The evolution of intellectual property strategy in innovation ecosystems
Uncovering complementary and substitute appropriability regimes
Holgersson, Marcus; Granstrand, Ove; Bogers, Marcel

Published in:
Long Range Planning

DOI:
10.1016/j.lrp.2017.08.007

Publication date:
2018

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC-ND

Citation for published version (APA):
The evolution of intellectual property strategy in innovation ecosystems: Uncovering complementary and substitute appropriability regimes

Marcus Holgersson a, *, Ove Granstrand b, Marcel Bogers c

a Chalmers University of Technology, Department of Technology Management and Economics, SE-412 96 Gothenburg, Sweden
b Chalmers University of Technology, Department of Mathematical Sciences, SE-412 96 Gothenburg, Sweden
c University of Copenhagen, Department of Food and Resource Economics, Unit for Innovation, Entrepreneurship and Management, Rolighedsvej 25, 1958 Frederiksberg C, Denmark

Abstract

In this article, we attempt to extend and nuance the debate on intellectual property (IP) strategy, appropriation, and open innovation in dynamic and systemic innovation contexts. We present the case of four generations of mobile telecommunications systems (covering the period 1980–2015), and describe and analyze the co-evolution of strategic IP management and innovation ecosystems. Throughout this development, technologies and technological relationships were governed with different and shifting degrees of formality. Simultaneously, firms differentiated technology accessibility across actors and technologies to benefit from openness and appropriation of innovation. Our analysis shows that the discussion of competitiveness and appropriability needs to be expanded from the focal appropriability regime and complementary assets to the larger context of the innovation ecosystem and its cooperative and competitive actor relations, with dispersed complementary and substitute assets and technologies. Consequently, the shaping of complementary and substitute appropriability regimes is central when strategizing in dynamic and systemic innovation contexts. This holds important implications for the management of open innovation, innovation ecosystems, platforms, and coopetition.

Introduction

An open approach to innovation has in both practice and research been increasingly accepted as a potential source of competitive advantage, enabling the use of external sources of innovation and external commercialization strategies (Granstrand, 1982; Granstrand and Sjölander, 1990; Chesbrough, 2003; Chesbrough et al., 2006; Dahlander and Gann, 2010; West and Bogers, 2014; Cassiman and Valentini, 2016). Every open innovation activity involves two or more actors, and a firm that engages in open innovation is part of a system of interconnected innovation actors, resources, activities, and institutions, connected by organizational and market relations. We here denote such a system an innovation ecosystem. With an increased awareness of the systemic nature of open innovation there have been several calls for research that brings increased understanding of the systems level of open innovation (Chesbrough and Bogers, 2014; West et al., 2014; Bogers et al., 2017).
One of several questions that deserves renewed attention when innovation takes place across systems of actors rather than in an integrated firm is the question of how firms appropriate or profit from innovation (Teece, 2006, Teece, forthcoming). Teece (1986) identified the role of appropriability regimes and complementary assets for explaining the distribution of profits from innovation in the 1980s. An important contribution was that when the appropriability regime is weak, i.e., when imitation is easy, it is important for innovators to establish positions in complementary assets in order to appropriate value from innovation. However, complementary technologies (as a subset of complementary assets) did not receive much explicit attention and the role of intellectual property (IP) strategy for the appropriability regime was not well developed at that time, as later described by Teece (2006). Today we know that IP strategy plays an important role for firms’ competitiveness (Granstrand, 1999; Pisano, 2006; Teece, 2006; Pisano and Teece, 2007; Somaya, 2012), and that innovation often takes place in complex multi-technology systems with complementary innovations spread across actors (Granstrand et al., 1997; Hall and Ziedonis, 2001; Bessen, 2004; Teece, 2009; Somaya et al., 2011; Granstrand and Holgersson, 2013; Teece, forthcoming).

Complex innovation settings require firms’ IP strategies to include protection and/or sharing of their own technologies on the one hand and access to others’ technologies on the other hand (Alexy et al., 2009; Somaya et al., 2011; Granstrand and Holgersson, 2013). In such settings IP strategy impacts appropriability both directly, through improved and protected sales and margins, and indirectly, for example through cross-licensing agreements, improved negotiation positions, standard-setting, blocking of others’ R&D, and improved provision of complementary innovations (Arundel et al., 1995; Duguet and Kabla, 1998; Granstrand, 1999; Bekkers et al., 2002a; Baldwin and von Hippel, 2011; Holgersson and Wallin, 2017). Moreover, a specific firm’s freedom to operate, i.e., the ability to do business without being excluded by the IP rights (IPRs) of others, is impacted by the firm’s own IP strategy as well as other firms’ IP strategies (Granstrand, 1999; Lemley and Shapiro, 2007; Bessen and Maskin, 2009; Somaya et al., 2011; Holgersson and Wallin, 2017; Jell et al., forthcoming).

In this paper, we want to shed light on the complexities that emerge when innovation takes place across actors in innovation ecosystems, implying an evolution of innovation and IP strategy across actors over time. We will present the case of mobile telecommunication systems in the period 1980–2015 to describe and analyze the co-evolution of strategic IP management and innovation ecosystems. Hereby, we extend and nuance the debate on IP strategy, appropriation, and open innovation in dynamic and systemic innovation contexts. The generational shifts in mobile telecommunication systems highlight the systemic and dynamic nature of such an open innovation context in which the role of IP strategy is not so much to provide short-term profits from single innovations as it is to ensure long-term competitive advantage within the innovation ecosystem—an advantage that is dependent on both internal and external innovations.

Our analysis leads to several contributions. First we identify the important roles of both collaborating and competing actors and both complementary and substitute technologies in innovation ecosystems. Second, we explicate the concept of appropriability, and relate it to the two distinct dimensions of technology governance and technology accessibility. Rather than focusing on the ease of imitation, this perspective acknowledges the fact that firms can use formal and/or informal technology governance to obtain high and/or low levels of technology accessibility in order to benefit from various forms of open and/or closed innovation in compliance with their business models (Chesbrough and Rosenbloom, 2002; Teece, 2010). Third, we illustrate how firms need to extend their focus from the focal appropriability regime to its complementary and substitute appropriability regimes, especially in systems technologies that are subject to standardization.

Theoretical background

The case of mobile telecommunications has been studied by one of the authors since the 1980s, and the use and development of theory have consequently evolved alongside the empirical studies over time. Transaction cost theory and its associated organization theory of the firm as developed by Williamson (1975) provided the main framework at the start of the longitudinal study. The growing use of various quasi-integrated organizational forms for conducting R&D and innovative activities, such as licensing and inter-firm collaborations, was identified early on and a typology of external technology acquisition and exploitation strategies was developed (Granstrand, 1982; Granstrand and Sjölander, 1990; Granstrand et al., 1992) based on contract theory (Grossman and Hart, 1986). Studies then showed that technology diversification and concomitant external technology acquisition and exploitation play pivotal roles for combining and recombining complementary technology resources into complementary multi-technology products, thereby generating recombinant growth (Granstrand and Oskarsson, 1994; Granstrand, 1998; Cantwell et al., 2004). Today such interorganizational innovation is often denoted “open innovation” —going back to Chesbrough (2003)—and it has received huge attention from both research and practice.

