PRIMARY RESEARCH

Identifying new technology areas based on firm's internal capabilities

Seung-Jun Shin [1] , Wonchul Seo [2*]

[1] Graduate School of Management of Technology, Pukyong National University, Busan, South Korea
[2] Division of Systems Management and Engineering, Pukyong National University, Busan, South Korea

Abstract. Previously, various studies have proposed several methods to discover new technology or product opportunities. There is, however, a problem in that they do not consider the firm's internal capabilities or take into consideration only the technical aspects by using only patent data. The search for technological opportunities should take into account the characteristics of the industry in which the technology is applied. Therefore, this study aims to present a systematic approach to identify possible opportunities for new technology areas with firm's internal capabilities taking into account the features of the industry. To do that, we first collect patent data and extract patent co-classification information from them. Second, we generate meaningful connections between technology classes by applying (ARM). Third, we combine the inter-industry linkage effects so that the connections can reflect a more industrial viewpoint. Finally, from the perspective of a specific firm, we derive new technology areas based on its internal capabilities in terms of technology classes. To show the applicability of the presented approach, we conduct a case study using patents granted in the Korean Intellectual Property Office (KIPO) between 2006 and 2014. This study is expected to contribute to suggesting an approach to identify new technology areas that a specific firm can practically utilize. Furthermore, it will be a basis for implementing a technology planning tool in that it can explore possible opportunities for new technology areas.

INTRODUCTION

Identifying new technology areas must be a key part for the firm's sustainable growth in that it explores new revenue sources. In this regard, various systematic methods have been presented to discover new technology or product opportunities such as identifying product opportunities based on patent data (Seo *et al.*, 2016) exploring technological opportunities using morphology analysis (Yoon, Park & Coh, 2014) and establishing a function-based technology opportunity discovery framework (Yoon *et al.*, 2015). They commonly rely on the patent data which have been recognized as an up-to-date reliable source of technological development (Chen & Chang, 2012). There is, however, a problem in that they do not consider the search for technological opportunities should take into account the characteristics of the industry in which the technology is applied. Therefore, this study aims to present a systematic approach to identify possible opportunities for new technology areas with firm's internal capabilities taking into account the features of the industry. The approach consists of three steps: 1) collecting patent data and extracting patent co-classification information, 2) generating meaningful connections between technology classes by applying Association Rule Mining (ARM), and 3) combining the con-

nections with the inter-industry linkage effects to derive new technology areas based on a specific firm's internal capabilities in terms of technology classes. To show the applicability of the presented approach, we conduct a case study using patents granted in the KIPO between 2006 and 2014. This study is expected to contribute to suggesting an approach to identify new technology areas that a specific firm can practically utilize. Furthermore, it will be a basis for implementing a technology planning tool in that it can explore possible opportunities for new technology areas.

GROUNDWORK
Patent Co-Classification Analysis
The attempts to identify new technology areas or classes usually utilize the patent citation information or co-classification information to measure the extent of the technological knowledge flows (Seo, 2015). However, the citation-based approach has been recognized as having limitations that cannot reflect recent trends since the latest patents have less chance to be cited by other ones (Yoon & Kim, 2011). Therefore, to derive meaningful connection relationships between technology classes, this study clarifies the knowledge flows based on the patent co-classification information.

A patent can be classified into multiple International Patent Classifications (IPCs). The co-classification analysis assumes that the idea of the invention within a patent can be applicable to various technology fields that correspond to the associated IPCs. From that, it formulates the knowledge flow relationships between technology classes. The patent co-classification analysis has been employed into various studies like measuring the technological knowledge flows that occur across industry boundaries (Verspagen, 1997). and analyzing technology impact networks to derive meaningful implications for R&D planning (Ko, Ko, Kim, Park & Yoon, 2014).

