Brockhoff, Klaus K.

Instruments for patent data analyses in business firms

Manuskripte aus den Instituten für Betriebswirtschaftslehre der Universität Kiel, No. 264

Provided in Cooperation with:
Christian-Albrechts-University of Kiel, Institute of Business Administration

Institute for Research in
Innovation Management
University of Kiel

No.264

Instruments for Patent Data Analyses in Business Firms

Klaus K. Brockhoff

Institute for Research in Innovation Management,
University of Kiel, Olshausenstraße 40, D-2300 Kiel 1, Germany,
Phone: + +-431-880-2165, Fax: + + 431 880 3349

Contents

1. Introduction 1
2. Purposes of Patent Analysis 3
3. Analyses of overall patent data of firms 5
4. Patents by technological fields 10
5. Patent Portfolio 15
6. Importance Weightings 18
7. Inventor Relations 27
8. Further Research 34

Literature 35

Abstract

The availability of patent data in data banks has largely increased the possibilities for patent analyses. In this paper we discuss overall measures of patent activity for individual firms. As the data are broken down by technology fields we introduce patent portfolios that represent the patent position of a firm relative to its competitors over all technology fields covered. We then use matrix algebra to suggest an alternative to standard procedures of importance weightings for patents, and we extent the analysis to show the relations between inventors and end-products via patents that are used in the end-products. This may have many applications, from inventor remuneration to human resource management. The paper concludes with suggestions for further research.

1. Introduction

Economists have long used patent data as indicators of inventive activity. They tested explanations and implications of inventive behavior in industries (Schmookler, 1966) or nations (Glisman/Horn, 1988) and they characterized "the nature of the technological environment in which firms operate" (Jaffe, 1989, p.87). This endeavour has been cumbersome because patent data had been available only in aggregate form or because it was necessary to scree individual patent documents. The availability of patent data from data banks has greatly advanced the possibilities of analyzing patenting behavior at those levels that appear as entries in patent documents.

A patent P could be characterized as a vector

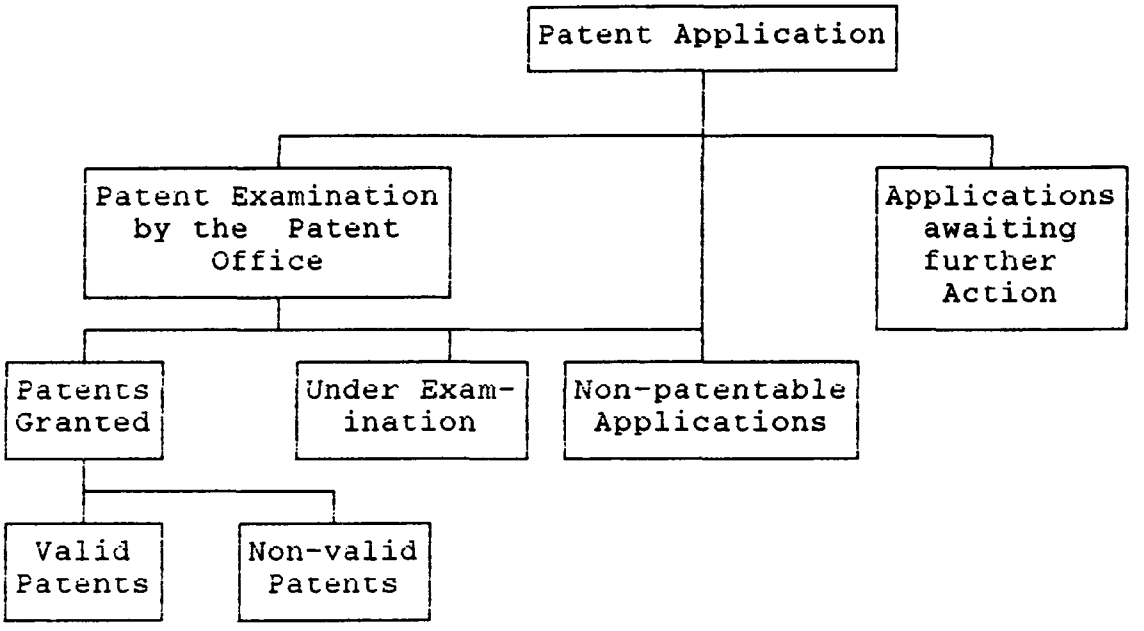
$$P = (T,C),$$

where T is a vector of calendar dates for certain events that have legal importance in the patenting process, such as the date of patent application. As the patent application is processed at the patent office, further information is added to the vector T, such as the date on which the patent is granted. C is a vector of patent characteristics, most of which are are not time-dependent. The items in C include the number of the patent, the inventor, the applicant, the classes or subclasses of a patent classification to which the patent has been assigned, citations of literature or other patents etc. As the patent is processed, the information contents of C may be changed.

At a given point in time any patent application has reached a certain stage of processing which is of interest for analysis. In Fig. 1 we sketch the different stages without going into

legal details.

Fig.1: A Sketch of the Development Stages of a Patent Application



Because of the short-term availability of application information in European patent documentations, we want to concentrate on European patent applications and patents for our purposes. As many non-Europeans file patents with European patent offices for competitive reasons, this does not seem to limit our observations severely.

The data banks vary with respect to the information that is available on the patents, and thus on the elements of the vectors T and C (Schmoch et al., 1988, pp. 77 et seq.). Partly this is due to differences in the legal process of a patent application in different countries. This means that certain types of analyses may require access to a specific data bank or that data from more than one data bank have to be pooled to

answer specific questions.

This, however, is not a major constraint. Specifically, it is not a constraint to analyzing patent data on a firm level where firms (rather than individual inventors) apply for patents.

2. Purposes of Patent Analysis

Patent analysis is of interest for many reasons: technology competition analysis, new venture evaluation, patent portfolio management, R&D management, and product area surveillance (Ashton/Sen, 1988).

It is interesting to note that no mention is made of forecasting in this list of items. In fact, Wilson (1987, p.73) cautions us not to expect too much from an analysis of patent data: "at best, patent information can tell you what your competitors were working on two to three years ago". Three counter arguments come to mind in assessing this statement:

(1) There are institutional differences in national patent systems. While the USPTO publishes only patents granted, the European Patent Office and the German Patent Office publish patent applications as soon as 18 months after the date of application, unless a patent has been granted earlier to the applicant. In a sample of 627 US patents granted to inventions in five "high technology" fields in 1985 it was found that 18.2% of the applicants had waited no longer than one year until their patent was granted, 58.1% waited up to two years, 94.3% no longer than three years (Schmoch et al., 1988, p.34). This indicates that Wilsons's statement may be more applicable

to the US patents than to those from Europe.

(2) Even if Wilson's statement is true the information that becomes available may not have been available before, and may thus be used to draw inferences by company specialists. It enhances the data that have been available before, and thus adds to the level of information.

(3) The raw information available in patents may need additionnal manipulation or interpretive analysis to use them for forecasting purposes (Campbell, 1983).

