

La qualité technologique des brevets : deux lectures structurales

The technological quality of patents: two structural readings

Raffaele Anedda¹

¹ Unité d'Economie Appliquée, ENSTA Paris, Institut Polytechnique de Paris, France, raffa.anedda@gmail.com

RÉSUMÉ. L'objectif de l'article est de proposer des lectures structurales de deux indicateurs de la qualité technologique des brevets : l'originalité et la généralité. Nous les transformons pour ce faire en des objets mathématiques – des « graphes » – et nous insistons sur des amendements et prolongements inspirés par cette nouvelle approche. Dans une première lecture, nous partons des indices originels et nous leur appliquons directement une représentation sous la forme de graphes. Dans une deuxième lecture, nous partons de matrices de flux technologiques et nous reconstruisons les indices de dispersion tels qu'ils ont été conçus initialement. Nous allons même un peu plus loin : nous proposons d'autres indices de la qualité technologique à partir de ces matrices, dont un qui nous semble de portée supérieure pour apprécier cette qualité.

ABSTRACT. The aim of this article is to propose the structural readings of two indicators of the technological quality of patents: originality and generality. To do so, we transform them into mathematical objects - graphs - and insist on amendments and extensions inspired by this new approach. In the first reading, we start from the original indices and we directly apply a representation in the form of graphs to them. In the second reading, we start from the technological flow matrices and reconstruct the dispersion indices as they were originally conceived. We go even further by proposing other indices of technological quality from these matrices, including one that seems to have a greater scope for assessing this quality.

MOTS-CLÉS. Brevet, technologie, théorie des graphes.

KEYWORDS. patent, technology, graph theory.

1. Introduction

L'OCDE [OCD 13] publie en 2013 un document de travail recensant une douzaine d'indices servant à estimer les qualités économique et technologique d'un brevet. La qualité technologique d'un brevet est l'impact potentiel de cette invention sur les développements technologiques ultérieurs. L'ensemble de ces indices ont pour particularité d'être mesurables à partir de données présentes dans les brevets évalués eux-mêmes, dans les citations de brevets antérieurs qu'ils font, ou encore dans les brevets ultérieurs les citant. L'intérêt de s'en tenir aux informations contenues dans les brevets est que cela assure une certaine homogénéité des informations et une comparabilité dans le temps et dans l'espace.

Un grand nombre de ces indicateurs identifiés par l'OCDE appliquent une simple logique de comptage : étendue technologique du brevet (nombre de sous-classes de la Classification Internationale des Brevets – CIB – auxquelles il est alloué), taille de sa « famille » (nombre d'offices dans lesquels l'invention est déposée), nombre de citations brevets amont (qu'il effectue) et aval (dans lesquels il apparaît), nombre de références qu'il fait à la littérature hors brevets (i.e. publications scientifiques), etc. L'objectif de notre article est de proposer des lectures structurales de deux de ces indicateurs de la qualité technologique des brevets, sans doute les plus populaires en économie de l'innovation : l'originalité et la généralité ([TRA 97], [HAL 01]).

Ces deux indicateurs vont au-delà de la logique de comptage. Il s'agit d'indices de dispersion des citations technologiques amont et aval. Une citation technologique peut être définie simplement comme une référence au contenu technologique des brevets cités (originalité) ou citant le brevet évalué (généralité). Le contenu technologique correspond aux codes de la CIB alloués aux brevets. Imaginons un brevet p qui cite cinq brevets antérieurs dans son état de l'art dont le contenu technologique n'est

constitué que du code CIB A. Dans ce cas, la dispersion des citations technologiques amont est nulle, le brevet p n'est pas original. A l'inverse si les brevets cités font référence à des codes CIB tous différents. Pour quantifier cette dispersion, les auteurs utilisent des opposés d'un indice d'Herfindahl.

Leurs concepteurs le reconnaissent volontiers : ces indices sont biaisés, et quelques contributions plus récentes proposent de raisonner sur des Herfindahl normalisés. L'essentiel des discussions autour de l'originalité et de la généralité s'est concentré sur ces biais (cf. p.ex. [KIM 19]) alors qu'en pratique, peu utilisent ces indices transformés (cf. p.ex. [MAR 18]). Il ne s'agit pas dans notre article de faire le point sur cette discussion technique ; à notre avis, les réponses formelles ont déjà été apportées (cf. p.ex. [JOS 06]). Nous souhaitons bien discuter de la portée des indices originels, mais en utilisant une toute autre perspective : nous allons explorer la signification structurale de ces indicateurs. Nous allons pour ce faire les transformer en des objets mathématiques – des « graphes » – et nous insisterons sur des amendements et prolongements inspirés par cette nouvelle approche.