Association Rule Mining
ARM aims at finding interesting correlations between items by figuring out how frequently they have occurred together (Liu & Hsu, 2004). It creates association rules that depict the hidden dependencies between co-purchased items. Three basic measures, support, confidence, and lift are considered to determine the rules' interestingness (Kim, Lee, Seol & Lee, 2011). The support measure describes the usefulness of the discovered rules by calculating the ratio of the occurrence frequency of items in the rules to the total number of transactions. The confidence measure ex-

plains the certainty of the rules by estimating the conditional probability that the consequent items of the rules appear in transactions given that the conditional items have already appeared. The lift measure represents the correlation feature of items in the rule by calculating the ratio of the confidence measure to the support measure of the consequent items. In this study, ARM is applied to generate meaningful relationships between technology classes. The confidence measure can account for the strength of the relationships. A lot of studies related to the patent analysis have adopted ARM such as determining the difference and impact of various technologies (Altuntas, Dereli & Kusiak, 2015) and classifying textual data into different classes for industrial knowledge management (Ur-Rahman & Harding, 2012; Sundar & Al Harthi, 2015).

Inter Industry Linkage Effect Analysis
Inter-industry linkage effect analysis aims at quantitatively analyzing inter-relationships among industries based on the inter-industry relation table. The table shows the input and distribution structure in the economy. The former explains how much each industry sector spends to produce goods and services and the latter describes how many products of each industry sector are used in the form of intermediate demand or final demand (Kim, Kim & Shin, 2016).

The table also shows the backward linkage effects and the forward linkage effects that represent the impact on the entire industry sector when the final demand for a product in an industry sector occurs and the impact that an industry receives when the final demand for a product in every industry sector occurs, respectively (Kwak, Yoo & Chang, 2005). This study attempts to consider the features of the industry for identifying possible opportunities for new technology areas. To do that, the inter-industry linkage effects will be combined with the generation of meaningful relationships between technology classes. A few studies have successfully incorporated the linkage effects into their own research analyses such as investigating the role of a certain industry in the national economy (Kwak *et al.*, 2005; Yoo, Lee & Lee, 2016) and estimating economic impact of infrastructure investment (Anas, Tamin & Wibowo, 2016).

PROPOSED APPROACH
Identifying possible opportunities for new technology areas with firm's internal capabilities must be a prerequisite for the firm's sustainable growth. To do that, this study presents a systematic approach which consists of 3 steps

(Figure 1): 1) extracting patent co-classification information from the patent data, 2) generating meaningful relationships between technology classes, and 3) combining the relationships with the inter-industry linkage effects to

identify new technology areas based on a specific firm's internal capabilities in terms of technology classes.
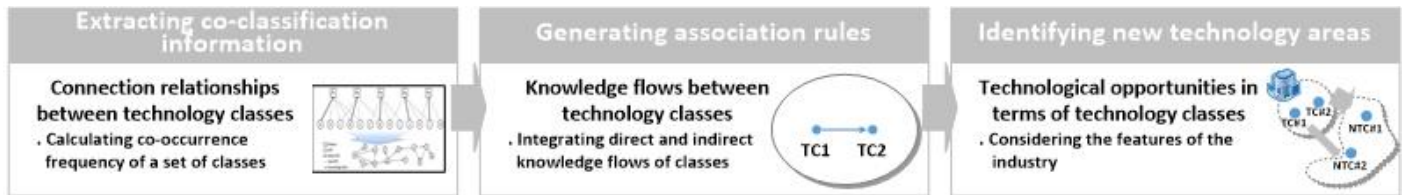

FIGURE 1. Procedure of proposed approach

Extracting Patent Co-Classification Information

This step collects patent data and extracts the co-classification information from the data. The extracted information will be input data for creating connection relationships between technology classes in the next step. To apply ARM, we should define the concept of transactions and items because ARM analyzes the co-purchased items in the transactions. This study intends to figure out the meaningful relationships between technology classes which are depicted by IPCs. Therefore, we assume that a patent document is a transaction and IPCs that the document is classified into are co-purchased items in the transaction. Under this assumption, we extract pairs of IPCs with the co-occurrence frequencies.

Generating Association Rules

This step generates connection relationships in the form of rules like A→B where A is a conditional component and B is a consequent component. Each component is represented by IPC since the purpose of this step is to reveal the hidden relationships between technology classes and the classes can be embodied in IPCs. Therefore, the generated rules can clarify the knowledge flows between the classes. To generate feasible rules applying ARM, restrictions on the three basic measures should be applied. First, for the support measure, we should figure out the number of transactions associated with the conditional and consequent technology classes to avoid creating rules for classes with a small frequency of occurrence.