Thus, while Wilson's argument may apply to some situations, notably to those industries with extremely short product life cycles, we do not accept it altogether. Even in these industries past patent data may be used to forecast patenting behavior.

In this paper we want to suggest some ideas on the use of patent information for the evaluation of a firm's technological position vis-à-vis its competitors. This goes beyond the purposes mentioned above. It will be shown in the next paragraph for the overall patents of a firm. In the fourth chapter we look at a break-down of these data by technologies. We introduce a patent portfolio as a representation of a firm's relative patent position in the fifth chapter. In the sixth chapter we want to look at patent weightings, and after this we are interested in the relations between inventors, patents and products of a firm. This leads to a unified view of some instruments already suggested in the literature as well as to interesting possibilities for patent evaluation, inventor remuneration, and human resource management. Finally, we

present suggestions for further research.

3. Analyses of overall patent data of firms

Analyses of overall patent data of a firm may be useful to discover patenting strategies. These need not be explicitly formulated. They may as well reflect traditional behavior or institutional arrangements within firms. In a time series analysis they may also serve to document such developments as the declining role of patentable hardware inventions as compared with an increasing role of non-patentable software innovations in some industries. Here, we want to concentrate on cross-sectional data. For some of the stages mentioned we refer to Fig. 1.

Cross-sectional data are analyzed for a certain period of time to compare firms on specific variables. Five variables have proven to be valuable in cross-sectional studies of overall patent data: (1) Total number of patent applications, (2) patents per application, (3) examination rate, (4) waiting time, and (5) concentration ratio. We will consider each one of these in turn.

(1) Total number of patent applications by competing firms within a certain period. Rather than an efficency measure, where patents are related to an input measure of inventive activity, such as R&D expenditure or number of researchers, this variable documents the total technological potential of a firm for the chosen time-period. This variable is considered because Wilson seems to observe that "market leaders patent heavily to make everybody else`s lives miserable" (Wilson,

1987, p.77).

Three qualifying remarks are necessary. First, it is well-known that there are national differences in patenting behavior which have to be considered if cross-national comparisons are involved. Typically, Japanese firms take patents on one claim, while "western" firms take multi-claim patents. It should be of interest to study similar behavioral differences that may exist within a group of competing firms. Second, non-patentable applications should be excluded from the number of applications as well as those for which examinations have not been initiated within the legally defined period. Third, while some firms may have concentrated their technological strengths on a narrow field, other firms may hold technological interests in very many areas. This raises the question as to whether and how to break down the total number into fields of competitive interest. Our fifth variable will look at this on an aggregate level.

Putting this argument in a more general perspective it could be suggested to limit our variable to the "important" patents only. Later we will see that there are many different approaches to identifying important patents. Therefore, we do not want to follow this suggestion at this stage.

(2) Patents per patent application (application success rate). This relative number indicates the degree to which an applicant has been successful in having patents granted, and, therefore, will adjust for the biassed impression given by the total number of patent applications.

This number is affected by the patenting behavior over time

and the length of the observation period. If a newcomer to the industry has only recently applied for patents or if an established firm has only recently stepped up its patenting behavior this may not be reflected properly in the number as the respective patents have not yet been granted. If a firm tends to apply for patent protection for more or less trivial inventions it will show a high activity under the first variable but a low activity under this variable. In comparing the application success rate between firms interesting diagnostic questions are raised. Some of these could be answered by comparing cumulative numbers of applications between firms (Wilson, 1987, p.20).

(3) Examinations in progress over patent applications minus patents granted (examination rate). The examination is a formal procedure to study whether an invention for which a patent application has been filed qualifies for a patent according to the criteria set forth in the patent law. Therefore, examinations could apply only to those applications for which no patents have been granted yet. This explains the denominator. The numerator is primarily determined by the applicant, although third parties could also initiate an examination. Within a seven year period after application the applicant choses to initiate the examination process, unless he loses the rights from the application.

The examination rate indicates the expected future patents, specifically if multiplied with the examination rate and the total number of patent applications (as defined above). It is descriptive of the relevant contents of the "pipe-line" of the

legal process leading to a patent. As such it could be indicative of the relative threat from the patent applications of different firms.

(4) Time period between date of application and granting of the respective patent (waiting time). This time period is characterized by a weaker legal position of the applicant as compared with the position after the patent has been granted. It is therefore assumed that applicants may want to shorten this period by initiating the examination as early as possible. A counter assumption is that the applicants postpone the examination as major costs due to holding the patent arise only after the patent has been granted.

The variable can be influenced by a number of factors besides a postponement of the examination. It may be related to the patent classes in which firms take out patents, as the process of patenting may take longer in some classes than in others. This again stresses the importance of going beyond an analysis of total patent data. The variable may also be related to the organization of the patenting process in the firm. It could be speculated that a free-lancing patent attorney may have a stronger motivation to speed up the patenting procedure than a intra-firm patenting department. It has also been suggested that a firm in the local neighborhood of a patent office may benefit from its location.

(5) The frequency distribution of patents over patent classes or subclasses (concentration ratio). This variable is of interest for many reasons. It reveals different technology strategies and it may serve to explain some of the other

variables, such as waiting time or success rate. Some firms may chose to concentrate their technological efforts on a very narrow range of patent classes, while other firms may chose to patent in very many fields. Entropy is a convenient measure of order, and is therefore chosen to represent the concentration ratio. If i=1,2,...,I is the number of patent classes in which a firm patents, and if f_i is the frequency of patents in the i-th class, we define entropy normalized by number of classes as

$$\sum_{i=1}^{I} f_i \mathrm{ld} f_i \; / \; \mathrm{ld} I .$$

A value of 0 indicates concentration of patents in one class, a value of 1 indicates equal frequencies in all classes used by the firm. This measure does not reflect the number of patent classes.

It is interesting to observe these variables for a small number of competing firms in one industry. We have selected a high-tech industry with competitors from Germany, Great Britain, Japan and the US. Results are presented in Tab. 1. We see that the Japanese firm 3 has the largest number of patent applications. If industry experts should discover that this firm follows the usual Japanese habit of restricting each patent to one claim, they may deflate the number as necessary. However, even if reduced to a third of its size, this company would hold more patent claims than firm 1. Firm 1 is relatively most successful in getting patents granted. It is very remarkable that the competitor 3 does not seem to be as successful. As this reflects the past, a better indication of future developments can be gained from looking at the exami-

nation rate. Here, firms 2 and 3 appear to be most active. If the examination rate, the appplication success, and the number of applications are multiplied, we find that firm 3 may be granted 152 patents in the future, firm 2 may expect to get 74, while the traditionally strong firm 1 will get only 28 patents. This information may be further specified if we look at the remaining variables. Firms 1 and 3 are successful in getting their patents granted fast. If the number of patents per year should be of relevance, the speed of firm 1 does not outweight the higher number of threats of firms 2 or 3 even though they take longer to get their patents granted. Firm 3 is the most concentrated in patenting, while firms 1 and 5 are the least concentrated. Little mention has been made of firm 5. It is obviously a start-up. It has applied for patents that are under examination, but it has not yet been granted a patent. Therefore, we cannot calculate the waiting time. Thus, it becomes obvious that the overall patent data can be used to analyze some strategic behaviors. However, they do not allow us to look more deeply into the technological fields covered by the competing firms. Therefore, we want to further look at a more detailed level of the data.