Open innovation research explains how firms can rely on external technologies to augment their internal innovation development or how they can tap into external partners to exploit internally developed technologies (Chesbrough and Bogers, 2014). This stream of research has paid much attention to how to leverage external sources of innovation with some key considerations being different mechanisms to obtain such innovation and how to integrate it into the organization (West and Bogers, 2014). At the same time, there are considerations about possible decreasing returns in terms of external search and appropriability (Laursen and Salter, 2006, 2014) and not only value-enhancing but also cost-increasing effects of open innovation (Faems et al., 2010). More generally, open innovation scholars have identified various mechanisms in relation to inbound knowledge flows and to some extent outbound knowledge flows, although these have largely been considered on the organization level with less attention to higher level attributes such as innovation systems (Chesbrough and Bogers, 2014; West et al., 2014). Recently, Bogers et al. (2017) proposed the innovation ecosystem as an important unit of analysis for future
The concept of open innovation is closely linked to Teece’s (1986) profiting from innovation (PFI) framework, which explains not only the role of complementary assets and appropriability regimes for the profitability of innovation, but also the decision between contracting and integration, i.e. the choice between open and closed innovation in some sense. According to Teece (1986), the ability of innovators to profit from innovation is impacted by both the appropriability regime at hand (which is related to the nature of the technology and to the legal means by which it can be protected) and by the access to complementary assets. The framework differentiates between three types of complementary assets—generic (general purpose), specialized (having a unilateral dependence between the asset and a focal innovation) and co-specialized (having a bilateral dependence)—be they tangible or intangible assets. Teece’s conclusion is that when imitation is easy, i.e., when the appropriability regime is weak, innovators need complementary assets, such as manufacturing assets, distribution assets, service assets, and complementary technologies, in order to appropriate value from innovation. Another conclusion is that “contracting [e.g., licensing] rather than integrating is likely to be the optimal strategy when the innovators appropriability regime is tight and the complementary assets are available in competitive supply (i.e. there is adequate capacity and a choice of sources)” (Teece, 1986, p. 293). Thus, the framework already from the outset addressed the relationship between appropriability and what would eventually be termed open innovation. However, the framework initially placed limited explicit focus on IP strategy and complementary technologies. These have received increasing attention more recently (Teece, 2006; Pisano and Teece, 2007; Teece, forthcoming).

Extant IP management literature increasingly accounts for the fact that firms adopt various types of open innovation strategies and thereby need to match their IP strategies with this (e.g., Alexy et al., 2009; Somaya, 2012; Granstrand and Holgersson, 2014; Manzini and Lazzarotti, 2016). This stream of literature has shown that, despite the fact that open innovation is related to knowledge sharing across firm boundaries, patents have an important role to play (Bogers, 2011; Zobel et al., 2016; Zobel et al., 2017) and are actually perceived as more important in open innovation settings than in closed innovation settings (Hagedoorn and Zobel, 2015; Holgersson and Granstrand, 2017). Open innovation, and the related knowledge diffusion, increases the risks of imitation (Veer et al., 2016), the risks of being blocked by others (Blind et al., 2006), and the need for safeguarding bargaining power over time, and consequently increases the value of protecting technologies with patents (Holgersson and Granstrand, 2017).

This literature typically takes the perspective of a focal firm, and there is a need to better understand the systemic nature and the long term strategy interactions and dynamics across actors (Teece, forthcoming). A firm that employs a very protective patent strategy may for example benefit from high margins in the short run, but suffer from a lack of complementary innovations or retaliatory patenting of complementors in the longer run, and a firm that employs an open source strategy may benefit from wide use of its technology, quick adaptations, and many contributors in the short run, while losing its competitiveness as an exclusive provider of the technology in the longer run. Thus, relationships in open innovation, technology markets, and innovation ecosystems are seldom symmetrical, and asymmetries may shift over time, for example due to technological changes and shifts between product and technology generations (Granstrand, 1999), due to strategic and business model changes (Baden-Fuller and Haefliger, 2013; Hacklin et al., forthcoming), or due to entries and exits of actors in the innovation ecosystem.

In the remainder of this paper we will focus on this complex co-evolution of strategic IP management and innovation ecosystems. We will do so by describing and analyzing the case of mobile telecommunications, an industry where patenting, licensing, and litigating have increased rapidly since its inception.

**Methodological approach**

**Case selection and virtues of multi-generation studies**

For the purposes of this article, a longitudinal case enables a rich picture of the studied phenomenon covering developments over long periods of time. As noted by Hargaden and Douglas (2001, p. 480) “historical case studies […] provide a perspective that covers the decades often necessary to observe an innovation’s emergence and stabilization”. It is moreover necessary to take a longitudinal perspective when studying strategic interplay that evolves over time. The case of mobile telecommunications was chosen due to several reasons. First, the development has from the outset taken place in an innovation ecosystem, which is the subject of study as a phenomenon that needs further research to be fully understood (von Krogh et al., 2012; Adner and Kapoor, 2016). Second, two of the authors are part of a research group having studied the development of this industry since the 1980s within several sub-studies, enabling ongoing access to primary data. Third, it is an empirically well researched case in innovation and technology management, thus containing rich additional secondary data relevant to this study (e.g., Bekkers et al., 2002a; b; Gandal et al., 2003; Sadowski et al., 2003; West, 2006; Dittrich and Duysters, 2007; Fuentelsaz et al., 2008; Bekkers and West, 2009; Di Minin and Bianchi, 2011; Hacklin et al., forthcoming).

There are and have been multiple competing technological standards in the mobile telecommunications sector over the years. This article describes the four generations commonly referred to as 1G-4G in the mobile telecommunication sector and associated with particular standards, starting off with the analog 1G Nordic telecommunication standard NMT, followed with significant overlaps by the digital 2G-4G standards being used in Europe, and becoming increasingly globalized over
generations. The case thus describes the development of four overlapping systems generations, namely NMT (1G), GSM (2G), UMTS (3G), and LTE-Advanced (4G).

This type of multi-generation study involving multiple generation shifts have several virtues in general, virtues that this paper will but cannot fully illustrate. First of all, formation of generations of product and/or process innovations over long periods of time is quite typical for high-tech industries and thus constitute an important phenomenon with ample opportunities for comparative studies across time and across industries. A second virtue is that generation shifts tend to involve more drastic or discontinuous inter-generational changes in the passage from one generation to another than more continuous intra-generational changes. Long range changes and their transitory dynamic mechanisms may thereby be more easily detected. In particular, detection of separate stages, periods or eras in business history may thereby be enabled. Third, managing technological transitions and disruptive innovation in connection with product/process generation shifts constitute one of the most important but difficult areas in strategic management. This is so since dynamic innovation-based Schumpeterian competition tends to intensify during such transitions, leading to reversals of technological leads and lags and losses of margins and market shares, including entries and exits. At the same time organizational and managerial behavior tends to become aligned with and foster incremental change in between such transitions. Fourth, many innovation studies focus on single technological transitions (or shifts) and disruptive innovation rather than a sequence of several. Such studies then tend to lose sight of the long-run dynamics of intermittent leadership with catching-up, forging ahead, falling behind, coming-back, etc., in a multi-technology context.