It can be implemented by specifying the minimum support value. Second, for the confidence measure, we should consider the potential conviction that the conditional classes can derive the consequent classes to make the rules' certainty reliable. It can be realized by defining

the minimum confidence value. Finally, for the lift measure, a value greater than 1 indicates a positive correlation so a constraint should be imposed to select only those rules for which the lift value is greater than 1.

Combining Industry Linkage Effects

This step combines the inter-industry linkage effects so that the generated connections between technology classes can reflect a more industrial viewpoint. The certainty of a connection between classes can be represented by the confidence value of the corresponding rule. This step tries to reflect the linkage effects on the certainty. However, since the linkage effects are defined at the level of industry sector, criteria for mapping the industry sectors to the technology classes are required. KIIP has provided these mapping criteria as a concordance table (Korea Institute of Intellectual Property, 2009).

So this study uses the concordance to convert the linkage effects defined at the level of industry sector into the level of technology classes. A rule has only one confidence value but each class associated with the rule has two linkage effect values; backward and forward effects. The backward effects indicate the impact of a demand on all other industry sectors and the forward effects represent the impact of demand in all industry sectors on individual sector. Therefore, for a rule like A→B, the rule's confidence value is aggregated with the forward effect of A and the backward effect of B to draw a comprehensive relationship between A and B. It enables the discovery of new technology areas that can be derived based on the capabilities of the conditional technology class. However, this is a rough identification that excludes the perspective of a firm exploring new technology areas. Since each firm has its own capabilities, it is necessary to consider the level of capabilities for each

technology class. Therefore, we first gather IPCs that the patents of an individual firm are classified into. These can be regarded as the firm's internal capabilities represented at the level of technology classes.

And then, we reflect the extent of these internal capabilities on the comprehensive relationships mentioned above. This will ultimately lead to discovery of new feasible technology areas that can be accessed based on the firm's internal capabilities.

ILLUSTRATION
Patent Co-Classification Information

To show the applicability of the proposed approach, we conduct a case study using the patents granted in the KIPO between 2006 and 2014. The total number of patents is 919,166 but to extract patent co-classification information, we only need patents that are classified into multiple IPCs. Considering the IPC sub-class level (i.e. 4-digit), we finally get 277, 461 patents that are used in this case study. The total number of paired IPCs is 24,881 and the number of their co-occurrence frequencies is 425,079. These frequencies will be input for generating association rules.

Connection Rules

Applying ARM to the co-occurrence frequencies of the paired IPCs, we generate association rules between IPCs. To determine the rules' interestingness, we specify constraints on the three basic measures. First, to examine the usefulness, the minimum support value is set. If the support values for the conditional and consequent IPCs exceed the pre-defined minimum value, we think that the relationship between them is sufficiently usable. To obtain enough association rules for our analysis, we set the minimum support value as 0.5%. Second, to examine the certainty, the minimum confidence value can be set. However, we will aggregate the confidence value with the inter-industry linkage effects to draw comprehensive relationships, and moreover, will also consider a firm's internal capabilities together to explore new feasible technology areas for that firm. Thus, this study does not place any constraint on the confidence value because even a relationship with a low certainty can become meaningful by aggregating other factors. Finally, to examine the correlation, we choose only those rules for which the lift value is greater than 1. From these constraints, we obtain 1,332 association rules (Table 1).

TABLE 1 . Generated association rules

Conditional IPCs	Frequency (Support)	Consequent IPCs	Frequency (Support)	Lift	Confidence
A61Q	4,691(1.69%)	A61K	17,609(6.35%)	15.60	99.02%
A61P	10,119(3.65%)	A61K	17,609(6.35%)	14.89	94.51%
C08K	5,517(1.99%)	C08L	8,946(3.22%)	25.29	81.55%
C12P	1,439(0.52%)	C12N	6,527(2.35%)	30.16	70.95%
C21D	2,549(0.92%)	C22C	3,135(1.13%)	62.32	70.42%
C12Q	3,026(1.09%)	C12N	6,527(2.35%)	27.70	65.17%
F21S	3,080(1.11%)	F21V	3,945(1.42%)	43.41	61.72%
C07D	4,408(1.59%)	A61K	17,609(6.35%)	9.44	59.91%
H04J	2,154(0.78%)	H04B	20,480(7.38%)	8.08	59.66%
G03F	2,743(0.99%)	H01L	19,248(6.94%)	8.36	58.00%