Tab. 1: Comparison of Overall Patent Analysis Variables for Five Competing Firms (1986-1990)

	Firm 1	Firm 2	Firm 3	Firm 4	Firm 5
Patent applications	241	186	774	53	13
Application Success Rate (%)	78.4	47.8	48.0	60.4	0.0
Examination Rate(%)	69.2	83.6	79.1	61.9	92.3
Waiting Time (yrs)	4.9	6.9	5.5	7.1	-
Concentration Rate	.867	.608	.557	1.00	.887

4.Patents by technological fields

It has been suggested that the total number of patents should be analyzed only after being broken down by technological fields. A first thought is to use the International Patent Classification for this purpose. However, this seems to have a number of drawbacks if used without further qualifications. The drawbacks stem from a multitude of behaviors. First, mistakes in filing patents to the classes or subclasses as well as printing errors in the documents used as source material for the data banks seems to be frequent (Hyams, 1978, p. 240). Second, the classification of a certain patent may be different in different countries, if individually filed with different patent offices (Hyams, 1978, p. 247). Third, firms may strategically try to have their patents classified into non-appropriate classes to camouflage them from competitor technology analyses (Greif, 1982, p.29). Fourth, new technological developments may not find appropriate sub-classes in the patent classification. Therefore, they may have to be assigned to "catch-all" classes (so-called X-notation in the German Patent Office) or to classes considered to be related to the new technology.

As an indicator of the frequency of these difficulties Greif (1982, p.28) reports that the German Patent Office re-classified 14.419 patents and assigned 598 patents to the X-notation between 1977 and 1980.

For the most part it is too costly and not always technologically feasible to scan the patent description itself in order to decide on its allocation to specific technological fields.

Rather, it may be feasible to define certain technological fields with respect to the activities of a given firm. It may then be asked which sub-classes of the IPC can be related to the field, whether these should be related to the field in conjunction with certain sub-classes or should not be related in conjunction with others. This can be supported by searching for relevant keywords in the patent documents (Schmoch et al., 1988, pp. 99-125). It is obvious that in this stage "a certain amount of manual intervention is required" (Wilson,1987, p.25). If the relationship between IPC classifications and the technological fields has been formulated by using Boolean operators patents can be automatically allocated to these fields.

In the medium sized Firm 1 (see Tab. 1) we defined 9 technological fields together with company experts. After running a computerized allocation of patents to these fields we found that 13 patents were allocated to more than one field, 27 patents were incorrectly allocated, and 37 patents could not be allocated to one of the technological fields. Except of 5 patents the incorrect and the failing allocations can be attributed to incorrect allocations of patents to patent sub-classes. It has become evident that this occurs more frequently with respect to some sub-classes than to others. Therefore, an analysis for the competing Firm 2 with 174 patents (out of its 186, due to running the analyses at different points in time) led to somewhat worse percentage shares of problems (Tab.2). The results could probably have been improved if we had added a search for keywords.

Tab.2: Report on attempts at automatic allocation of patents to technological fields for 2 firms

	Firm 1		Firm 2	
	number	%	number	%
Total number of patents	241	100	174	100
Non-allocatable	37	15	52	30
(of which "others" X)	5	.	36	.
Double allocations	13	5	7	4
Allocations (Total)	217	90	129	74
(of which mis-alloc.)	27	.	11	.
(Percentages do not add to 100.)				

We suggest that a classification by technological fields is more relevant for the evaluation of competitve positions than a classification by patent classes. A measure of "closeness" of the technological position of firms (Jaffe, 1989, p. 88) may therefore be calculated for technological fields rather than for patent classes[1]. Considering the firms listed in Tab. 2 together with the Japanese competitor (Firm 3), we arrive at the measures of closeness in Tab.3. The first two competitors

Tab.3: Measure of closeness (according to Jaffe, 1989) for three competitors

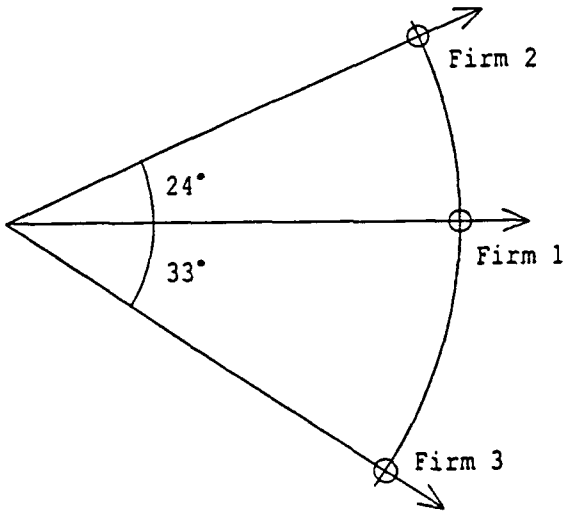
	Firm 1	Firm 2	Firm 3
Firm 1	1.0	0.9138	0.8436
Firm 2		1.0	0.6197

[1] Jaffe suggests that if k and 1 are indexes for two firms, and if f and i are interpreted as above, we should calculate:

$$P_{kl} = \frac{\Sigma f_{ki} f_{li}}{(\Sigma f_{ki}^2)^{1/2} (\Sigma f_{li}^2)^{1/2}}$$

seem to be closer to each other than both with respect to their Japanese competitor. The first firm is closer to the Japanese firm than the second. If we interpret the measure of closeness as a cosine value of angles, we can show the relationship between the firms almost perfectly on a circle of unit radius in the plane (Fig. 2).

Fig.2: Technological Relationships between three Firms in an Industry



Given the classification of patents by technological fields it makes sense to draw figures of cumulative numbers of patent applications or patents. Interestingly, this can reveal a strategy of hiding patents in remote patent classes. Firm 1 discovered in this way that its competitor Firm 3 was accumulating more patents than Firm 1 in 8 out of 9 technological fields between 1981 and 1985. The same information had not been available before. Firm 3 had been considered an actively

patenting firm, however not posing a severe competitive threat as it seemed to be patenting in too many remote classes. The additional information supports the data in Tab. 3.