Empirical data and analysis

The empirical base of the case includes secondary data from a wide range of sources, including research articles, journal articles, annual reports, press releases, etc., complemented with primary data from interviews and patent statistics, all in all in order to provide as much data and as much opportunities as possible for triangulation (Jick, 1979; Langley, 1999). The interviews related to the research project behind this article, consist of well over a hundred interviews with R&D, IP, and/or standardization managers of firms such as Ericsson, Google, Huawei, Intellectual Ventures, Microsoft, Motorola, Nokia, NTT DoCoMo, Samsung, Sony, and Telia, including interviews from former projects dating back up to three decades as part of a number of studies on the developments in the telecommunication industry. These studies have resulted in a number of articles, books, and theses, where additional information is available (cf. Granstrand et al., 1992; Granstrand, 1993; Oskarsson, 1993; Bohlin and Granstrand, 1994; Oskarsson and Sjögberg, 1994; Bohlin, 1995; Granstrand et al., 1997; Granstrand, 1999; Granstrand and Lindmark, 2002; Lindmark, 2002; Granstrand and Holgersson, 2012; Holgersson, 2012; Granstrand, forthcoming). For clarity we here refer to earlier publications whenever applicable.

We use patent data as an additional source of information. The patents necessary for the use of a standard are commonly called standard-essential patents (SEPs), relating to essential technologies without which the system does not function. We here also use the more general concept of systems-essential patents/technologies, denoting patents/technologies that are necessary for the functionality of a technological system that may or may not be standardized. Technological systems consist of several complementary components (with underlying technologies and patents) and the complementarity relation between two components could be strong (essential) or weak (non-essential), and asymmetric/symmetric. Two or more components could moreover be technically substitutable, i.e., they could bilaterally replace each other. In case substitutable components are jointly essential in the sense that the system will not function without at least one of them, they are no longer individually essential.

Given this complex interdependence between essentiality, complementarity, and substitutability, it is far from clear how to define a patent as either essential or non-essential for a standard despite the increasing importance of such a distinction. Essential patents for the standards under study are currently self-reported by companies to standardization organizations such as the European Telecommunications Standards Institute (ETSI). Due to essential patents being self-reported, there are many non-essential patents being included in the records as a result of companies’ incentives to promote their roles as innovators in order to increase royalty income and bargaining power. Hence, the number of reported essential patents must be distinguished from the number of patents that are objectively evaluated essential, where evaluated essential patents are fewer and a subset of the reported essential patents. While the reported essential patents can be found in the records of ETSI, these must be objectively evaluated to receive a measure of the number of evaluated essential patents, and we here use the results from the evaluations by Fairfield Resources International (2005, 2007, 2009a, b).

Empirical data from our primary sources is combined with data from secondary sources to create the case description. Data from different (types of) sources is then used to confirm events and relationships between events (Yin, 1994; Dubois and
Evolution of the innovation ecosystem and strategic IP management in mobile telecommunications

Mobile telecommunication technologies and other information and communication technologies (ICTs) have since the 1980s revolutionized the way we work, interact, and live our lives more generally by extending our ability to process information and to communicate without physical proximity. They have been major drivers of human development and welfare during the last few decades and major drivers behind the emergence of a new type of economy with increased emphasis on intellectual assets, capital, and property (Granstrand, 2000). Due to the large-scale interdependent investments needed for ICT developments, the systemic nature of ICT-based innovations, and the network effects that are inherent characteristics of these technologies and the related products and services, ICT developments significantly benefit from inter-organizational collaborations. Such benefits include cost sharing and developing and setting standards that ensure interoperability for reaping economies of scale, scope and speed. We will here describe how the innovation ecosystem and the strategic IP management evolved in the mobile telecommunications industry. This section is primarily structured according to different technological generations. See Fig. 1 for a timeline with some of the main events.

The collaborative development in the first generation (NMT)

The development of analog mobile telecommunication technologies in Sweden in the 1960s through the 1980s is a successful case of international collaboration among national telecom operators in the Nordic countries, and a family of related companies. This innovation ecosystem consisted of system operators (incumbent service distributors), suppliers of terminals and base stations, distributors/retail dealers (of terminals), and users (such as public authorities and private users), see Fig. 2.

The organization of the development of the technical system NMT was by design mixing cooperation and competition within a hierarchy with collaborating operators as a governing body (in contrast to pure collaborative open innovation). Thus, the development was not administered within a large, integrated firm, nor did it emerge as a result of pure, competitive market forces. Experience from other countries like Germany and Japan did not speak in favor of a high degree of vertical integration, while a high degree of uncoordinated competition as in the US appeared to be less efficient in allocating resources in systems development and commercialization. Rather, it appears that the quasi-integrated, or in other terms semi-open, nature of the development organization was a critical factor behind the success of the 1G NMT system, being the earliest cellular system in the world with a unique international roaming feature and portable terminals, which quickly became adopted by various user groups, primarily for use in automobiles.

The actors together had sufficient capabilities within development and technology transfer across several key technologies, especially radio technology and switching technology. The geographical and cultural proximity between the Nordic operators and suppliers facilitated such technology transfer. In order to govern the required knowledge sharing among relevant actors, a special ‘working group’ was established as a center or hub for idea formation, implementation, and coordination in the network of actors. It was a deliberate policy to promote supplier cooperation and operator adoption
of the NMT system through wide and open communication as well as avoidance of patenting, to keep obstacles for adoption and diffusion as low as possible, at the same time as the domestic markets were protected by monopolies (Granstrand, 1993). The idea was that widespread adoption would lead to increased demand for systems equipment, which through economies of scale in turn would lead to lower costs, benefiting everyone. The NMT group paid much attention to organizing this interaction early on in an open manner in order to have an industry base ready to manufacture when demand started to grow.

The Swedish telecom operator had shifted from a (weak) pro-patent regime in the 1950s to an anti-patent regime in the 1960s (due to problems with internal inventor remunerations, i.e. internal governance problems) and then insisted on abolishing patenting when the NMT group was formed. At the same time some standard-setting organizations (SSOs) in general insisted on patent-free or at least royalty-free technical standard designs, which in turn fostered defensive/prophylactic publishing and strategic disclosure. The established telephone industry structure of the 1970s and 1980s with national operators moreover had procurement procedures that favored national suppliers. Much of the development work also took place among the operators whose output markets were protected. This led to less importance to protect and control the technological resources. All in all the incentives to patent were weak. The result was that very little in NMT was covered by patents, despite the fact that much specification and R&D work was patentable. The discarding of patenting in turn led to a certain neglect of patent matters in general (Granstrand, 1999).

A shift to patenting in the second generation (GSM)

The analog 1G market in Europe grew rapidly in the 1980s, but with different systems throughout the European countries there were limited possibilities for roaming. The Nordic countries made an exception with their joint NMT standard. As a next step in the European telecommunications development, the European Conference of Postal and Telecommunications Administrations (CEPT) decided to create a pan-European digital 2G mobile phone standard (Besen, 1990; Gandal et al., 2003). Besides ensuring roaming, a joint standard was to increase economies of scale in equipment production (Gandal et al., 2003), and later it turned out that the uniform de jure standard in Europe also implied a significantly faster diffusion than for 2G in the US (Fuentelsaz et al., 2008). In 1987, the national operators in CEPT agreed upon using the Global System for Mobile Communications (GSM) in a memorandum of understanding (MoU). The creation of ETSI by the European Commission in 1988 was followed by a transfer of the GSM standard efforts from CEPT to ETSI in 1989 (Bohlin, 1995; Lindmark, 2002; Gandal et al., 2003).