Inter-Industry Linkage Effects

To identify new technology opportunities in terms of technology classes, the industrial spillover effects need to be taken into consideration. Moreover, the internal capabilities of a firm that tries to capture and utilize those opportunities should also be considered. This step aims to reflect these two factors, inter-industry linkage effects and the firm's internal capabilities, in the certainty of the association rules generated in the previous step. Using the concordance presented by KIIP, we can convert the linkage effects defined at the level of industry sector into the level

of technology classes. And then, we combine the forward linkage effect value of the conditional class in the rule and the backward linkage effect value of the consequent class by calculating their average. This combined linkage effect value is finally aggregated with the confidence value of the rule. Since the two effect values have different deviations, normalization is performed separately to prevent distortion caused by the deviations. We simply use the min-max normalization algorithm but it has a problem of normalizing the minimum value to zero. It can be perceived as not having any effect. Therefore, we arbitrarily normalize the

minimum value to 0.0001. Through this process, we can finally derive the comprehensive relationships between technology classes that reflect an industrial perspective as shown in Table 2. We now proceed to the next step to re-

flect the firm's internal capabilities. The information about the firm's technological capabilities is naturally compiled by its patent portfolio. This allows us to measure the level of capabilities based on the size of the patent portfolio.

TABLE 2 . Comprehensive relationships combining inter-industry linkage effects

Conditional IPCs	Consequent IPCs	Confidence	Forward Linkage Effect(Cond. IPCs)	Backward Linkage Effect(Cons. IPCs)	Comprehensive Relationship
G06Q	G06F	30.53%	1.0000	0.0644	0.1625
H04M	H04B	50.53%	0.2516	0.2494	0.1266
C12P	C12N	70.95%	0.0613	0.2568	0.1129
G06K	G06Q	21.00%	0.0445	1.0000	0.1097
A61Q	A61K	99.02%	0.1421	0.0783	0.1091
G06Q	H04W	21.00%	1.0000	0.0100	0.1060
B82B	B82Y	47.63%	0.0361	0.3468	0.0912
C12Q	C12N	65.17%	0.0028	0.2568	0.0846
G06Q	G06K	15.11%	1.0000	0.1182	0.0845
G06F	G06Q	15.81%	0.0363	1.0000	0.0819

TABLE 3 . Feasibility aggregating comprehensive relationships with firm's internal capabilities

Conditional IPCs	Consequent IPCs	Feasibility	Feasibility Rank	Total Sum of Feasibility
G06F	H04M	0.004773	2	0.010036
H04L	H04M	0.002177	5	
H04B	H04M	0.001351	13	
H04W	H04M	0.001106	15	
H04N	H04M	0.000376	31	
G08B	H04M	0.000253	36	
G06Q	G08G	0.004809	1	0.008995
G06F	G08G	0.001752	10	
G08B	G08G	0.000886	17	
G06K	G08G	0.000740	19	
H04N	G08G	0.000423	24	
H04W	G08G	0.000385	29	
G06Q	G06T	0.003989	3	0.006907
G06F	G06T	0.001602	11	
H04N	G06T	0.000752	18	
G06K	G06T	0.000516	22	
G08B	G06T	0.000048	68	

To concretize these capabilities from the viewpoint of the technology class, it is worth considering IPCs that the firm's patents are classified into in that they specify the technology fields in which the inventions based on the capabilities can be practically applied. It means that we can identify which technology areas the firm's capabilities focus on from the IPCs. Therefore, we first build a patent portfolio of a certain firm and then gather IPCs from the portfolio. Using the frequency information of these IPCs, we estimate the

extent of the firm's internal capabilities that will be aggregated with the comprehensive relationships drawn above. The comprehensive relationships represent new technology areas as consequent classes that can be feasible from the internal capabilities for conditional classes. The extent of the capabilities possessed by a certain firm is measured at the level of individual technology class and aggregated with the strength of the relationships. We only use the firm's capabilities for the conditional classes because the

firm needs to have enough capabilities for the conditional classes to properly derive some outputs in the consequent classes. This will lead to identify technology areas that can be derived from the internal capabilities possessed by the firm.