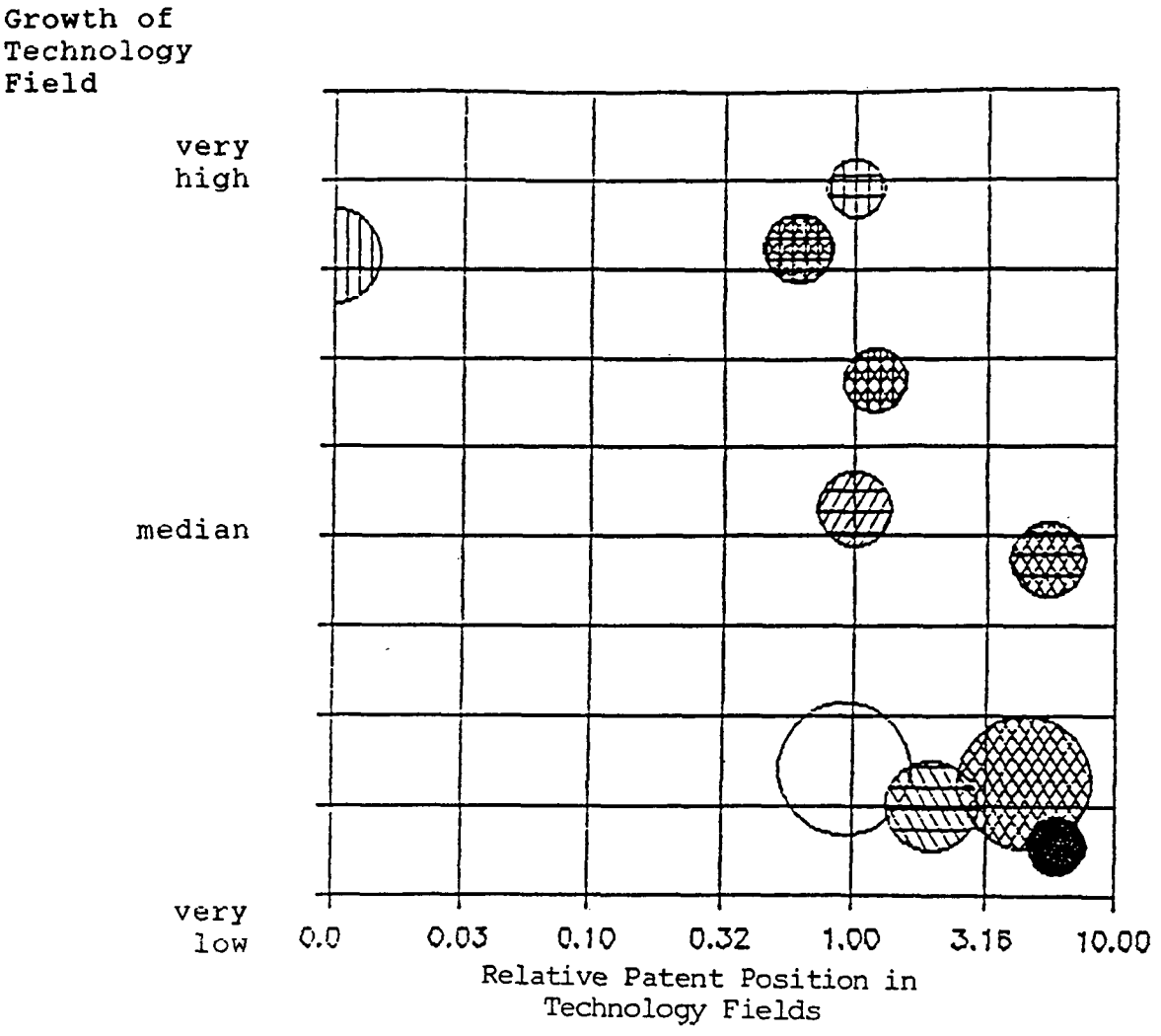
5. Patent Portfolio

Another way of interpreting a firm's patent position is by drawing its patent portfolio. This is a two-dimensional representation of the patents or the patent applications of a firm with consideration of its technological competitor's position. A portfolio is shown in Fig. 3.

On the abscissa we show the relative patent share. This is the number of patent applications by one firm relative to the number of applications of its most active competitor in a certain field of technology. Both firms are considered equally active at a value of 1. The firm considered is relatively more active at higher values and relatively less active than its competitors at lower values. As in market portfolios the abscissa values are largely determined by the behavior of the firm in question.

On the ordinate we show growth of technological fields as measured by patents. Here, we have calculated the growth during the past four years relative to the growth in the preceding 16 years, thus covering the 20 year patenting period (for other growth measures see: Marmor/Lawson/Terapane, 1979; Faust, 1989). The partition of this period is artifical. It reflects the idea that recent changes in a growth trend should become recognizable. The ordinate values are largely determined by all firms that patent in the respective technological

Fig.3: Patent Portfolio for an Electronics Manufacturer



fields.

The size of the technological fields is drawn such that it reflects the relative number of patent applications in the field as related to the total number of patent applications by the firm under consideration. This reflects the technological effort by the firm and its distribution among different fields.

As we can see in the example the sample firm has put a lot of technological effort into slowly growing fields, where it was able to attain a strong technological position relative to its competitors. Within the faster growing fields the firm is about as strong as its competitors in four fields. Given the size of these four fields the firm does not seem to put much technological emphasis on these fields. In one field the firm owns a relatively large number of patents in a strongly growing field, which is however dominated by a much stronger competitor.

Interesting questions to ask are: Should the distribution of R&D effort be redirected to achieve a different patent portfolio ? What would be an "ideal" representation of a patent portfolio ? How will the portfolio look in three or five years from the present ?

We are not yet equipped to answer the first two questions as we lack a thorough relationship between R&D, patent output and company success of a detailed level such as employed in this analysis. To answer the third question we can at least give a few hints. The future growth of the technology fields can probably be inferred from applying technological forecasting

techniques, either quantitative or qualitative, such as expert ratings.

The relative patent position can be inferred from two sources. First, the average waiting time can be used to forecast the future position of present-day patent applications. These can than be corrected by the past experience on the application success rate. As all these data are available on a firm specific basis they can be broken down to individual fields of technology. Secondly, the future stream of patent applications can be estimated from information on R&D expenditures and other sources of competitor technology intelligence. This may then lead to a forecasted patent portfolio, given the present R&D policies. Unwanted portfolio structures would then have to be corrected by changes in the R&D policy.

