Winter 12-4-2009

Monitoring The Evolutionary Patterns of Technological Advances Based On the Dynamic Patent Lattice: A Modified Formal Concept Analysis Approach

Changyong Lee

Yongtae Park

MONITORING THE EVOLUTIONARY PATTERNS OF TECHNOLOGICAL ADVANCES BASED ON THE DYNAMIC PATENT LATTICE: A MODIFIED FORMAL CONCEPT ANALYSIS APPROACH

Changyong Lee[1] and Yongtae Park[2]
Department of Industrial Engineering, School of Engineering
Seoul National University, Republic of Korea
[1]akuta7@snu.ac.kr; [2]parkyt@cybernet.snu.ac.kr

Abstract

The strategic importance of monitoring changes in technology has been highlighted for achieving and maintaining firms' competitive positions. In this respect, among others, patent citation analysis has been the most frequently adopted tool. However, it is subject to some drawbacks that stem from only consideration of citing-cited information and time lags between citing and cited patents. In response, we propose a modified formal concept analysis (FCA) approach to developing dynamic patent lattice that can analyze the complex relations among patents and evolutionary patterns of technological advances. The FCA is a mathematical tool for grouping objects with shared properties based on the lattice theory. The distinct strength of FCA, *vis-á-vis* other methods, lies in structuring and displaying the relations among objects in the amount of data. The FCA is modified to take time periods into account for the purpose of technology monitoring. Specifically, patents are first collected and transformed into structured data. Next, the dynamic patent lattice is developed by executing a modified FCA algorithm based on patent context. Finally, quantitative indexes are defined and gauged to conduct a more detailed analysis and obtain richer information. The proposed dynamic patent lattice can be effectively employed to aid decision making in technology monitoring.

Keywords: Technology monitoring, patent analysis, formal concept analysis, dynamic patent lattice,

Introduction

The recent decade has seen markets shifting rapidly and unlimited proliferation of technologies, resulting in product life cycles becoming ever shorter [1][2]. It has become the norm for successful companies to have consistently to innovate for survival [3]. Under such turbulent environment, the strategic importance of monitoring changes in technology has likewise been highlighted as the technology is reckoned as a critical asset for success. Technology monitoring has been defined in many different ways. According to the EIRMA [4], it is referred to identification and assessment of technological advances critical to the firm's competitive position. Although variations may exist among researchers regarding to the definition and scope of technology monitoring, what the literature has in common is that it plays a crucial role in defending against potential threats and exploiting promising opportunities arising from technological environment [5]; consequently, there are rising attempts to formalize the technology monitoring process using suitable models, methods, and tools.

In this respect, patent documents are an ample source of technical and commercial knowledge for supporting the technology monitoring process [6][7]. Almost 80% of all technological information can be found in patent publications [8]. It can also be easily accessed through commercial and public databases. Patent analysis therefore has long been considered as a useful analytic tool for technology monitoring. An analysis of technological information in the patent documents is visualized as a patent map or network, allowing the complex patent information to be understood easily and effectively [9] and highlighting the crucial elements of knowledge on technologies, competitive positions [10][11], and infringement risk [12]. Moreover, it helps identify technological details and relationships, inspire novel industrial solutions reveal business trends, and decide investment policy [13][14].

A variety of methods for patent analysis have been introduced and among others, patent citation analysis has been the most frequently adopted tool. However, it is subject to some drawbacks as follows. Firstly, the scope of analysis and richness of potential information are limited since it takes only citing-cited information into account [15]. Secondly, regarding to the first problem, it has no capability of considering internal relationships among patents. Only existence or frequency of citations is taken into account, which may lead to superficial or even misleading analysis. Finally, it is difficult to grasp the up-to-date trends of technology because the time lags between citing and cited patents are more than ten years on average [16]. In order to overcome the limitations mentioned above, keyword-based patent analysis has been proposed as an alternative of patent citation analysis. Despite all the possibilities offered by the keyword-based patent analysis, it also has some limitations because of the difficulties in monitoring the evolutionary patterns of technological advances as time goes on. Only simple methods such as cluster and co-word analysis have been utilized. The keyword-based patent map thus limited to visualizing the static view on current status of technology as a snapshot.