However, in order for this result to have a novelty in a given firm, the identified technology areas must be new to that firm. If the firm already has a high level of capability for the technology area that has been uncovered, it cannot be novel to the firm. Therefore, we try to discard some relationships in which the consequent classes are already highly associated with the firm. It ultimately allows us to discover new feasible technology areas that can be accessed based on the firm's internal capabilities. In this case study, we use an IT and system integration consulting company, LG CNS, Inc.

Specialized in smart factory, healthcare, banking and finance solutions, cloud computing and IoT as a target firm. We believe that LG CNS, Inc. is suitable for the illustration of this study due to its focused business areas. We collect the firm's patents granted in the KIPO between 2006 and 2014. The number of distinct IPCs is 52 but a number of them have few occurrence frequencies. After eliminating some IPCs that have few frequencies, we finally get 18 IPCs. They are regarded to be technology classes for which the firm has enough capabilities. We quantify the extent of capabilities of each class by normalizing the number of occurrences. And then, we aggregate it with the comprehensive relationships presented in Table 2 to finally identify new feasible technology areas that can be derived from the firm's capabilities. The final result of this case study is shown in Table 3.

RESULTS AND DISCUSSION

As shown in Table 3, our approach has presented H04M (telephonic communication), G08G (traffic control systems) and G06T (image data processing) as technology areas where the firm can make new entry based on its already existing capabilities. First, the capabilities of G06F (electric digital data processing) and H04L (transmission of digital information) will be a basis for entering into the field related to telephonic communication systems. As the target firm possesses the capabilities of digital data and information processing, it will be a good opportunity to develop communication systems between all objects based on the IoT technologies for various research and industry fields like manufacturing, home appliances, health care and smart grid. Second, the capabilities of G06Q (data processing sys-

tems or methods) and G06F will be a basis for advancement into the technology area of traffic control systems. The target firm has already established IT infrastructure for traffic management of roadway, railway and aviation like automatic fare collection and fleet management systems. These systems have mainly focused on the transportation cards, variable message signs and real-time monitoring of roads and railways.

Therefore, these accumulated experiences and skills can enable the implementation of smart traffic control systems that control and optimize the traffic flows. Finally, the capabilities of G06Q, G06F and H04N (pictorial communication) will allow the target firm to move into another technology area of image data processing. It must be useful in applications where the interpretation of image data is important like diagnosis of disease. The target firm has also already developed healthcare systems including hospital information systems and mobile healthcare solutions. Upon these experiences and capabilities, it will be a potential technological opportunity to provide a platform for medical data analysis or diagnosis consulting solutions.

CONCLUSION

This study presented a systematic approach to identify possible opportunities for new technology areas with firm's internal capabilities taking into account the features of the industry. It first extracts patent co-classification information from the patent data. Second, it generates meaningful connection rules between technology classes by applying ARM.

Finally, it creates comprehensive relationships by combining the rules with the inter-industry linkage effects and measures feasibility values by aggregating the comprehensive relationships with firm's internal capabilities. To show the applicability of the presented approach, we had conducted a case study using patents granted in the KIPO. This study is expected to contribute to suggesting an approach to identify new technology areas that a specific firm can practically utilize.

Furthermore, it will be a basis for implementing a technology planning tool in that it can explore possible opportunities for new technology areas. Despite the contributions, further challenging issues still remain. We only dealt with the identification of new technology opportunities in terms of technology classes.

However, to describe the discovered opportunities in more detail, the approach should be discussed in terms of detailed technologies by analyzing the invention descrip-

tions in patent documents. Moreover, we only used the patent co-classification information to make relationships between technology classes.

It is required to use citation or co-word information to form the relationships more meaningful.