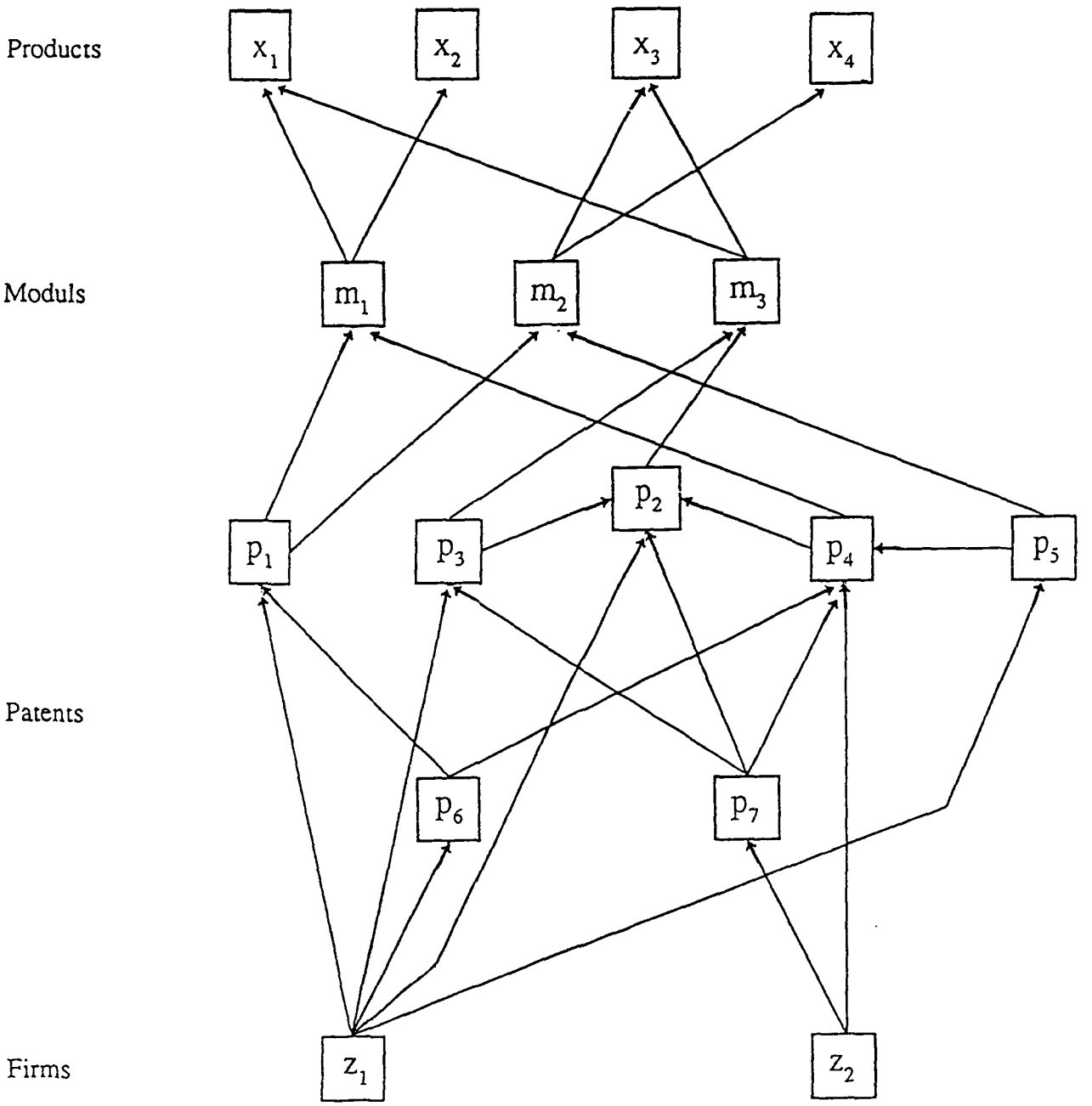
6. Importance weigthing

It is trivial to argue that simply counting the number of patents, even if broken down by categories like technological fields, does not reflect differences in importance of patents. The term importance can have many meanings to many people. Thus, citation frequencies (Carpenter/Narin/Woolf, 1981; Narin/Carpenter/Woolf, 1984; Ashton/Sen, 1988) or concentration on foreign country patents (Wilson, 1987, p.24) or consideration for patent renewals that involve increasing renewal cost with increasing patent life (Bosworth, 1973; Pakes/Schankerman, 1984) have been adopted as indirect measures of patent importance. For managers the term importance implies the long term contribution to company profits. It cannot be

taken for granted that the measures just mentioned are correlated with actual profits or expected net present values of products. Therefore, we suggest a different procedure to arrive at an indicator for patent importance. This will be demonstrated by use of a small example.

Consider a company that produces products $x_1, x_2, \ldots, x_4$ (see Fig.4). These products are produced by using certain modules of pre-assembled parts, m_1, m_2, m_3, which is depicted by arrows leading from the modules to the end-products. If we were interested to know how often each module is needed for the production of one item of the end-product we would put the respective number next to the arrow. This could be of interest in calculating licence payments. Next, we assume that five different patents are used in the modules, $p_1, p_2, \ldots, p_5$. The arrows leading from the patents to the modules indicate that a certain patent may be used in more than one module. The numbers next to the arrows may be considered as importance weights of the patents. Here, we assume that these are all equal to one. However, the weigths could take on different values assigned by subjective reasoning or by some formal procedure. Some of the patents are related to other patents (among them p_6 and p_7) by way of citation. Another set of arrows indicates whether the patents are owned by the company under consideration (z_1) or one of its competitors (z_2). Relations with inventors or other information contained in the data set C of a patent could be used at this stage.

Fig. 4: A Model of Relationships between Products, Modules, Patents, Inventors and Firms.



These relationships can now be depicted in an incidence matrix M, where the source nodes are shown in the rows and the sink nodes are shown in the columns:

M=

	x_1	x_2	x_3	x_4	m_1	m_2	m_3	p_1	p_2	p_3	p_4	p_5	p_6	p_7	z_1	z_2
x_1																
x_2																
x_3																
x_4																
m_1	1	1														
m_2			1	1												
m_3	1		1													
p_1					1	1										
p_2							1									
p_3							1		1							
p_4					1				1							
p_5						1					1					
p_6								1			1					
p_7									1	1	1					
z_1								1	1	1		1	1			
z_2											1			1		

For practical reasons we want to partition the matrix M into several sub-matrices, namely a (m,x)-production submatrix Q that shows the relationships between modules and end-products, a (p,m)-patent usage sub-matrix U that shows the relationships between patents and modules (or end-products), a (p,p)-patent-citation sub-matrix P that shows relationships among patents, and a (z,p)-patent ownership matrix A (assignees) that relates patents to their owners. In our example we have:

$$M=\begin{pmatrix} 0 & 0 & 0 & 0 \\ Q & 0 & 0 & 0 \\ 0 & U & P & 0 \\ 0 & 0 & A & Z \end{pmatrix}$$

(Z is a (z,z)-matrix with 0 entries (except if we were interested in depicting ownership relationships between firms), and 0 represents sub-matices of respective size filled with 0`s.) The different matrices can be used to produce interesting

diagnostic information. This can be demonstrated by a few examples.

Ashton/Sen (1989, p.40) present an inter(firm)citation matrix for a specific technological field. This matrix can be produced from the information given above by calculating Ax(PxA`). In our example we find

$$A\times(P\times\grave{A}) = \begin{pmatrix} 2 & 2 \\ 3 & 1 \end{pmatrix}$$

where the cited company is found in the rows and the citing company is located in the columns.

This tells us that both companies cite their own patents (which is shown by the positive values on the main diagonal). This can be easily verified: For company z_1 we find that its patent p_6 is cited by p_1, and that its patent p_3 is cited by p_2. Similarly, for company z_2 we find that its patent p_7 is cited by p_4. Furthermore, company z_2 cites two patents from company z_1, p_6 and p_5 in its patent p_4. Finally, p_7 is cited by p_2 and p_3 and p_4 is cited by p_2 in company z_1.