Nous n'appliquerons pas les nouveaux indices d'originalité et de généralité à des données de brevets réelles. Nous nous focaliserons sur la méthode et des exemples simples afin de bien faire comprendre la démarche de construction et d'interprétation des indices. De ce point de vue, nous proposerons ici deux lectures complémentaires. Dans la première, nous partirons des indices originels et nous leur appliquerons directement une représentation sous la forme de graphes (section 2). Dans la deuxième lecture, nous partirons de matrices de flux technologiques et nous reconstruirons les indices de dispersion tels qu'ils ont été conçus initialement. Nous irons même un peu plus loin : nous proposerons d'autres indices de la qualité technologique à partir de ces matrices, dont un qui nous semble de portée supérieure pour apprécier cette qualité (section 3).

2. Originalité et généralité des brevets sous la forme de graphes

Les indices d'originalité et de généralité d'un brevet correspondent donc à des opposés d'indices d'Herfindahl. [TRA 97] et [HAL 01] les définissent de la manière suivante. Si l'on note O_p la valeur de l'originalité du brevet p comprise entre 0 et 1, s_{pj} le pourcentage de références fait par le brevet p au code technologique j parmi les C codes contenus dans les brevets cités par p , alors :

$$O_p = 1 - \sum_{j=1}^C s_{pj}^2 \quad [1]$$

De manière symétrique, si l'on note G_p la généralité du brevet p comprise entre 0 et 1, s_{pk} le pourcentage de références au code technologique k fait par les brevets citant p parmi les N codes contenus dans ces brevets, alors :

$$G_p = 1 - \sum_{k=1}^N s_{pk}^2 \quad [2]$$

L'originalité fait appel à des informations contenues dans des brevets passés, que l'on peut dater à leur enregistrement en office, alors que la mesure de la généralité fait appel à des brevets enregistrés après p , ce qui pose des problèmes propres liés au délai pertinent à retenir pour apprécier sa valeur. Dans les deux cas, les informations utiles extraites des brevets concernent les technologies qui leur sont allouées. Notons T l'année de l'évaluation de la qualité. Les figures 1 et 2 sont des représentations sous forme de graphes des indices d'originalité et de généralité.

Dans ces deux figures, chaque brevet cité par p ou citant p est représenté par un cercle, et leur contenu technologique – leurs codes CIB – correspondent aux lettres A, B, ... F. Les relations de ces brevets à p sont dirigées vers p (originalité) ou ont pour origine p (généralité). Ce sont des flux allant des brevets cités vers les brevets citant qui sont ici reproduits : des connaissances technologiques anciennes qui servent à produire des inventions ultérieures.