REFERENCES

Altuntas, S., Dereli, T., & Kusiak, A. 2015. Analysis of patent documents with weighted association rules. *Technological Forecasting & Social Change*, 92: 249-262. **DOI**: 10.1016/j.techfore.2014.09.012

Anas, R., Tamin, O., & Wibowo, S.S. 2016. Applying input output model to estimate broader economic impact of transporta- tion infrastructure investment. *Journal of The Institution of Engineers*, 97(3): 323-331. **DOI**: 10.1007/s40030-016-0170-2

Chen, Y.S., & Chang, K.C. 2012. Using the entropy-based patent measure to explore the influences of related and unrelated technological diversification upon technological competences and firm performance. *Scientometrics*, 90(3): 825-841. **DOI**: 10.1007/s11192-011-0557-9

Kim, C., Lee, H., Seol, H., & Lee, C. 2011. Identifying core technologies based on technological cross-impacts: An Association Rule Mining (ARM) and Analytic Network Process (ANP) approach. *Expert Systems with Applications*, 38(10): 12559- 12564. **DOI**: 10.1016/j.eswa.2011.04.042

Kim, Y.J., Kim, C.Y., & Shin, Y.J. 2016. The effects of ubiquitous health care service on the south Korean economy: Using in- put–output analysis. *Information Systems Frontiers*, 3(5): 1-12. **DOI**: 10.5859/KAIS.2016.25.4.1

Ko, S.S., Ko, N., Kim, D., Park, H., & Yoon, J. 2014. Analyzing technology impact networks for R&D planning using patents: Combined application of network approaches. *Scientometrics*, 101(1): 917-936. **DOI**: 10.1007/s11192-014-1343-2

Korea Institute of Intellectual Property. 2009. *Proposition and application of inter industry IP relation table*. **URL**: goo.gl/iWQMHY. Last accessed on 15 October 2014.

Kwak, S.J., Yoo, S.H., & Chang, J.I. 2005. The role of the maritime industry in the Korean national economy: An input–output analysis. *Marine Policy*, 29(4): 371-383. **DOI**: 10.1016/j.marpol.2004.06.004

Liu, D.R., & Hsu, C. 2004. Project-based knowledge maps: Combining project mining and XML-enabled topic maps. *Internet Research*, 14(3): 254-266. **DOI**: 10.1108/10662240410542689

Seo, W. 2015. Analyzing causality of technological knowledge spillovers: Patent analysis approach. *Procedia Manufactur- ing*, 2(1): 485-489. **DOI**: 10.1016/j.promfg.2015.07.083

Seo, W., Yoon, J., Park, H., Coh, B.Y., Lee, J.M., & Kwon, O.J. 2016. Product opportunity identification based on internal capa- bilities using text mining and association rule mining. *Technological Forecasting and Social Change*, 105: 94-104. **DOI**: 10.1016/j.promfg.2015.07.083

Sundar, C.S., & Al Harthi, F.N.S. 2015. Impact of capital structure on firm's profitability with reference to companies listed on MSM (Muscat Securities Market). *International Journal of Business & Administrative Studies*, 1(1): 23-28. **DOI**: 10.20469/IJBAS.10004

Ur-Rahman, N., & Harding, J.A. 2012. Textual data mining for industrial knowledge management and text classification: A business oriented approach. *Expert Systems with Applications*, 39(5): 4729-4739. **DOI**: 10.1016/j.eswa.2011.09.124

Verspagen, B. 1997. Measuring intersectoral technology spillovers: Estimates from the European and US patent office databases. *Economic Systems Research*, 9(1): 47-65. **DOI**: 10.1080/09535319700000004

Yoo, J., Lee, M.K., & Lee, W.S. 2016. Asymmetrical corporate responses to economic information: Applying the firm size ef- fect. *Journal of Administrative & Business Studies*, 2(1): 29-34. **DOI**: 10.20474/-jabs2.1.3

Yoon, B., Park, I., & Coh, B.Y. 2014. Exploring technological opportunities by linking technology and products: Application of morphology analysis and text mining. *Technological Forecasting & Social Change*, 86: 287-303. **DOI**: 10.1016/j.techfore.2013.10.013

Yoon, J., & Kim, K. 2011. Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks. *Scientometrics*, 88(1): 213-228. **DOI**: 10.1007/s11192-011-0383-0

TAF
Publishing

Yoon, J., Park, H., Seo, W., Lee, J.M., Coh, B.Y., & Kim, J. 2015. Technology opportunity discovery (TOD) from existing technolo- gies and products: A function-based TOD framework. *Technological Forecasting & Social Change*, 100: 153-167.

— This article does not have any appendix. —