In response, the primary purpose of this study is to propose a modified formal concept analysis (FCA) approach to developing dynamic patent lattice that can analyze complex relations among patents and evolutionary patterns of technological advances. The FCA is a mathematical tool for grouping objects with shared properties based on the lattice theory. The distinct strength of FCA, *vis-á-vis* other methods, lies in structuring and displaying the relations among objects in the amount of data. For the purpose of technology monitoring, the FCA is modified to take time periods into account. Specifically, patents in a technology field of interests are first collected and transformed into structured data. Next, the dynamic patent lattice is developed by executing a modified FCA algorithm on the basis of patent context. Finally, some quantitative indexes are defined and gauged to conduct a more detailed analysis and obtain richer information.

The main contributions and potential utilities of this study are twofold. First and foremost, this study theoretically contributes to the technology monitoring research, by proposing an algorithmic approach that can structure, analyze, and visualize the evolutionary patterns of technological advancement. The proposed approach overcomes the drawbacks of patent citation analysis that stem from only consideration of citing- cited information and time lags between citing and cited patents. Second, this study is exploratory in that a modified FCA algorithm is first proposed, which can be utilized in many real world problems.

The rest of this paper is organized as follows. As an introductory statement, the general background of technology monitoring, patent analysis, and FCA is first reviewed in Section 2. The proposed approach is explained in Section 3. Finally, this paper ends with conclusions in Section 4.

Background

Put theoretically, FCA is integrated together with patent analysis under a systematic framework for the purpose of technology monitoring. They are used together only rarely, and thus most readers will be comfortable with one or some, but perhaps not all of them. We therefore touch briefly on what they are and how they are combined in this study.

Technology monitoring

Technology monitoring draws more attention in both theory and practice for establishment of technological forecasting and planning. The technology monitoring may reinforce the dominance of firms in the market or open up new one [17]. It is also widely recognized that an inadequate response to technological change may lead to the demise of established company [18]. The main reasons for business failure in the market are insufficient information on trends of technology and managerial incompetence [19]. Furthermore, previous research in different industries shows that the ability to monitor technological changes is one of crucial factors in managing the risk of organizational failure [20]. For these reasons, the decision makers endeavor to discover the current status of technology, and to anticipate future events critical to the firm's competitive position [21], which is called technology monitoring.

A number of methods have been proposed to monitor the technological changes and forecast future events such as consensus method, Delphi method, structural models, and scenarios and technological vigilance. These methods can be grouped into three types: qualitative procedure, quantitative procedure, and combined procedure [22]. Firstly, the consensus method and Delphi method are qualitative procedures that primarily hinge on human intuition and individual experience. These may be distorted and biased due to the subjectivity [23][24]. Secondly, the quantitative procedure, such as structural model, may eliminate these subjective factors. This model isolates certain factors affecting the technology development process and mathematically explains some of the functional relationships among factors. However the procedures may tend to be abstractions. Omissions of certain factors that are not judged to be relevant for model construction may occur. Lastly, the combined procedure helps to identify threats and opportunities, but it requires time-consuming data collection work and is difficult to obtain objective information [19].

Patent analysis

Patent documents are as an ample source of technical and commercial knowledge [25]. The patent analysis provides a unique opportunity to satisfy the need for conceptual or qualitative analyses of technological change [26] and empirically explains most aspects of technological innovation [27]. Recent years thus have seen a huge increase in the use of patent analysis. The patent analysis has been employed for identification of economic effects of technological innovation [28], assessment of national technological competitiveness [29] assessment of individual firms' technological competitiveness [30][31], R&D activity prioritization [32], identification of technological change effects on performance [33], and identification of technological opportunity [7].