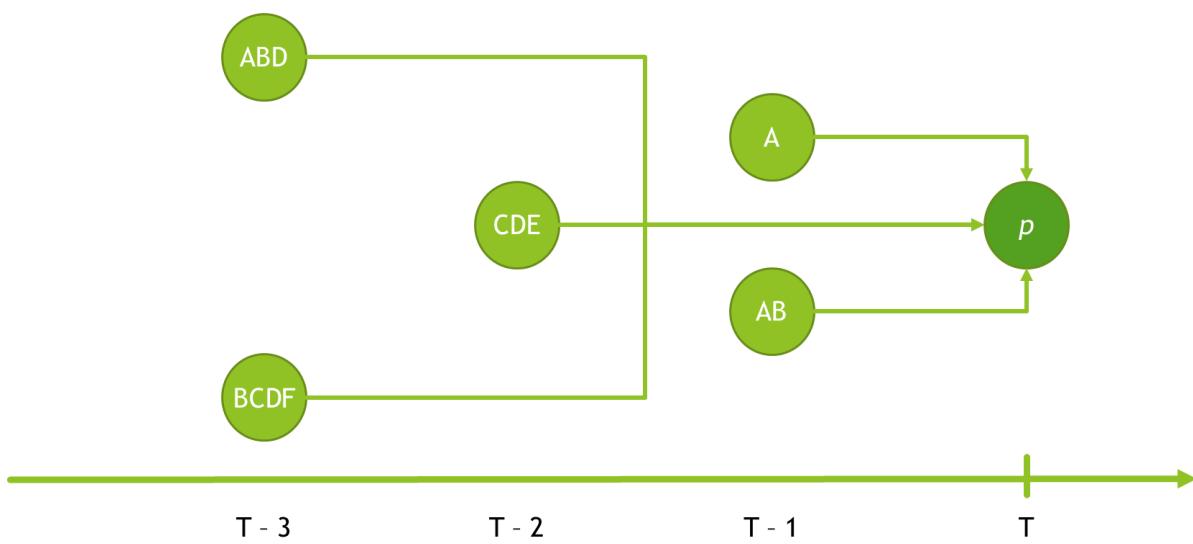


Figure 1. Exemple de graphe d'originalité

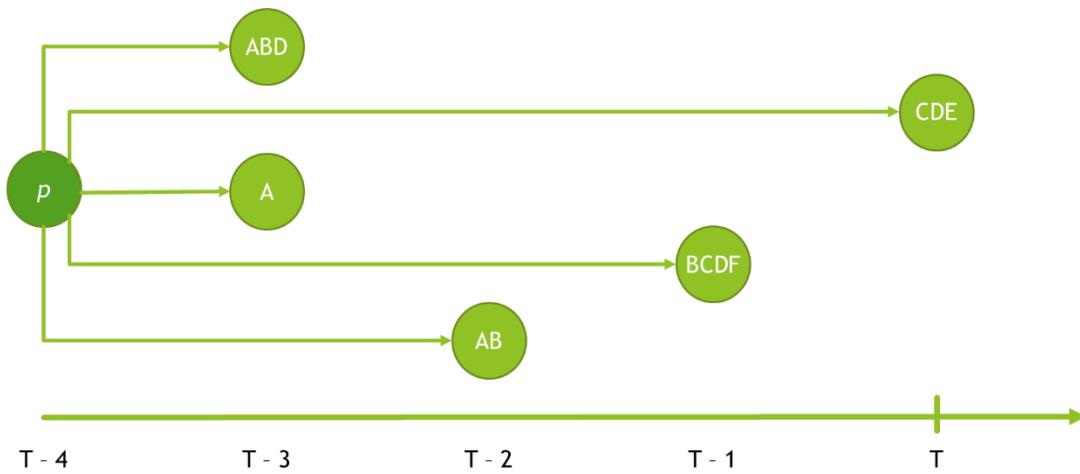


Figure 2. Exemple de graphe généralité

D'un point de vue formel, les raisonnements sur O_p et G_p sont symétriques. Dans la suite, nous utiliserons l'indice d'originalité pour illustrer nos arguments. En tout cas, tels que les exemples sont construits, la mesure des deux indices est aisée. La figure 3 représente la répartition des contenus technologiques des brevets de la figure 1 en lien avec p , et la valeur de O_p s'en déduit : 0,81 en arrondi, indiquant une originalité importante. En effet, avec le biais de la mesure noté par [TRA 97], avec six technologies liées à p , la valeur maximale possible pour O_p est de 0,84.

Nous sommes donc en présence de deux graphes de citations. L'exploitation de données relatives aux citations de brevets pour analyser la diffusion de connaissances a donné lieu à une littérature foisonnante à la suite des travaux fondateurs de [JAF 93]. Le raisonnement de base est double. Premièrement, la relation entre le brevet cité et le brevet citant peut dépeindre partiellement les flux de connaissances entre des technologies, des organisations ou des inventeurs. Deuxièmement, il est possible d'ajouter une dimension géographique à cette relation. La localisation géographique des inventeurs peut être par exemple utilisée pour cartographier la distribution et la diffusion spatiales des flux ou des retombées de connaissances ([JAF 93], [JAF 99], [JAF 00], [JAF 02], [PER 05], [RAV 13], [VER 99], [VER 04]). Dans cette tradition, les travaux empiriques utilisent généralement les informations géographiques et technologiques fournies dans les brevets. Les flux de citations peuvent

être cartographiés à des niveaux d'agrégation distincts, par exemple au niveau du secteur, du domaine technologique, de la région ou du pays.