The incidence matrix shows only direct links between pairs of variables. We are interested to know which variables are indirectly linked to other variables. In graph theory this is equivalent to searching for the length of a path between any to nodes of a graph. Starting from the M-matrix we convert every 0 to infinity and let all values on the main diagonal equal 0. The first operation indicates that there exists no path between the variables involved, while the latter indicates that the path from a variable to itself is equal to

zero. This leads to a so-called distance matrix. For any nxn distance matrix D we define C = D ⊗ D, with

$$c_{ij} = \min_{k} (d_{ik} + d_{kj}), \qquad i,j,k=1,\ldots,n,$$

where c and d are matrix elements of C and D, respectively, and where n is the number of rows or columns of the D-matrix. If we rename C to $D^{[2]}$ we can continuously proceed in the same way to $D^{[3]}$ and so forth. In each distance matrix $D^{[n]}$ we find matrix elements that tell us the length of the path between the associated variables. As we are interested in the maximum length of such a path we proceed with the calculations until we find $D^{[m+1]}=D^{[n]}$.

We apply this procedure to the patent usage sub-matrix U and the patent-citation sub-matrix P. We find that we need to calculate only up to m=3. We then arrive at the following solution (where the columns are labeled in the same way as the rows):

$$D^{[3]} = \begin{array}{|ccccccccc||c|} \hline 0 & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & m_1 \\ \infty & 0 & \infty & \infty & \infty & \infty & \infty & \infty & \infty & m_2 \\ \infty & \infty & 0 & \infty & \infty & \infty & \infty & \infty & \infty & m_3 \\ 1 & 1 & \infty & 0 & \infty & \infty & \infty & \infty & \infty & p_1 \\ \infty & \infty & 1 & \infty & 0 & \infty & \infty & \infty & \infty & p_2 \\ \infty & \infty & 1 & \infty & 1 & 0 & \infty & \infty & \infty & p_3 \\ 1 & \infty & 2 & \infty & 1 & \infty & 0 & \infty & \infty & p_4 \\ 2 & 1 & 3 & \infty & 2 & \infty & 1 & 0 & \infty & p_5 \\ 2 & 2 & 3 & 1 & 2 & \infty & 1 & \infty & 0 & p_6 \\ 2 & \infty & 2 & \infty & 1 & 1 & 1 & \infty & \infty & 0 & p_7 \\ \hline \end{array}$$

This can be interpreted such that for instance p_5 is related to p_2 via a path of length 2 as it is cited by p_4, which in turn is cited by p_2. The patent p_6 enters module m_3 by a path of length 3, as it is cited by p_4 which in turn is cited by p_2, which then enters m_3.

We could have used the total M-matrix as a starting point which would have revealed more relationships.

In this manner it can be easily shown whether there exists a path between a certain patent and an end-product. Should the patent be owned by a competitor this could mean that the uninterrupted production of the end-product is dependent on the uninhibited right to use the competitor`s patent. In a company with a great number of patents the multitude of relationships can hardly be detected without the help of a procedure such as the one outlined above. Certainly, the instrument is a many-sided sword. Each competitor may apply the procedure and use it to his benefit, for instance to demonstrate a competing firm`s product on his own patent.

Therefore, it may be of interest to find out the number of paths by which any two variables in the M-matrix are connected with each other, irrespective of the length of a path. The more such possibilities exist to connect one variable to another, the more vulnerable would be a sink variable if the source variable is a competing company or a patent owned by a competitor. Again, it is a standard procedure to derive this information.

It is well-known from production management (Vaszonyi, 1954) that the information to solve the problem is contained in a matrix:

$$G = (I - M)^{-1},$$

where I is an identity matrix.

Given our example, we find the following result:

$$G = \left[\begin{array}{cccc|ccc|ccccccc|c}
1 & & & & & & & & & & & & & & \\
 & 1 & & & & & & & & & & & & & \\
 & & 1 & & & & & & & & & & & & \\
 & & & 1 & & & & & & & & & & & \\
\hline
1 & 1 & & & 1 & & & & & & & & & & \\
 & & 1 & 1 & & 1 & & & & & & & & & \\
1 & & 1 & & & & 1 & & & & & & & & \\
\hline
1 & 1 & 1 & 1 & 1 & 1 & & 1 & & & & & & & \\
1 & & 1 & & & & 1 & & 1 & & & & & & \\
2 & & 2 & & & & 2 & & 1 & 1 & & & & & \\
2 & 1 & 1 & & 1 & & 1 & & 1 & & 1 & & & & \\
2 & 1 & 2 & 1 & 1 & 1 & 1 & & 1 & & 1 & 1 & & & \\
3 & 2 & 2 & 1 & 2 & 1 & 1 & 1 & 1 & & 1 & & 1 & & \\
5 & 1 & 4 & & 1 & & 4 & & 3 & 1 & 1 & & & 1 & \\
\hline
9 & 4 & 8 & 3 & 4 & 3 & 5 & 2 & 4 & 1 & 2 & 1 & 1 & & 1 \\
7 & 2 & 5 & & 2 & & 5 & & 4 & 1 & 2 & & & 1 & & 1
\end{array}\right]$$

Looking at the last row we can immediately see that only the end-product x_4 is not related to any of the competitor's patents (as there is a 0 entry in the respective row and column). There are 7 ways by which product x_1 is related to the competing firm. In fact, as we know from the M-matrix firm z_2 holds rights to the patents p_4 and p_7, we can see in the respective rows and the first column that x_1 is connected twice with the patent first mentioned and five times with the latter patent.

Thus, judging by the number of paths between products or modules and competitor's patents we can detect the vulnerability of one's own products, if the legal position gives rise to doubts. Similarly, the rest of the matrix could be interpreted.

Quite obviously, the present value of the products may be quite different. Let us assume that it is contained in a vector V. This information could be used to analyze the dependability of a firm on its own and on foreign patents. As

shown above we can calculate a matrix G that gives the number of paths that exist between any two variables. (Alternatively, one could work with a matrix that contains all paths between the variables up to their maximum length.) We transform this matrix by setting any finite element in this matrix to 1 and any infinite element to 0. This matrix shall be called $\underline{G}$. It shows in its columns for each product to which elements it is related. In the rows it shows for each patent to which products it is related.

In addition to this matrix we can calculate $\underline{G}xV$. The resulting vector shows us the present value summed over all products that is dependent on any one of the patents. This information should not be mistaken for the value of a patent. However, it might serve as to rank order patents by their importance.

The following example is derived from the matrix M that was converted into G for a patent by products sub-matrix:

$$G = \begin{array}{c} \begin{matrix} x_1 & x_2 & x_3 & x_4 \end{matrix} \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \infty & 1 & \infty \\ 2 & \infty & 2 & \infty \\ 2 & 1 & 1 & \infty \\ 2 & 1 & 2 & 1 \\ 3 & 2 & 2 & 1 \\ 5 & 1 & 4 & \infty \end{bmatrix} \end{array} \qquad \underline{G} = \begin{array}{c} \begin{matrix} x_1 & x_2 & x_3 & x_4 \end{matrix} \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \end{array} \begin{matrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \\ p_7 \end{matrix}$$

$V' = (100, 20, 55, 10)$ $(\underline{G}xV)' = (185,155,155,175,185,185,175)$

This shows that some patents are related to all end-products, while others are not. A considerable share of the total net present value of 185, namely 175, is in part dependent on the patents owned by the competing firm. This share may be interpreted as an index of potential vulnerability, unless provisions have been taken to legalize the dependability.