Patent data contain dozens of items for analyses, which can be grouped into two categories: structured and unstructured items [6]. The structured items are consistency in semantics and format across patent documents (e.g. patent number, filing date, inventors, and assignees) while the unstructured items are text of contents having different structures and styles (e.g. descriptions and claims). In the structured data analysis, the bibliographic fields of patent documents are utilized to explore, organize, and analyze a large amount of historical data in order that researchers can find hidden patterns to support their decision making. However, the scope of analysis and the richness of information are limited since the only bibliographic fields are employed, despite the potential utility of unstructured items. The unstructured data analysis is aimed at extracting and analyzing the technological information from the unstructured items of patent documents. Data mining techniques, especially text mining, have been widely employed for knowledge discovery from textual information.

Formal concept analysis

The FCA is a mathematical tool that can structure and visualize the relations among objects with shared properties to make them more understandable. The method was first proposed by Wille [34] based on the lattice theory of Birkoff [35]. The distinct strength of FCA, *vis-á-vis* other methods, lies in structuring and displaying the relations among objects in the amount of data. Recent years thus have seen a huge increase in the use of FCA for various problems such as ontology engineering [36], knowledge discovery in database [37], service engineering [38], collaborative recommendation [39], software engineering [40], and case-based reasoning [41].

The basic notations of FCA are summarized as follows. First, the formal context is defined as $K=(G, M, I)$, where G is a set of objects, M is an set of attributes, and I is a binary relation of G and M. The binary relation represents which attributes describe an object and vice versa. Second, the formal concept is referred to (O, A) which satisfies $intent(O)=A$ and $extent(A)=O$ where $O \subseteq G$, $A \subseteq M$ in a context (G, M, I). Finally, the concept lattice is developed by formal concepts and relations among concepts. The set of all the formal concepts of a context is denoted by $B(G, M, I)$. The structure of $B(G, M, I)$ is given by order relations between super- and sub-concept represented by $\leq$ and defined as:

$$(A_1, B_1) \leq (A_2, B_2) \text{ if } A_1 \subseteq A_2 \text{ (which is equivalent to } B_2 \subseteq B_1)$$

Proposed approach

In this section, we examine the overall process of proposed approach, giving a brief explanation of each stage at the same time. The proposed approach is comprised of five stages, as shown in Figure 1. First, a technology field of interests is selected and related patent documents are collected. Second, the patent documents are transformed into structured data for further analyses. Third, patent context is constructed in terms of year of publication, patent number, and occurrence of keywords through using text mining technique. Fourth, the modified FCA algorithm is conducted to structure and visualize the evolutionary patterns of technological advances. Finally, some quantitative indexes are defined and gauged to conduct a more detailed analysis and obtain richer information.

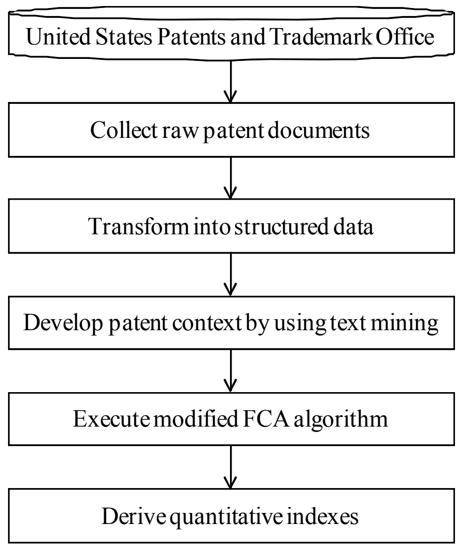

Figure 1.
Overall process

Data collection and transformation

Patent documents in a technology field of interests are collected based on various search conditions from patent database. The patent documents need to be preprocessed because they are semi-structured data in the form of electronic documents, which are merely expressed in text