In this development more actors and technologies became involved. However, a lax patent mentality initially remained among the European national service providers as well as equipment suppliers. A kind of club mentality had developed in the telecom community in Europe, with a gentleman’s agreement to be generous to each other when it came to patents
In the late 1980s, in parallel with the liberalization of European telecommunications opening up for competing systems providers (e.g., Karlsson, 1998), the situation changed drastically when Motorola (a newcomer on the European scene) started to use patents aggressively against the GSM group, according to some as a result from information leakage from the GSM standardization work.

Bekkers et al. (2002) summarize a few other reasons for Motorola’s patent strategy. First, US firms were more aware of the role of intellectual property rights (IPRs) than European firms in the late 1980s (Granstrand, 1999; Hall and Ziedonis, 2001). Second, while Motorola heavily patented its GSM technology throughout the early development, the European firms used limited patenting while relying on their gentleman’s agreement as described above (Granstrand, 1999). Third, being a foreign actor on the European market where domestic firms were generally favored presumably meant that Motorola had to use licenses to generate income from its R&D investments. Fourth, not having digital switching systems from which continuous sales could be generated also implied an increased focus on license revenues (Iversen, 1999).

Eventually Motorola entered into cross-licensing agreements with a limited number of selected parties, including Alcatel, Ericsson, Nokia, and Siemens (Bekkers et al., 2002b). For companies with no essential or other patents to trade, licensing costs became a high barrier to entry (West, 2006; Bekkers and West, 2009) – Bekkers et al. (2002a) report total royalty fees of 29% of the costs of a GSM handset for new entrants. All in all, the conduct of Motorola triggered a new era of heightened patent awareness and activity in (European) telecommunications, leading to an irreversible track of new patenting and competitive behavior in general, on which service providers also embarked, pushed by privatization, liberalization, competition, internationalization and globalization in general. A former CEO of Ericsson described the behavior of Motorola as a shock that led to Ericsson initiating its strategic IP work, and it had a similar impact also on Nokia (Granstrand and Holgersson, 2012). Although Motorola initially collected large royalties from the GSM group, updates in the GSM standard that were necessary to Motorola eventually enabled at least Ericsson to relax its license royalty payments to Motorola.

Blocking threats in the third generation (UMTS)

The GSM standard became a major success with widespread diffusion worldwide. The European strategy for the third generation (3G) of mobile telecommunication systems (which was to support a wider range of applications) followed a similar path as for 2G, including a uniform standard and the same frequency band throughout Europe (Fuentelsaz et al., 2008). ETSI selected UMTS, a combination of W-CDMA and TD/CDMA, for the third generation, after essentially compromising between Nokia and Ericsson on one hand (W-CDMA) and Alcatel and Siemens (TD/CDMA) on the other hand (Bekkers and West, 2009). Since UMTS became a global project involving not only European actors, but also American and Asian ones, the standardization work was eventually transferred from ETSI to the worldwide standardization organization for third generation wireless communication, the 3rd generation partnership program (3GPP), adopting similar policies as ETSI (Leiponen, 2008).

Some learning how to handle IPR issues in standardization had taken place among standardization bodies, and companies had also learnt how to use patents more strategically in standardization, and the number of patents continued to increase (see Table 1). An IPR policy had been established in ETSI, and this was transferred also to 3GPP, requiring that in order to be included in standards, the patents should be licensable on fair, reasonable and non-discriminatory (FRAND) terms, a policy first proposed in 1994.3 However, the FRAND terms set no cap on royalty rates, thus leaving important residual rights (Hart, 1995) controlled by the patent holders. The actual access to the distributed technologies therefore remained limited and highly dependent on individual patent holders.

As described above ETSI decided to opt for the UMTS standard throughout Europe as a compromise between Ericsson and Nokia on the one hand and Alcatel and Siemens on the other hand. The American firm Qualcomm, supporting the competing standard CDMA 2000 following the cdmaOne 2G standard, realized that there was a great risk that UMTS would turn out to be the dominant worldwide 3G standard. UMTS was designed to be backward compatible with GSM. Since GSM dominated the 2G world market, Qualcomm feared that its own 3G technology would be in an unfavorable position, and possibly be locked out from the 3G market. US companies, especially Qualcomm, supporting CDMA 2000, then tried to influence ETSI to modify UMTS in order to make it backward-compatible with cdmaOne 2G standard and thereby simplify development of components supporting both UMTS and CDMA 2000 (Bekkers and West, 2009), threatening to otherwise block UMTS with their CDMA patents (Table 2 gives an indication of the importance of Qualcomm’s patents in UMTS/CDMA). In the end Qualcomm, who was not part of the UMTS standardization work, asked for roughly 5% of sales in royalty fees from all UMTS equipment manufacturers, claiming that its patents were essential for UMTS.

The main European companies (notably Ericsson) then threatened to block CDMA2000 in the US, unless there was reciprocity in licensing, and Ericsson started to plan US litigations. Thus, Ericsson’s stake of patents related to CDMA2000 was central here as a means for retaliation threats. The industry seemed to be deadlocked by a conflict that could not be resolved. However, Ericsson and Qualcomm negotiated a solution behind the scenes. On March 25, 1999, the two companies announced an agreement that included Ericsson acquiring Qualcomm’s systems division and associated R&D operations. By doing so,

3 A similar rule applies in the framework of the UN organization International Telecommunications Union (ITU) that has been the main international standardization body in telecom.
Lucent, Cisco, Clearwire, Intel, Samsung and Sprint Ally stated that they were to create a patent pool, the Open Patent Alliance, what we can call multi-generational patents. Such multi-generational patents preserve incumbency across generation shifts, previous generations also had gap.

As was the case with several of Ericsson’s patents in the shifts to 3G and 4G.

Top five evaluated essential patent family holders (patent shares within parentheses).

Competing standards in the fourth generation (LTE-Advanced)

The Internet Protocol based Long Term Evolution (LTE) was in 2004 suggested by NTT DoCoMo to succeed GSM and UMTS as a fourth generation standard. Although LTE is commonly called 3.9G, due to too low download rates, it is closely related to its successor LTE-Advanced, which is a true 4G standard. At this point the standardization work was well established, and patents had become a natural mechanism of governance in the innovation ecosystem. As seen from Table 1 the total number of firms with standard-essential patents (SEPs) as well as the total number of SEPs increased from 1G to 3G, and seemingly also to 4G as indicated by other data. Some SEPs read on several standards in the 1G-4G sequence and therefore become what we call multi-generational patents. Such multi-generational patents preserve incumbency across generation shifts, as was the case with several of Ericsson’s patents in the shifts to 3G and 4G.