7. Inventor relations

The relation of inventors to their patents as well as to the end-products is of considerable importance. Relations to patents can be used to describe inventor productivity, while relations to the end-products can be used as a bases for the remuneration of inventors. Information on these relations could also be used for human resource management. For instance, personnel departments may want to know about a researcher who is either an active inventor or who produces patents that serve as valuable input to other patents and thus to end-products. This information could direct efforts to retain such personnel in case they intend to leave the firm or to hire personnel from the outside. Thus, there are a number of indirect links between inventors and the competitiveness of firms. The relation between inventors and patents can be described in different ways that are most easily demonstrated if we assume that more than one inventor contributed to a patent. Then it is possible to set up a claimant`s 0-1-matrix that simply describes whether a certain inventor contributed to a patent (1) or not (0). The claimant`s matrix is different from a contributor`s matrix, that tells us something about the share with which an inventor contributed to a patent. Here, the shares of all inventors listed on a patent document have to add up to 1. The determination of these shares is an evaluative process that could be supported by the analytic hierarchy process (AHP) (Saaty, 1980).

The following example may be illustrative of the distinctions between claimant`s and contributor`s matrices, where e_1, e_2, e_3

stand for three inventors.

Claimant`s matrix E

	e_1	e_2	e_3
p_1	1		
p_2	1		
p_3	1	1	
p_4			1
p_5		1	
p_6		1	
p_7			1

Contributor`s matrix $\underline{E}$

e_1	e_2	e_3
1		
1		
0.7	0.3	
		1
	1	
	1	
		1

The difference between the two matrices is that while in the claimant`s matrix the first two inventors each hold a claim to the third patent, while in the contributor`s matrix the share of contribution of each inventor is given. Inventors 1 and 2 have successfully worked on three inventions each, while inventor 2 originated two patents.

If we multiply E`xA` we arrive at a matrix that tells us which inventor belongs to which firm and how many patents the inventors took for the firms. The result is simply:

	Firm 1	Firm 2
Inv. 1	3	0
Inv. 2	3	0
Inv. 3	0	2

Given a larger data set one may calculate on the basis of this result the number of patents per inventor in each firm as a rough measure of inventive productivity.

Next, we are interested in the relation between inventors and end-products. This can be deduced from the multiplication of a product-patent-relationship and a patent-inventor-relationship. The most obvious product-patent-relationship is the one that tells us which patents enter immediately into the products. The term immediately is meant to exclude those patents

that are merely related to end-products via citations. Fig. 4 shows that an immediate relation in this sense is given by a path of length 2 between a patent and an end-product. From a (x,p)-distance matrix $D^{[2]}$ and E we get:

$$D^{[2]} \times E = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 0 & 1 \\ 3 & 2 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

This matrix shows the number of claims of each inventor towards the end-products. Although we can read from $D^{[2]}$ that the first five patents are related in 12 ways to the four end-products, we discover 14 relations of the inventors to the end-product. This is due to the fact that two inventors were involved in one patent, namely p_3. This leads to a double-counting.

If we want to exclude this we have to adaopt $\underline{E}$ rather than E. This leads to:

$$D^{[2]} \times \underline{E} = \begin{pmatrix} 2.7 & 0.3 & 1.0 \\ 1.0 & 0.0 & 1.0 \\ 2.7 & 1.3 & 0.0 \\ 1.0 & 1.0 & 0.0 \end{pmatrix}$$

Here we find that the inventors 1 and 2 share their claims with respect to the end-products 1 and 3. Thus, any double-countings can be avoided.

The number of claims, even if corrected for double-countings, does not tell us anything about the value of the inventor`s contributions to the end-products and a possible income resulting from this contribution. Neither do we want to discuss at length the many ways of determining the value of a patent, nor do we want to deal with the legal or the behavioral issues of letting employed inventors share some of the proceeds of their inventions. We simply assume that such a share has to be calculated, as under the German law for the protection and emuneration of the employed inventor. This law mentiones three different calculation procedures that are widely employed in practice: Analogy to sales dependent licence payments, calculation of benefits to the company, estimation (Hoffmann/Bühner, 1979). Although the estimation approach is not detailed any further it could again be supported by the AHP (Saaty, 1980). Here, we concentrate on the first approach in a simplified version.

The law suggests that the individual remuneration is the product of the "value" of the invention and the individual participation share of the inventor in the invention (Richtlinien, 1959, §39; for a critical review see: Hoffmann/Bühner, 1979. p. 581). The latter is derived from the individual contribution to the definition of the problem, the contribution to the solution of the problem, and from the position within the firm. The participation share is smaller if the individual`s position is defined such as to expect him or her to be involved in inventive activities, and larger if this could not be expected. Thus, individuals in higher hierarchi-

cal positions could expect only a smaller participation share in a given invention as compared with somebody contributing in a comparable way but from a hierarchically lower position, other things being equal. We assume that this information is contained in our $\underline{E}$ matrix.

The value is the product of an industry-specific licencing rate and the sales. As the law offers a wide range of suggestions for the licencing rate, we assume that this rate could vary by products. In addition to that we should keep in mind that some products may be dependent on more than one patented invention, which would call for a distribution of the licencing share over the contributing patents. These share values are contained in a matrix V that shows a positive share value at those matrix positions where the $D^{[2]}$-matrix shows a value of 1:

$$V = \begin{pmatrix} 0.01 & 0.01 & 0.02 & 0.01 & 0 & 0 & 0 \\ 0.03 & 0 & 0 & 0.02 & 0 & 0 & 0 \\ 0.01 & 0.01 & 0.01 & 0 & 0.01 & 0 & 0 \\ 0.04 & 0 & 0 & 0 & 0.02 & 0 & 0 \end{pmatrix}$$

As we can see by summing within the rows the licence shares vary between 4% and 6% of sales. We can now calculate VxE, which gives us the individual share of the inventor in the licence share. If we have a sales vector S, which could vary between years, we find the remuneration to the inventor from S`xVxE.

For S` = (100, 200, 1000, 300) we find:

$$V\times E = \begin{pmatrix} 0.034 & 0.060 & 0.010 \\ 0.030 & 0 & 0.020 \\ 0.027 & 0.013 & 0 \\ 0.040 & 0.020 & 0 \end{pmatrix} \qquad \grave{S}\times V\times E = \begin{pmatrix} 48.4 \\ 25.0 \\ 5.0 \end{pmatrix}$$

Thus, the first inventor could receive 48.4, while the third might be entitled to a payment of 5.0. It will be recalled that the third inventor was assumed to be employed by a competing firm. Thus, this payment could be thought of as the current year`s licence payment to the competing firm, which might in turn remunerate its inventor from this payment.