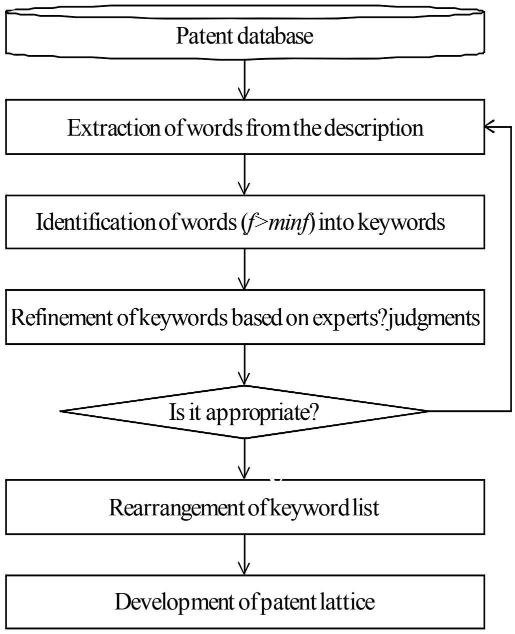

Figure 2.
Procedure of development of keyword list

format. A structured patent database is constructed for further analyses. The database constructed includes not only structured items but also unstructured ones for structured and unstructured data analyses.

Development of patent context

A patent context is developed to be utilized as an input at the next stage. The patent context consists of three parts: year of publication, patent number, and occurrence of keywords. Repetitive trials between experts and computer-based approach are required to define the keyword list. Figure 2 explains steps to elicit the keywords from documents and to fill the patent context. Text mining is first conducted to find words with high frequency and then the words are refined based on the experts' judgments. Finally, a set of final keywords are rearranged to consider the abbreviation, synonyms, singular, and plural forms of words. A patent context is exemplified in Table 1. The occurrence of keywords in a patent document is represented as a binary value. In the patent context, "V" means that the patent includes the corresponding keywords, while the blank means the patent does not.

Table 1.
Example of patent context

Year	Patent #	K1	K2	K3	K4
2007	P1	V			
	P2	V	V		
2008	P3	V			
	P4		V	V	
	P5	V	V		
2009	P6			V	
	P7			V	
	P8	V	V	V	V

Execution of modified FCA algorithm

```
RPC = Sort (patent, year, patentContext)
for i = 1 to # of patent
  tempList = Read(RPC, i)
  if ( tempList consists of all new keywords )
    Make_Node(tempList)
    // Make a node for ith patent without linkage
  else if ( tempList consists of new an existing keywords )
    Make_Node( Find_New_Key(tempList) )
    Link( Find_Related_Node(tempList) )
    // Make a node for ith patent and link up with related patents
  else
    if ( there exists nodes with same property in the same year )
      Update_Property(existingNode)
      // Add the patent to existing nodes
    else Link( Find_Related_Node(tempList) )
next i
```

Figure 3.
Pseudo-code of modified FCA algorithm

A general lattice consists of nodes and arcs that link two nodes on the basis of order relations between super-concept and sub-concept. In this case, it only provides the current status of technology without time periods. The FCA is therefore modified to take time into account to analyze the evolutionary patterns of technological advances, as shown in Figure 3. The modified FCA algorithm is carried out by considering both publication year and similarity between patents in terms of the occurrence of keywords.

The basic concepts of modified FCA algorithm are summarized as follows. Firstly, by contrast to conventional FCA, patents published earlier than the target patent are only considered to develop the dynamic patent lattice. Secondly, the order relations are derived based on the cosine similarities among concepts as well as shared properties. There are three possible types of order relations: all new keywords, combination of new and existing keywords, and all existing keywords. A new concept is generated when the target patent consists of all new keywords without existing ones. If the target patent includes the existing keywords, it is linked with the concepts having the maximum