With increasing focus on multimedia and data transfer, and decreasing focus on pure voice transfer, the mobile telecommunication industry was converging with the computer industry, from which an alternative standard, WiMAX, was being developed by Intel, Cisco and others. Once again new players, this time from the computer industry, seemed to impact the established group of firms working with GSM/UMTS/LTE, albeit in a different way than previously. On June 8, 2008, Alcatel-Lucent, Cisco, Clearwire, Intel, Samsung and Sprint Ally stated that they were to create a patent pool, the Open Patent Alliance, to give access to WiMAX-related patent licenses at limited and predictable cost. However, Ericsson and other firms in the GSM/UMTS/LTE group raised concerns about the fact that the WiMAX group, including Intel, wanted to create a communication standard from which little revenues were generated to innovators in the telecommunications field while at the same time enabling income in other areas due to dominant positions in for example processor technologies (Brismark and Alfalahi, 2008). Hence, the strategy of the new competitors was this time not related to blocking power, but rather consisted of

Ericsson gained access to Qualcomm’s cdmaOne competence and, perhaps more importantly, its patent portfolio. According to a former CEO of a competing firm, this also led to Qualcomm shifting strategies and becoming more proactively involved in the UMTS group of standardization, leaving behind its previous strategy of waiting at the side with a few essential patents in order to generate future license fees by blocking power. 

---

4 Thus, LTE could be viewed as a so-called gap filler technology, temporarily filling a time gap on the market in a shift between two major generations. Previous generations also had gap filler technologies, e.g. GSM Edge and NMT-900.

5 The 2G system GSM was in the early 1990s covered by over 2000 patents of which ca 30 were standard-blocking patents (Granstrand, 1999) while the 4G LTE system was covered by over 4.700 declared SEPs in the 2010s (EU Court of Justice Judgment, 2015 in Huawei v. ZTE).

6 Some SEPs also read across competing standards and constitute another type of multi-standard patents.

enabling access to their communication technologies on the one hand, while strictly controlling important complementary assets in order to generate revenues on the other hand. However, Ericsson clearly stated that it would not give away its WiMAX-related patents and licenses for free to the WiMAX group (Brismark and Alfalahi, 2008).

At the same time, the LTE group tried to decrease entry barriers by suggesting caps to total royalty rates, as for example expressed by Ericsson, Nokia, and Alcatel-Lucent agreeing upon a single-digit maximum aggregate royalty rate for LTE. However, the LTE group mainly relied upon bilateral agreements rather than patent pooling, and statements of royalty rates required from individual firms, as summarized by Stasik (2010), aggregated to much more than a single-digit. This indicates the difficulties with setting cap royalties in standards with multiple patent holders, inhibiting access to technologies for outsiders. This is also exemplified by Research In Motion’s (RIM) acquisition of at least 66 patents from Ericsson in 2008 for an estimated price of $172M (Stasik, 2010). Stasik argues that this was the price RIM had to pay for being able to compete in the wireless industry, since the patents would enable improved possibilities for cross-licensing.

The smartphone patent war

The previous section ended with describing RIM’s purchase of patents to increase its competitiveness in the wireless industry. Another, and much larger, purchase of patents was made by Google in 2011 when it purchased Motorola Mobility and its patent portfolio for $12.5 billion to “help protect the Android ecosystem”, as stated on Google’s webpage. Google's Android operating system and the supporting handset manufacturers competed with Apple in the smartphone industry, and Google needed patents to increase its retaliatory power to fend off litigation threats from Apple and others. A number of handset manufacturers using the Android operating system appreciated this move, as exemplified by a statement by J.K. Shin, President of Samsung’s Mobile Communications Division: “We welcome today’s news, which demonstrates Google’s deep commitment to defending Android, its partners, and the ecosystem.” A few years later, in 2014, it was announced that Google was to sell the Motorola Mobility business for only $2.9 billion (compared to the acquisition price at $12.5 billion), while keeping most of the patent portfolio, illustrating the value and importance of the patents in Google’s initial acquisition.

Patent-based competition in smartphone handsets and operating systems was thus progressing in parallel with the development of the technological standards. The patent-based competition in the area of smartphones actually became even fiercer than in the area of the underlying communication standards, and the situation developed into what has been commonly denoted a smartphone patent war. Between 2009 and 2015 a vast amount of patent infringement lawsuits were filed in various jurisdictions, including many lawsuits filed by so called patent trolls or non-practicing entities (NPEs). In a study of the US patent litigations in the smartphone industry involving at least one of the firms Motorola, Microsoft, Apple, and Samsung, Graham and Vishnubhakat (2013) investigated the 133 involved patents. Many of these were software-related, and contrary to the belief that the smartphone patent war was driven by low-quality patents, they found that among the 21 patents with court decisions including indications of validity, 17 were valid (or likely valid), and only 4 were invalid (or likely invalid). In another study, Lloyd et al. (2011) identified 298 litigated smartphone patents, and analyzed these patent litigations. In that study, Motorola, Qualcomm, and Apple were found to be the most frequent plaintiffs, while Apple, Nokia and New York Times were the most frequent defendants, see Table 3. The three most frequently disputed smartphone technologies (in number of litigated patents) were mobile data access (75), touch screen technology (27), and mobile data transmission (25). The set of top plaintiffs differ from the set of top patent holders in the standards, which illustrates that the smartphone patent war is not primarily fought with patents that are necessary to build upon the standard, but rather with patents on non-standard complementary technologies and patents.

One of the most reported fights in the smartphone patent war was the one fought by Apple and Samsung, involving patents related to their smartphone product lines (iPhone and Galaxy, respectively). Samsung’s (and many other Korean firms’) IP awareness had by and large been awakened by a patent infringement case in the 1980s involving ten US patents on DRAM held by Texas Instruments (Lee and Kim, 2010). Just as in the cases of Ericsson, Nokia, and other European firms (e.g., Granstrand, 1999; Granstrand and Holgersson, 2012), Samsung’s patent awareness was raised by having to defend itself against a major actor from the US, where patent awareness and IP strategies had been established much earlier, in turn leading to Samsung adopting a strategy for patent protection, establishing an IP division, and making efforts to increase in-house innovativeness (Lee and Kim, 2010). Thus, when the litigations in the smartphone industry picked up pace Samsung was no newcomer in the patent field. In an interview in 2011, Samsung’s head of litigation explained why Samsung was a
frequent defendant in patent litigations: It was a large, profitable, and successful firm with one of the world’s largest and broadest product ranges. The smartphone industry was furthermore a huge industry with high growth and good profitability.

Samsung, with more than 200 employees in its central IP organization and at least another 300 working with IP prosecution in their R&D organization (as of November 2011), had according to its head of litigation an IP litigation strategy that was much more defensive than offensive (interviews reveal a similar situation in Huawei, another Asian newcomer in the field). This was confirmed by a NY Times article which identified 127 smartphone patent lawsuits involving Samsung. In only four of these Samsung was the plaintiff while the other 123 lawsuits involved Samsung as defendant, many of these being filed by NPEs.\textsuperscript{15} According to its head of litigation, one of the explanations of this asymmetry was that Samsung was at very high risk when initiating litigation, since there are many opportunities to countersue such a big company, plus the fact that most competing firms also were Samsung’s customers in that they procured components from Samsung. This was for example the case with Apple.

This section started with the description of how Google in 2011 acquired Motorola Mobility. The main reason for this acquisition was to get hold of the patent portfolio of Motorola Mobility. However, Google also inherited a number of lawsuits involving Apple in this acquisition. These lawsuits were finally settled in May 2014 (in a settlement that did not include a cross-license agreement), and Google and Apple then also stated that they had “agreed to work together in some areas of patent reform”.\textsuperscript{16} Future has to tell whether this was the beginning of the end of the smartphone patent wars, or the end of the beginning.