In our example inventor 3 is payed very little as compared with inventor 1, because his invention that led to p_7 does not relate immediately to any one of the end-products. Similarly, some other patents appear to be rather influential in serving as a technological knowledge base for other patents, as evidenced by the citing patterns. It may be wise to reward inventors for this contribution in order to enhance creativity and the sharing of information. For this purpose one might like to identify those inventors that have contributed to the end-products by being cited in the patents that relate immediately to the end-products.

The patent-end-product sub-matrix from G gives the number of paths from any patent to any end-product. If we subtract from this matrix the respective sub-matrix that gives the number of immediate relationships between the patents and the end-products, $D^{[2]}$, we arrive at a frequency matrix of citations. One could limit this procedure by looking at the difference be-

tween $D^{[3]}$ and $D^{[2]}$, only (which would make no difference in the present example, however, this observation is not to be generalized). The transpose of this difference matrix could then be multiplied by either E or $\underline{E}$ to discover "citation stars" among the inventors. They contribute to the end-products indirectly through their patented knowledge. This is discovered by summing up over the rows in the resulting matrix (or equivalently by multiplying the resulting matrix from the right with a properly sized matrix filled with 1`s). The development of the matrices is shown in the next formula:

$$(G\text{-}D^{[2]})' \times \underline{E} = \begin{pmatrix} 0 & 0 & 1 & 1 & 2 & 3 & 5 \\ 0 & 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 & 1 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0.7 & 0.3 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0.7 & 5.3 & 6.0 \\ 0 & 3.0 & 1.0 \\ 0.7 & 3.3 & 5.0 \\ 0 & 1.0 & 1.0 \end{pmatrix}$$

Summing up, we find that the first inventor has influenced the end-product with 1.4 citations, the second inventor with 12.6 and the third inventor with 13.0. This inventor is the overall citation star, while the internal citation star is inventor 2. Going back to Fig. 4 the reason for this result is easily recognized. However, it should have become obvious that in more realistic and more complicated cases one would have difficulties in extracting the same information from either checking through the patent documents or from looking at a drawing similar to Fig. 4.

8. Further research

The ideas presented above lead to immediate extensions. Thus, total patent positions could be described by more and other variables. Until now we do not have a standard set of variables that by empirical proof could be considered as valid descriptors of the technological position of a firm. We have no tested models that would help to explain the duration of various stages in the processing of patent applications. This could be approached on the basis of an econometric techniques known as duration analysis (Kiefer, 1988).

The patent portfolio could include other variables. As with market portfolios it should be tested which variables prove to be success related.

The discussion on patent-inventor-relations could be extended by considering the relative importance of the citations in the (p,p)-triangular matrix of patent citations or by looking at relative frequencies of citations for purposes of comparing between firms or between technology fields. Given the presentation above it should not cause great difficulties to develop appropriate measures.

Finally, if more data are calculated from the patent files, it is easy to extend the ideas presented with respect to importance weightings of patents. Given an additicnal sub-matrix that relates inventors to the patents, a firm may chose the approach outlined above to relate the inventors automatically to the products and use this information as a basis for its incentive payments. Thus, there is a multitude of interesting applications to the procedures sketched here.

Literature:

Ashton,W.B./Sen,R.K.,1988, Using Patent Information in Technology Business Planning -I, Research-Technology Management, Vol.31, Nov/Dec. pp.42-46.
Bosworth,D.L.,1973, Changes in the Quality of Inventive Output and Patent Based Indices of Technological Change, Bulletin of Economic Research, Vol. 25, pp. 95-103.
Campbell,R.S.,1983, Patent trends as a technological forecasting tool, World Patent Information, Vol.5, pp.137-143.
Carpenter,M.P./Narin,F./Woolf,P.,1981, Citation rates to technologically important papers, World Patent Information, Vol. 3, pp. 160-163.
Faust,K.,1989, Neue technologische Trends im Licht der internationalen Patentstatistik und die Orientierung der deutschen Forschung, München (ifo-institut für Wirtschaftsforschung).
Glisman,H.H./Horn,E.-J.,1988, Comparative Invention Performance of Major Industrial Countries: Patterns and Explanations, Management Science, Vol. 34, pp.1169-1187.
Greif,S.,1982, Angebot und Nachfrage nach Patentinformationen. Die Informationsfunktion von Patenten. Göttingen (Schwartz).
Henn,R./Künzi,H.P.,1968, Einführung in die Unternehmensforschung, Vol. 2, Berlin et al. (Springer-Verlag).
Hoffmann,F./Bühner,R.,1979, Die Vergütung von Arbeitnehmererfindungen aus betriebswirtschaftlicher Sicht, Die Betriebswirtschaft, Vol. 39, pp.573-584.
Hyams,M.,1978,Problematik und spezielle Merkmale der Derwent Patentdienstleistungen, in: Internationales Symposium Patentinformation und Patentdokumentation, New York et al. 1978, pp. 236 et seq.
Jaffe,A.B.,1989,Characterizing the "technological position" of firms, with application to quantifying technological opportunity and research spillovers, Research Policy, Vol. 18, pp.87-97.
Kiefer,N.M.,1988, Economic Duration Data and Hazard Functions, Journal of Economic Literature, Vol. 26, pp. 646-679.
Marmor,A.C./Lawson,W.S./Terapane,J.F., 1979, The Technology Assessment and Forecasting Program of the United States Patent and Trademark Office, World Patent Information, Vol. 1, pp. 15-23.
Narin,F./Carpenter,M.P./Woolf,P.,1984, Technological Performance Assessment Based on Patents and Patent Citations, IEEE Transactions on Engineering Management, Vol. EM-31, pp. 172-183.
Pakes,A./Schankerman,M.,1984,The Rate of Obsolescence of Patents, Research Gestation Lags, and Private Rate of Return to Research Resources. In: Griliches,Z., R&D, Patents, and Productivity, Chicago, London (Univ. of Chicago Press),pp.73-88.
Richtlinien für die Vergütung von Arbeitnehmererfindungen im privaten Dienst, vom 20.Juli 1959, Bundesanzeiger No. 156.
Saaty,T.L., 1980, The Analytic Hierarchy Process, New York (McGraw Hill).
Schmoch,U. et al., 1988, Technikprognosen mit Patentindikatoren. Köln (Verlag TÜV Rheinland).
Schmookler,J.,1966, Invention and Economic Growth, Cam-

bridge/Mass. (Harvard Univ. Press).
Vaszonyi,A.,1954, The Use of Mathematics in Production and Inventory Control, Management Science, Vol.1, pp.70-85,207-233.
Wilson,R.M.,1987, Patent Analysis using Online Databases -I. Technological Trend Analysis. World Patent Information, Vol. 9, pp.18-26.
Wilson,R.M.,1987, Patent Analysis using Online Databases - II. Competitor Activity Monitoring. World Patent Information, Vol.9, pp. 73-78.