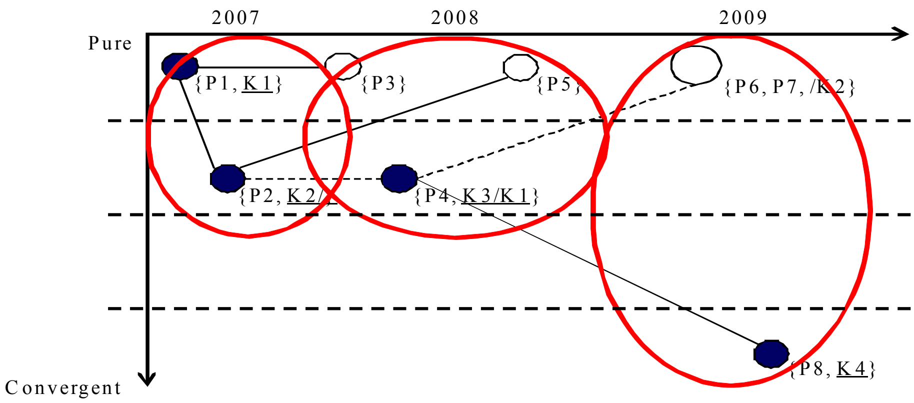

Figure 4.

Example of dynamic patent lattice

similarity and having the similarity greater than pre-defined threshold with the annotations of keywords changed (added and removed keywords). Thirdly, nodes in the dynamic patent lattice differ from one another with respect to types of concepts and number of patents in a concept. In more detail, concepts having new keywords are represented by colored circles while the concepts that only include the existing keywords are described by empty circles. The size of nodes is proportional to the number of patents that consists of a concept. Finally, regarding to the arcs, a solid line means that differences exist between concepts while a dotted line refers to the order relations between super- and sub- concepts.

The suggested dynamic patent lattice is shown in Figure 4. The horizontal axis represents the time periods while the vertical axis shows the complexity of technology. In the dynamic patent lattice, the technological advances are divided into three types: improvement, convergence, and substitute. First, the "improvement" is exemplified in the relation between P1 and P3 that share same keywords. The P3 is an advanced version of P1, which may improve some aspects of P1. Second, the "convergence" is found in the relation between P1 and P2. The P2 includes K2 besides all the aspects of P1. Finally, the "substitute" is depicted in the relations between P2 and P4. K3 is added to P4 while K1 is removed from P2.

Derivation of quantitative indexes

The proposed dynamic patent lattice shows intuitive knowledge on the patterns of technological advances and characteristics such as complexity of technology To conduct a more detailed analysis and obtain richer information, quantitative indexes, however, need to be operationally defined and gauged. Among various indexes, three major dimensions and related indexes are proposed, as summarized in Table 2. Firstly, the "importance" dimension measures the contribution of a subjective technology to the technological advances. Specifically, the local importance implies the direct influences of a subjective technology to others while the global importance considers indirect influences as well as direct ones. Secondly, the "newness" refers to the status in life cycle of a subject technology by calculating the average publication year of directly linked technologies in the past. Finally, the "promise" measures the potential attractiveness of a subject technology by gauging the emerging and declining rates of keywords.

Table 2. Dimensions of analysis and related indexes

Dimension	Index	Object and definition
Degree of importance	Technology centrality	Measure the direct influences of a subjective technology to other technologies by gauging the density of direct linkages
	Technology propagated centrality	Measure the direct and indirect influences of a subjective technology to other

Conclusions

We proposed a modified FCA-based dynamic patent lattice that can analyze the complex relations among patents and evolutionary patterns of technological advances. Patent documents in a technology field of interests are first collected and transformed into structured data. Next, a modified FCA algorithm is executed based on patent lattice. Finally, quantitative indexes are defined and gauged to conduct a more detailed analysis and obtain richer information.

The proposed approach can be utilized together with the conventional citation-based patent map, as a technology monitoring and benchmarking tool. The dynamic patent lattice and quantitative indexes may enable in-depth analysis and thus aid decision making in technology monitoring. The main contributions and potential utilities of this study are twofold. First and foremost, this study theoretically contributes to the technology monitoring research, by proposing an algorithmic approach that can structure, analyze, and visualize the evolutionary patterns of technological advancement. The proposed approach overcomes the drawbacks of patent citation analysis that stem from only consideration of citing- cited information and time lags between citing and cited patents. Second, this study is exploratory in that a modified FCA algorithm is first proposed, which can be utilized in many real world problems.