\textbf{Analysis and discussion}

The case provides several insights about the dynamics of strategic IP management, appropriation, open innovation, and innovation ecosystems. The different technological generations constitute developments within an innovation ecosystem with clearly identifiable cooperative and competitive relations among actors as well as complementary and substitute relations among technologies and products. The innovation ecosystem was governed in a hybrid or quasi-integrated form which falls between an integrated corporate innovation system and a disintegrated market organization (cf. Granstrand, 1982; Williamson, 1991, 1996). Standard-setting can use pure market mechanisms, giving a de facto standard, or pure management mechanisms, giving an administered standard, or a mix of these mechanisms as in our case. The latter typically involves some form of open collaborative innovation between competing firms (coopetition) with some form of licensing among them.

Thus, implicit and explicit IP contracting were used for governing this collaboration. Below, we will first develop the concepts of technology governance and technology accessibility, concepts that help us better understand open innovation and appropriability in innovation ecosystems. We will then expand Teece’s PFI framework with the addition of complementary and substitute appropriability regimes.

\textbf{Governance and accessibility for appropriation and open innovation}

The development in the case explicates the concept of appropriation in innovation ecosystems, and the link between complementary assets and IP strategy was illustrated in several instances. For example, in the NMT generation, the involved actors used limited patenting, and relied upon monopolistic market protection for capturing value from R&D investments. The limited patenting endured also in the early GSM generation, when the actors relied upon implicit contracting and gentlemen’s agreements in technology development and coordination, until Motorola entered into GSM with an offensive

---

\textsuperscript{15} Fighters in a patent war, New York Times, October 7, 2012.

\textsuperscript{16} See, e.g., http://www.reuters.com/article/2014/05/16/us-apple-google-settlement-idUSBREA4FOSO20140516.
patent strategy in parallel with market liberalization. Motorola lacked some complementary assets (Iversen, 1999), and the firm was thus more dependent on the appropriability regime to profit from innovation through extensive patenting and licensing. While Teece (1986) argues that firms turn to complementary assets when the appropriability regime is weak for their technologies, this case shows that firms with relative lack of complementary assets can shape and tighten their appropriability regimes to profit from innovation. Similarly, a subsequent event in the case, the creation of the WiMAX standard, illustrates how actors with strong positions in complementary assets may want to weaken the appropriability regime of a systems technology in order to enable quick and widespread diffusion of it while relying on the complementary assets for their competitive advantage in order to profit from innovation.

While the above addresses complementary assets in general, we now move to complementary technologies more specifically. After Motorola’s disruption of the European industry, mobile telecommunications became an area for extensive patenting and explicit technology license contracting, as well as litigations. In innovation ecosystems, the appropriation strategies of different actors are highly interdependent, due to the fact that the innovations and technologies are interdependent and complementary. Once one actor starts a patent offense, everyone else has to follow to create retaliatory power (Granstrand, 1999; Hall and Ziedonis, 2001), thus resulting in an irreversible strategy shift in the innovation ecosystem (Jell et al., forthcoming).

The case moreover shows that when the mode of technology control changes so that formal legal instruments such as patents increasingly are being used, the mode of coordination may also need to change, leading to an increased level of formal patent licensing and to some extent patent litigation. We denote this as a shift from informal to formal technology governance in the innovation ecosystem, meaning that more formal control and coordination mechanisms are being used. This distinction between formal and informal technology governance relates to the distinction between explicit contracting enforceable by law, implicit contracting enforceable by markets (Klein et al., 1978) or social norms (Ostrom, 1999; Ostrom et al., 1999), and possibly no contracting at all. The distinction also relates to different streams of literature related to open innovation. While open innovation scholars generally focus on collaborating and trading and licensing IP (Chesbrough, 2003; West et al., 2014), user innovation and open source scholars have emphasized openness of the technology itself, in terms of its characteristics of patent licensing and to some extent patent litigation. We denote this as a shift from informal to formal technology governance, in which accessibility to certain technologies became limited. The aggregate license cost became a significant barrier to entry for new firms, due to many and disperse IPRs and IPR holders. Various actors in the innovation ecosystem then raised concerns about the difficulties to collect or assemble all necessary IPRs to produce telecommunication equipment, and the negative effects this had on the industry as a whole. To mitigate this the FRAND requirements to license standard-essential patents at fair, reasonable, and non-discriminatory terms were formally established through ETSI. This is an example of how institutions (such as ETSI and FRAND) can be explicitly or implicitly formed to mitigate transaction costs (North, 1990).

Although formal technology governance is often used for limiting accessibility to technologies in order for innovators to collect monopolistic rents (e.g., Arrow, 1962; Teece, 1986; Granstrand, 1999; Scotchmer, 2004), formal technology governance can thus also be used in order to ensure a high level of technology accessibility. Needless to say, intangible assets can be accessed by several parties simultaneously. The accessibility to IP is therefore an important decision dimension for IP strategy, which also makes it more complex than the management of physical property rights since owners of IPRs can use access and veto rights specifically and sometimes independently (Bel, 2013).

The accessibility of available technologies is here taken to mean a characteristic describing how easy or cheap technologies can be accessed and used by non-owners. Seen from the point of view of an actor or coalition of actors, technology accessibility could be inbound or outbound, where the inbound accessibility of a technology depends on the outbound technological accessibility of other actors. Depending on the mode of technology governance employed by an actor or coalition of actors, implicit and/or explicit contracting is used to transfer or share different types of IPRs, and the related contracts are associated with costs and prices (although they might sometime be zero) that impact the accessibility for non-owners. Technology accessibility can be governed by the use of different types of license contracts, such as the FRAND licensing requirements in our case or the General Public License (GPL) in open source software. The GPL is a formal type of contract aimed to protect the rights of users to view, modify, and distribute the technology/code (O’Mahony, 2003), i.e. to enable high accessibility.

17 Our concept of (outbound) technology accessibility is thus closely related to the distinction between selling and (freely) revealing knowledge/technologies in outbound innovation processes (Dahlander and Gann, 2010). While the former is typically related to technology trade of various forms (Arora et al., 2001), the latter emphasizes benefits from freely revealing information about innovations (Harhoff et al., 2003), meaning that “exclusive intellectual property rights to that information are voluntarily given up by the innovator” (Baldwin and von Hippel, 2011, p. 1401).

18 Note that the patent system intends not only to offer temporary private rights to a technology, but also to diffuse patented technologies through patent publications.

...
As the case illustrates, technology accessibility is an actor-specific characteristic of technologies in innovation ecosystems. Technologies can be more or less accessible to different non-owners/non-holders, and technology owners/holders can differentiate accessibility among a set of potential accessing actors. For example, a patent holder can offer a cheap license to non-competitors while not offering a license at all to competitors, who in turn might prefer another licensor of a substitute or try to invent around. Alternatively, a group of owners of complementary standard-essential technologies can cross-license their technologies more or less freely within the group, while effectively locking out anyone lacking standard-essential patents or bargaining power more generally (cf. UMTS). Similarly, firms can selectively share subsets of their technologies at low or no cost, while other technologies are licensed at high cost or kept completely proprietary (Henkel, 2006; Henkel et al., 2013), indicating differentiated accessibility not only across actors but also across technologies. This illustrates the multitude of combinatorial possibilities across open and closed forms of innovation in an innovation ecosystem.

