A novel method for retrieving specialisation profiles – The case of patent agent firms

Kazimir Menzel [b], Lutz Maicher [a, b, c, *]

[a] Fraunhofer IMW, Neumarkt 9-19, Leipzig, Germany
[b] Technology Transfer Research Group, Friedrich-Schiller-Universität, Jena, Germany
[c] Michael Stifel Center for Data Driven and Simulation Science (MSCJ), Jena, Germany

ARTICLE INFO

Article history:
Received 18 November 2016
Received in revised form
18 September 2017
Accepted 14 November 2017

ABSTRACT

Introduction: For a broad variety of innovation intelligence tasks, like finding the patent agent being specialised in a technology, specialisation profiles of entities are of particular interest. Such a profile indicates for each entity in which activities it is markedly specialised. Conventional approaches suffer two shortcomings, though. First, they tend to consider only the activity profiles of the respective entities, while neglecting information about all other entities' activities. Second, they often lack in considering entity specific characteristics, when assigning an activity to an entity's specialisation profile or not.
Method: To address these shortcomings, we introduce the RSI-specialisation, a novel method for retrieving dichotomous and idiotypic specialisation profiles. Operationalising relative specialisation, the novel method is rooted in the theory of comparative advantage. In order to contrast it with the approaches based on absolute specialisation, it is compared with the baseline method ENF- specialisation, which has its theoretical roots in the concept of effective number of components.
Analysis: Both methods are demonstrated by applying them to the case of the specialisation of patent agent firms. RSI-specialisation and ENF-specialisation are applied to a data set containing all EPO patent applications in 2014 and 2015.
Results: Compared to the baseline approach, the RSI-specialisation reduces noise from market effects to a greater satisfaction. Besides being less dependent on agent size and market structure, it reduces interpretation to the most essential question for an applicant, i.e. why he should opt for one rather than another provider. In addition, it also guarantees that a potential applicant can find in any field a specialised patent agent firm.
Conclusion: We find that the novel RSI-specialisation promises to be a robust and reliable method for retrieving dichotomous and idiotypic specialisation profiles. Patent agent firms aside, it could be applied to a multitude of different domains, like the specialisation of other professional workers, of politicians, experts in consulting firms or even to users in online-shops and patients in clinical trials.

1. Introduction

As the importance of intellectual property rights (IPR), like patents, increases, the innovator faces now more than ever the task of not only transforming his invention into a viable product but also to secure and monetise his IPRs. As a result, the market for intellectual property services is on the rise [1]. This confronts both, IP service providers like patent agent firms and innovators with new challenges in market orientation and matchmaking [2]. Successful innovators need the patent agent firm with the most suitable profile. Today the matchmaking process is based mostly on personal recommendations instead of objective assessments. Improved knowledge about the idiotypic specialisation profile of a patent agent firm could thus help the innovators to make better informed decisions in contracting services. By *idiotypic* we mean that the specialisation profiles are derived from individually defined characteristics of the patent agent firm.

For the prospective patent applicant, it is important to ensure that the application process is successful and swift, to keep the cost of his application low, and delivers results of high quality, i.e.

* Corresponding author. Friedrich-Schiller-Universität, Ernst-Abbe-Platz 2, D-07743 Jena, Germany.
E-mail address: lutz.maicher@uni-jena.de (L. Maicher).

https://doi.org/10.1016/j.wpi.2017.11.005
0172-2190/

patents that hold in court. All these aspects depend considerably on the technological expertise of the patent agent firm entrusted with the application. It is therefore important for the applicant to correctly identify specialists in the technology field(s) of her invention.

On the side of the patent agent firms, the increasing number of market participants, the ongoing internationalisation of IPR, and especially the increase of technological complexity implies a broader competition across national markets and technology areas. One viable strategy in patent agent firms' business models is the road towards a *niche*. This implies a focus on few technology areas, like pharmaceuticals, to differentiate themselves from the competitors. However, as technological development and demand for services can shift swiftly from one field to another [3] or, even more important, touch different technology areas at the same time. Therefore, it might be a healthy strategy for patent agent firms to maintain a certain level of *diversification* across technology fields. As a result, these patent agents offer a wider array of expertise than a simple niche-oriented approach might suggest.

However, how can innovators solve the information retrieval problem of finding the patent agent firm with the best suited specialisation profile? While most patent service providers communicate their areas of specialisation in one way or another on their websites and in marketing material, researching these heterogeneous sources might turn out a time consuming and expensive endeavour for a prospective patent applicant.