Despite all the possibilities offered by this new and algorithmic approach to technology monitoring, this study still has some limitations that stand in the way of our future research plans. Firstly, this study only focuses on technology monitoring; how to make a strategic decision has not dealt with. To fill the missing link, the proposed approach can be integrated together with technology roadmapping. Secondly, automated supporting systems need to be developed to save the time and cost and increase the efficiency of proposed approach. These topics can be fruitful areas for future research. A case study will be also included in the future research show the feasibility and utilities of proposed approach.

Acknowledgement. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2009-0085757)

References

[1] Ali, A., Krapfel, R., and Labahn, D. "Product innovativeness and entry strategy: impact on cycle time and break-even time," *The Journal of Product Innovation Management*, 12(1), 1995, pp. 54-69.

[2] D'Aveni, R. "Coping with hypercompetition: utilizing the new 7s framework," *Academy of Management Executive*, 9(3), 1995, pp. 45-57.

[3] Menor, L., Tatikonda, M., and Sampson, S. "New service development: areas for exploitation and exploration," *Journal of Operation Management*, 20(2), 2002, pp. 135-157.

[4] European Industrial Research Management Association (EIRMA), Technology monitoring for business success, 2000, Working Group 55 Report.

[5] Nosella, A., Petroni, G., and Rossella, S. "Technological change and technology monitoring process: evidence from four Italian case studies," *Journal of Engineering and Technology Management*, 25(4), 2008, pp. 321-337.

[6] Lee, S., Yoon, B., and Park Y. "An approach to discovering new technology opportunities: keyword-based patent map approach," *Technovation*, 29(6/7), 2009, pp. 481-497.

[7] Yoon, B. and Park, Y. "A systematic approach for identifying technology opportunities: keyword-based morphology analysis," *Technological Forecasting and Social Change*, 72(2), 2004a, pp. 145-160.

[8] Blackman, M. "Provision of patent information: a national patent office perspective," *World Patent Information*, 17(2), 1995, pp. 115-123.

[9] World Intellectual Property Organization (WIPO), Patent map with exercises (related), In: WIPO - MOST inter- mediate training course on practical intellectual property issues in business, 2003, Theme 16.

[10] Abraham, B.P. and Morita, S.D. "Innovation assessment through patent analysis", *Technovation*, 21(4), 2001, pp. 245–252.

[11] Liu, S. and Shyu, J. "Strategic planning for technology development with patent analysis," *International Journal of Technology Management*, 13(5/6), 1997, pp. 661–680.

[12] Daim, T., Rueda, G., Martin, H., and Gerdsri, P. "Forecasting emerging technologies: Use of bibliometrics and patent analysis," *Technological Forecasting and Social Change*, 73(8), 2006, pp. 981–1012.

[13] Campbell, R.S. "Patent trends as a technological forecasting tool," *World Patent Information*, 5(3), 1983, pp. 137–143.

[14] Jung, S. Importance of using patent information, In: World Intellectual Property Organization (WIPO) - Most intermediate training course on practical intellectual property issues in business, 2003.

[15] Yoon, B. and Park, Y. "A text-mining-based patent network: analytical tool for