Whether an innovation is easy or not to imitate is thus not the one main characteristic of an appropriability regime. Imitation is just one way in which an actor can access technologies developed by others. For example, several events in our case illustrate how actors in the innovation ecosystem willfully try to increase accessibility to their joint systems technology in order to decrease transaction costs and promote the continuous development and use of it through open innovation. They may do so by informal as well as by formal technology governance, as exemplified in the different technology generations.

To summarize this section, several stages or phases in the evolution of IP management can be identified in the mobile telecommunications case, see Table 4. The innovation ecosystem evolved from an initial stage of an anti-patent regime. As a response to offensive IP strategies with harsh licensing requirements and litigation threats, companies with weak patent portfolios or soft enforcement of them had to start building defensive patent portfolios and enter a stage of pro-patenting. Some of these companies then eventually started to try to appropriate more value from their expensive patent portfolios by using their portfolios more offensively, thereby entering a stage of pro-licensing and eventually pro-litigation. While the shift to formal technology governance may be irreversible, a litigious IP space, sometimes referred to as a patent war, may nevertheless eventually cool down. Recurrent contracting in an industry tends to mitigate litigation and new entrants with no record or appreciation of recurrent contracting in an industry might therefore be more litigious initially. The litigation propensity may then decrease and licensing propensity increase as time goes by and recurrent contracting develops. Litigation in itself also serves a communication and coordination function across actors, although at high cost.

### Table 4

<table>
<thead>
<tr>
<th>Stage</th>
<th>Mainly during</th>
<th>Some stage characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-patent</td>
<td>1G (NMT)</td>
<td>Market regulation; Group of Nordic actors collaborating; No/limited patenting; Strategic disclosure to foreclose patenting; Neglect of patent issues</td>
</tr>
<tr>
<td>Pro-patent</td>
<td>2G (GSM)</td>
<td>Deregulation; New entrants, upsetting the collaborative club spirit and gentleman’s agreement; Defensive patenting</td>
</tr>
<tr>
<td>Pro-licensing</td>
<td>3G (UMTS)</td>
<td>New entrants, threatening to block standards; Offensive patenting; Licensing and focus on FRAND requirements</td>
</tr>
<tr>
<td>Pro-litigation</td>
<td>4G (LTE-Advanced)</td>
<td>IP acquisitions for defensive and offensive use; New entrants; Smartphone patent war</td>
</tr>
</tbody>
</table>

**Complementary and substitute appropriability regimes**

The case illustrates how an innovation ecosystem may contain both cooperative and competitive relations between actors, and both complementary and substitute relations between technologies, patents, components, applications, and systems. These features have to date not been sufficiently included in the analysis of appropriation, open innovation, and innovation ecosystems, and it allows us to move forward in our theorizing about appropriation and IP strategy in innovation ecosystems.

In innovation ecosystems the accessibility and appropriability of a specific focal technology may be impacted not only by patents on the focal technology, but also by patents on complementary as well as on substitute technologies/standards. For example, the cross-licensing of complementary patents played an important role to enable accessibility among the main innovators in the innovation ecosystem within the different standards, and Ericsson’s patents related to the competing

---

19 The labels of the stages are short and highlight the key new feature of IP management. As with all staged evolutions, the stages are not distinct and disjoint in time but are approximate and typically overlapping with variations across nations and companies. Nor are they synchronous with the generation shifts but rather emerge during a generation, in fact due to critical events mainly related to the offensive IP strategies of new entrants, such as Motorola (2G), Qualcomm (3G), and Apple (4G).

20 The concepts of complementarity and substitutability are applicable to multiple levels of a system.

21 The innovation ecosystem concept has for example been used to emphasize how a focal firm can build and use an ecosystem of external actors to reach its business objectives (Chesbrough et al., 2014), how a focal firm can build a platform upon which external actors can innovate and design complementary products, services, and technologies (Gawer and Cusumano, 2014), and how a focal firm’s offerings are synthesized with other firms’ offerings in a coherent customer solution (Adner, 2006).
(substitute) CDMA2000 standard in the third generation was instrumental for its accessibility to strongly complementary patents related to UMTS. Thus, in innovation ecosystems complementary appropriability regimes (related to complementary components/technologies/standards) as well as substitute appropriability regimes (related to substitutes) play important roles, in addition to the specific focal appropriability regime.

In fact, non-standard complementary technologies, patents, and designs may be more valuable than standard-essential technologies and patents, despite their non-essentiality. When a systems technology is subject to standardization and collaboration across competitors, the appropriability regime for standardized technologies is often weakened by design, as illustrated by the case. This is done to mitigate transaction costs and hold-up problems, which arise partly due to large investments in specialized and/or co-specialized complementary assets that fuel opportunism (Williamson, 1975; Teece, 1986). While standardization may fortify the essentiality of the standardized technologies, thereby strengthening their appropriability, the matched licensing requirements are designed to limit hold-ups and they thereby limit appropriability. Several reports now point at patent holdout problems rather than patent holdup problems in the mobile telecommunication industry (Galetovic et al., 2017; Teece, 2017).

This creates interesting dynamics in the innovation ecosystem, as profit distribution, and consequently incentives for investments, shift from systems-essential technologies to non-standard complementary technologies, where a tight appropriability regime can be shaped through IP strategy. An important strategy to profit in innovation ecosystems is thus to invest in complementary technologies with large user/consumer value, which are in turn related to their own appropriability regimes and complementary assets. For example, the smartphone patent wars are not primarily fought with standard patents, as described above, but rather with patents on complementary technologies (such as user interfaces). The most successful firms are now those who can build on the standards and develop differentiated and proprietary complementary technologies with large value for users and consumers, i.e. firms like Apple and Samsung with products building directly on the standards or firms like Facebook and Google whose products and services are enabled by the underlying ICTs.

Finally as to appropriation, innovations and innovators may be dynamically distributed in the innovation ecosystem, especially when several system generations appear over time as illustrated in our case where the main locus of innovation shifted from incumbent service providers to equipment suppliers, new entrants and to some extent to end users. For example, in the development of the 1G system NMT the main innovators in the innovation ecosystem were the service providers, not the end users, distributors or equipment suppliers. Moreover, the main concern of the national state-owned service providers was generation of consumer surplus and social return on investment rather than private return on investment plus the prestige associated with being a technologically savvy and pioneering service provider. Similar social rather than private concerns guided the development of 2G system GSM, although now equipment and component providers played a more prominent role as innovators while the service providers continued to play the role of systems integrators (ecosystems integrators). The development was moreover speeded up in order not to let the by then ongoing deregulation in Europe end up in a situation with competing standards, which was conceived of as reducing end user appropriation of innovation value. In addition to that, international diffusion of the system was promoted by making the GSM technology easily accessible in order to promote end user appropriation. The issue was actually not an original PFI issue, i.e. the “business model” of the GSM group was not a matter of capturing as much value as possible for the innovators. Thus appropriation of innovation values and the reaping of economies of scope may be a more complex and multi-faceted dynamic issue in an innovation ecosystem than in a traditional innovation context. The single focal innovator’s appropriation problem for a single core innovation in the original PFI framework then has to be extended to accommodate an interdependent set of appropriation problems in the innovation ecosystem with different groups of actors over long periods of time.