Through the IPC classes assigned to each patent application, patent data indirectly discloses whether a patent agent firm has been active in a given technological field. Not only the search for an appropriate patent agent firm, but a multitude of decisions are informed by data sets with such a common matrix structure: a set of entities (like persons, companies, or animals), a set of activity types being realised by the entities (like involvements in initiatives, number of followers, interests in products, or patents in specific IPC classes), and for each combination the observed number of how often the entity has realised the specific activity [4]. One common type of retrieval problem against these data structures is about specialisation; hence, about which entities are concentrating their involvement in which activities.

Retrieving specialisation profiles are a mean to approach these questions [5]. We define a specialisation profile of an entity as the rank-ordered list of activities in which the entity is relatively more active than in its own average, as well markedly more active than the average entity in the respective activity. Simplified, a specialisation profile is the selection of all activities, in which an entity is considerably more active than could be expected from the average behaviour of all entities.

Obviously, a naïve approach of counting activities yields in low-information specialisation profiles. Previous work shows, that simple patent counts are very weak proxies of patent strength and competitive impact [6,7]. We will refer to this kind of approaches as absolute specialisation. These approaches are challenged by the fact, that an observed level for a specific activity could be considered as being low, if it is realised by a big entity in a popular activity type. At the same time, a similar level could be considered as being high, if it is realised by a small entity in a niche activity. Thus, given the same amount of involvements, in the first case, the activity shouldn't be part of the specialisation profile of the entity, in the second case, it should be.

In order to overcome these deficiencies, we introduce a novel method of relative specialisation measure, the *relative specialisation index* (RSI). This method retrieves *idiotypic* and *dichotomous* specialisation profiles. Idiotypic means that the profiles only contain activities that markedly diverge from the average activity, analysed from the perspective of the entity as well as from the perspective of the activity itself. Dichotomous means that for each activity is binarily decided whether it is part of an entity's specialisation profile or not.

With the RSI-specialisation we present a novel method for retrieving specialisation profiles, which can be used in a multitude of applications. First, it allows for better informed indexing functionalities. Given a specific activity, such an index allows to retrieve all entities that are markedly specialised in this activity. Furthermore, the index allows to analyse the popularity of specific activities, based on the number of entities specialised in them. Another interesting aspect are combinations of activities within specialisation profiles, in order to apply co-occurrence-based techniques for similarity detection or clustering.

In this paper, we demonstrate how the RSI-specialisation method can be used to retrieve *idiotypcial* and *dichotomous* specialisation profiles of the patent agent firms in the market. By various examples we will demonstrate, that these profiles inform the matchmaking process between the firms and the innovators in a better way than the baseline approach. The RSI-specialisation is already built into the IP Industry Base (IPIB, http://s.fhg.de/ipib), an online database about the market for IP service providers. Based on the weekly updated activity matrix of the full market (disclosed by EPO patents), the specialisation profiles are calculated for all relevant patent agent firms in the market and updated on a weekly base.

The remainder of the paper is structured as follows. In the next section, we introduce and formalise the concept of absolute and relative specialisation. We then introduce in section 3 the RSI-specialisation method, a generic approach of retrieving idiotypic and dichotomous specialisation profiles for entities. In the subsequent sections, we apply this measure to the EPO patent activity data for the years 2014 and 2015 and compare it to our baseline index, the ENF-specialisation. In the last section, we demonstrate the application of the RSI specialisation profiles, and how they are calculated on a weekly base for the full market.

2. Two concepts of specialisation

Specialisation in an economic sense is defined as prioritising one field of activity over one or many others. To our best knowledge, there exists no body of literature on the technological specialisation of patent agent firms. However, things are slightly different for similar analyses of the service specialisation of technology firms [8, 734–738], but the general literature on specialisation is abundant [9; 10, 33–38], even by using patent data as empirical basis [11–13].

For a given patent agent firm, the resulting distribution of activities over technology fields constitutes an activity profile. Such a profile informs in absolute terms how much an agent concentrates its activities on certain technology fields. However, activity levels depend on the size and the age of an agent, or they depend on the market demand for technology fields or the location of a patent agent. By consequence, information obtained by simple counting activities into an activity profile may be noisy. As these noise factors operate on individual and global level, it is necessary to filter them out.

In the literature, two general concepts of specialisation are employed. One compares absolute counts in given fields or for given entities. One example is the definition of “distinctive scientific specialisation” by [8, 741-2]. The second approach is to take into account the size of an entity as well the market structure and filter these effects out. This latter concept is called *relative specialisation*, while we will refer to the former concept as *absolute specialisation*. Below the differences between both notions will be discussed in detail.