high-technology trend," *Journal of High Technology*, 15(1), 2004b, pp. 37-50.
[16] Hall, B., Jaffe, A., and Trajtenberg, M. "The NBER patent citation data file: lessons, insights and methodological tools," *National Bureau of Economic Research*, 2001, Working Paper, 8498.
[17] Lichtenthaler, E. "Managing technology intelligence processes in situations of radical technological change," *Technological Forecasting and Social Change*, 74(8), 2007, pp. 1109-1136.
[18] Henderson, R. and Clark, K.B. "Architectural innovation: the reconfiguration of existing product technologies and the failure of established firms," *Administrative Science Quarterly*, 35(1), 1994, pp. 9-30.
[19] Lemos, A.D. and Porto, A.C. "Technological forecasting techniques and competitive intelligence: tools for improving the innovation process," *Industrial Management and Data Systems*, 98(7), 1998, pp. 330-337.
[20] Utterback, J.M. *Mastering the Dynamics of Innovation*, Harvard Business School Press, Boston, Massachusetts, 1994.
[21] Porter, A.L., Roper, A.T., Mason, T.W., Rossini, F.A., and Banks, J. *Forecasting and Management of Technology*, Wiley, NY, 1991.
[22] Choi, C. and Park, Y. "Monitoring the organic structure of technology based on the patent development path," *Technological Forecasting and Social Change*, 76(6), 2009, pp. 754-768.
[23] Godet, M. "Reducing the blunders in forecasting," *Futures*, 1983, 15(3), 181–192.
[24] Betz, F. *Forecasting and planning technology, Handbook of Technology Management*, McGrow-Hill, NY, 1996.
[25] Harhoff, D., Scherer, F.M., and Vopel, K. "Citations, family size, opposition and the value of patent rights," *Research Policy*, 32(8), 2003, pp. 1343-1363.
[26] Ernst, H. "Patent portfolios for strategic R&D planning," *Journal of Engineering and Technology Management*, 15(4), 1998, pp. 279-308.
[27] Jaffe, A. and Trajtenberg, M. *Patents, citations and innovations - a window on the knowledge economy*, MIT Press, Cambridge, MA, 2002.
[28] Griliches, Z. "Patent statistics as economic indicators: a survey, Journal of Economic Literature," 28(4), 1990, pp. 1661-1707.
[29] Tong, X. and Frame, J. "Measuring national technological performance with patent claims data," *Research Policy*, 23(2), 1994, pp. 133–141.
[30] Mogee, M.E. and Kolar, R.G. "International patent analysis as a tool for corporate technology analysis and planning," *Technology Analysis and Strategic Management*, 6(4), 1994, pp. 485-503.
[31] Narin, F. and Noma, E. "Patents as indicators of corporate technological strength," *Research Policy*, 16(2-4), 1987, pp. 143-155.
[32] Levitas, E.F., McFadyen, M.A., and Loree, D. "Survival and the introduction of new technology: a patent analysis in the integrated circuit industry," *Journal of Engineering and Technology Management*, 23(3), 2006, pp. 182-201.
[33] Hirschey, M. and Richardson, V.J. "Are scientific indicators of patent quality useful to investors?," *Journal of Empirical Finance*, 11(1), 2004, pp. 91-107.
[34] Wille, R. Restructuring lattice theory: an approach based on hierarchies of concepts, In: Rival, I. - *Ordered Sets*, Reidel, Dordrecht, 1982.
[35] Birkhoff, G. *Lattice Theory*, American Math. Society College Pulishers, Providence, 1973.
[36] Formica, A. "Ontology-based concept similarity in formal concept analysis," *Information Sciences*, 176(18), 2006, pp. 2624-2641.
[37] Stumme, G., Wille, R., and Wille, U. "Conceptual Knowledge discovery in databases using formal concept analysis methods," *Proceedings of the 2nd European Symposium on Principles of Data Mining and Knowledge Discovery*, 1998, pp. 450-458.
[38] Aversano, L., Bruno, M., and Penta, M. D. "Visualizing the evolution of web services using formal concept analysis," *Proceedings of Eighth International Workshop on Principles of Software Evolution*, 2005, pp. 57-60.
[39] Boucher-Ryan, P. and Bridge, D. "Collaborative recommending using formal concept analysis," *Knowledge-Based Systems*, 19(5), 2006, pp. 309-315.
[40] Tilley, T., Cole, R., Becker, P., and Eklund, P. "A survey of formal concept analysis support for software engineering activities," *Lecture notes in computer science*, 3625, 2005, pp. 250-271.
[41] Dı´az-Agudo, B. and Gonza´lez-Calero, P.A. "Classification-based retrieval using formal concept analysis," *Proceedings of the 4th International Conference on Case-Based Reasoning*, 2001, pp. 173-188.