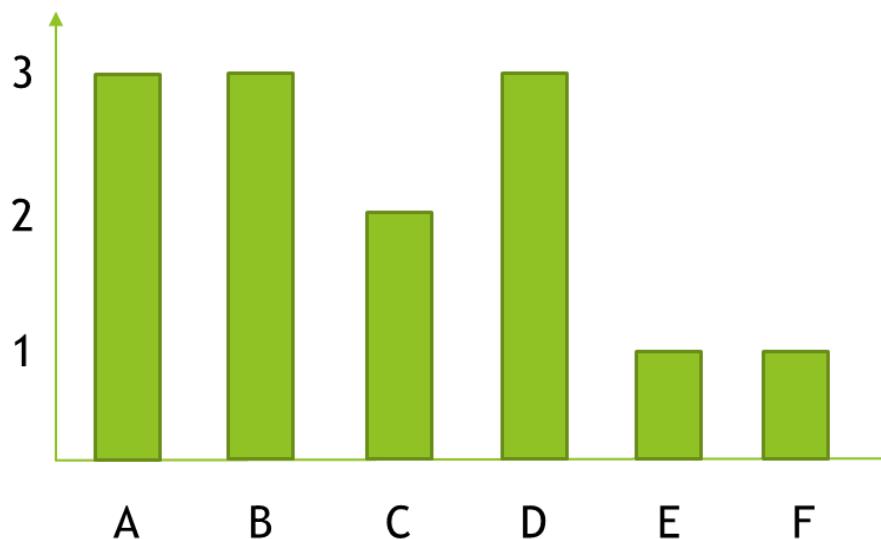


Figure 3. Répartition des contenus technologiques des citations du brevet p dans la figure 1

Si on approfondit l'analogie avec cette tradition de recherche, la manière dont les graphes des figures 1 et 2 sont construits pose une double question. Tout d'abord, la direction des flux ne permet pas de lier entre elles des technologies car on ne sait pas quelles sont celles qui sont allouées au brevet p , alors que cette information est disponible. Ensuite, une technologie particulière contribue de la même manière à la valeur de O_p qu'elle soit seule allouée à un brevet antérieur ou bien qu'elle soit combinée avec d'autres. Par exemple, dans la figure 1, en période $T - 1$, la technologie A contribue deux fois à la valeur de O_p , alors que pour un brevet cité elle ne fait sens pour p que lorsqu'elle est associée à la technologie B. Dans la littérature citée au paragraphe précédent, une manière de traiter ce dernier point consiste à accorder un poids différent aux technologies selon le nombre d'entre elles qui sont allouées dans les brevets. Cette technique de pondération de la contribution des technologies citées s'appelle « comptage fractionné » (*fractional counting*). Elle possède un avantage qui nous semble essentiel : si le brevet comprenant les technologies A et B en $T - 1$ dans la figure 1 attribue à chacune d'elles un poids de 0,5, cela signifie que cette citation particulière vaut 1, soit autant que celle qui, à la même date, ne comprend que la technologie A. Au total, à cette même date, la technologie A cumule un poids de 1.5, alors que la technologie B ne contribue à p que pour un 0.5. Le poids cumulé de A et B vaut donc bien le nombre de citations de p en $T - 1$.

En approfondissant la lecture en termes de graphe des citations de p , d'autres éléments peuvent être introduits. Transformons d'abord le graphe de la figure 1 en un graphe de relations technologiques, en posant que le brevet p combine les technologies A et B (figure 4).

Le premier élément additionnel a trait à la radicalité de l'invention p . [SHA 01] mesure la radicalité d'un brevet par le nombre de technologies CIB auxquelles les brevets cités sont alloués, mais dans lesquelles le brevet p lui-même n'est pas classé. La radicalité ici repose sur la capacité des inventeurs à faire reposer leur nouveauté sur des champs de connaissances différents de ceux auxquels elle s'exerce. La figure 5 plaque ce critère de radicalité aux flux reportés dans la figure 4.