Conclusions and managerial implications

This paper has described and analyzed the co-evolution of strategic IP management and innovation ecosystems, and in so doing extended and nuanced the debate on IP strategy, appropriation, and open innovation in dynamic and systemic innovation contexts. The case illustrates the need for including both collaborators and competitors and both complements and substitutes in the analysis and management of innovation ecosystems. Our analysis then explicates the concept of appropriability with the concepts of technology governance and technology accessibility. Rather than focusing on the ease of imitation, this perspective acknowledges the fact that firms can use IP strategy to shape weak or strong appropriability regimes in order to benefit from various forms of open and closed innovation. Our focus on complementary and substitute technologies moreover contributes to the PFI framework (Teece, 1986) within dynamic and systemic innovation contexts.

More specifically, we identify the role of complementary and substitute appropriability regimes for the appropriation of a focal innovation and for competitiveness in general. When systems-essential technologies are standardized and/or distributed across actors there is often a need to increase technology accessibility to improve the efficiency of the ecosystem. This in turn leads to a weakened appropriability regime of systems-essential technologies and a need to turn to complementary assets to profit from innovation (Teece, 1986). One management solution is then to invest in complementary technologies and tighten the related complementary appropriability regimes.

When the management solution to weak appropriability regimes for systems-essential technologies is to turn to investments in complementary technologies with tighter appropriability regimes, an economic consequence is
underinvestment in systems-essential technologies. This in turn indicates the need for including complementary and substitute assets and appropriability regimes also in policy analysis. This is an area where more research is needed.

These conceptual and theoretical developments have several implications for managers. First, the sequence of stages of IP management outlined above implies an extension of traditional IP management mainly oriented around patent protection, prosecution, monitoring, clearance, infringement and related matters to involve also patent licensing, acquisition, and litigation. This evolution is probably unavoidable as interdependent technologies and patents become more abundant and dispersed in an industry with higher stakes involved, which calls for foresight, competence enhancement, and preparedness in IP management. The evolution is moreover likely irreversible due to the arms race logic in patent portfolio races, which calls for more offensive and value-oriented IP management, in short shifting the focus of IP management from protection to value.

Second, increasingly multi-technological and multi-actor systems technologies call for extending IP strategy attention to the larger innovation ecosystem with its cooperative and competitive actor relations and its dispersed complementary and substitute assets, especially if compatibility standards are or could be expected to be heavily involved. Our focus on complementary technologies and complementary appropriability regimes as well as substitute technologies and substitute appropriability regimes incorporate the systemic nature of technologies that must be considered when shaping appropriability regimes. Multiple levels of complementary technologies are conceivable, each of which has its related set of complementary assets, in a system of assets and technologies in recurring shapes similar to fractals. Consequently, appropriation is systemic and must be treated systematically.

Third, the systemic nature of appropriability regimes enables strategic combinations of different types of appropriability regimes, in other words combinations of open and closed modes of innovation. This means that technology accessibility can be differentiated across different technologies and across different actors in order to balance the benefits from complementors and complements with the drawbacks of competition and substitutes.

Fourth, when engaged in technological standards, investments and patenting in standard technologies are central. However, when technologies are formally standardized and/or widely distributed across actors, there is typically a need to allow for increased accessibility to systems-essential technologies, which in turn weakens the related appropriability regime. Therefore, it becomes increasingly important to invest in and patent complementary technologies in order to profit from licensing or from sales of differentiated products and services. Thus, when a technology becomes standardized with related licensing regulations, proactive patenting in both the standard-essential and the non-standard complementary technologies is important to build competitive advantage.

Fifth and finally, control of complementary and substitute technologies through patents, especially through multi-generational and multi-standard patents, is important in standardization (in which ample disclosures are typically required), in turn requiring more elaborate patenting, licensing and patent acquisition strategies for appropriation, e.g. for creating freedom to operate within a standard and promoting its diffusion to dominance among users and suppliers while possibly limiting freedom to operate and diffusion for competing standards. Thus, the use of complementary technologies and patents for appropriation has to be complemented with the use of substitute technologies and patents.

While our analysis focuses on IP strategy in innovation ecosystems and open innovation, it is also relevant for future research and management of coopetition (e.g., Brandenburger and Nalebuff, 1996; Gnyawali and Park, 2011; Fernandez et al., 2014; Boucknen et al., 2015) and platforms (e.g., Gawer, 2014; Gawer and Cusumano, 2014; Bogers et al., 2017), respectively. Our case is in fact an illustration of coopetition in a large innovation ecosystem, and of the important role of IP in the management of coopetition. Moreover, the management of systems-essential (platform-essential), complementary, and substitute appropriability regimes is conducive for setting up and appropriating value from platforms. The contributions in this paper may be useful in also this context. For example, following our analysis a platform owner may benefit from tightly controlling the platform technology, while enabling substitutes on resource inputs as well as on downstream complements, i.e., substitute complements, in order to increase user/consumer value and decrease user/consumer costs while at the same time appropriating a majority of the value.

We can conclude that studying and managing complex innovation ecosystems are complex tasks. Hopefully this paper can contribute to bringing some order to both types of complexities.

Acknowledgements

The work on this paper has been undertaken within the projects Intellectual property management in digitalizing businesses at Chalmers University of Technology and Management, Economics and IP Law of Open Distributed Innovation at IMIT. The financial support from Vinnova and Jan Wallanders och Tom Hedelius stiftelse (Grant P2008-0220:1 and W2013-0063:1) is gratefully acknowledged. We thank Dr. Robert Myers at Fairfield Resources International for his assistance with data collection, and David Teece, Sarah van Santen, Martin Wallin, participants of the Open and User Innovation Workshop in Vienna (2011), the R&D Management Conference in Grenoble (2012), and the Ratio Colloquium for Young Social Scientists (2014), and seminar participants at the University of Manchester, University of Amsterdam, and Aalborg University for their helpful comments.
Appendix. Patent measures

The distribution of essential patents over different actors is used to calculate the concentration of essential patents in the standards with the Herfindahl index:

\[ H = \sum_{i=1}^{N} s_i^2 \in \left[ \frac{1}{N}, 1 \right] \]

with

\[ s_i = \frac{n_i}{\sum_{j=1}^{N} n_j} \]

where

\[ N = \text{total number of firms with essential patents} \]
\[ n_i = \text{number of essential patents held by firm } i \]

The Herfindahl index can then be normalized:

\[ H^* = \frac{H - 1/N}{1 - 1/N} \in [0, 1] \]

The concentration in terms of \( H \) and \( H^* \) can be used as measures of patent distribution within the standard. The concentration lies between \( 1/N \) and 1 when using the Herfindahl index and between 0 and 1 when using the normalized index. A low concentration indicates that the patents are fairly evenly distributed over many patent holders that need to cooperate by somehow sharing their technologies. A high concentration indicates that there are a few firms holding the majority of the patents, and less boundary crossing transactions may be needed. The number of patent holders is also a useful measure of this distribution.

References