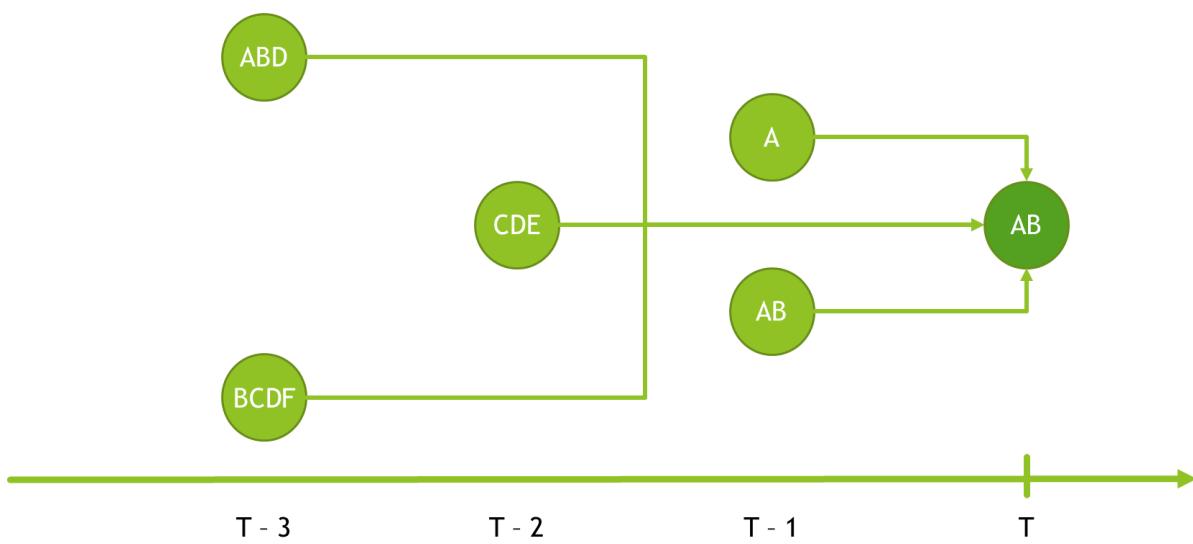


Figure 4. Graphe de citations entre technologies

En maintenant les règles de calcul de l'originalité présentées plus haut, nous proposons de nouveaux indices de la qualité technologique des brevets qui prennent en compte les liaisons technologiques, les pondérations et la radicalité. Ainsi :

$$\tilde{O}_p = 1 - \sum_{j=1}^{\tilde{C}} s_{pj}^2 \quad [3]$$

est le nouvel indice d'originalité où s_{pj} est le pourcentage des références *pondérées* fait par le brevet p au code technologique j parmi les \tilde{C} codes *originaux* contenus dans les brevets cités par p .

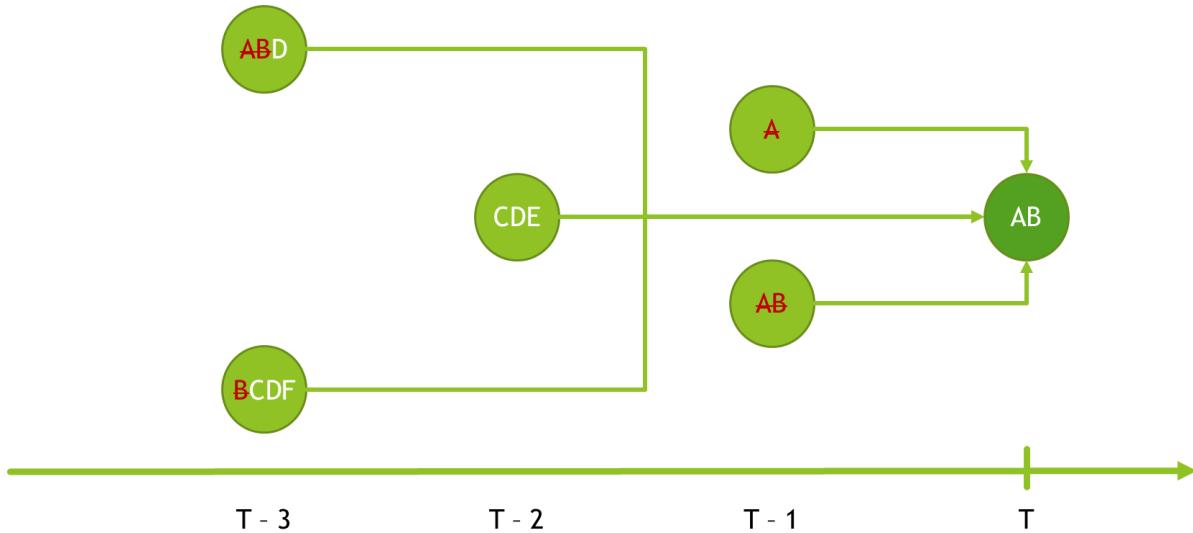


Figure 5. Radicalité du brevet p

De la même manière :

$$\tilde{G}_p = 1 - \sum_{k=1}^{\tilde{N}} s_{pk}^2 \quad [4]$$

est l'indice de généralité modifié où s_{pk} est le pourcentage de références *pondérées* au code technologique k fait par les brevets citant p parmi les \tilde{N} codes *originaux* contenus dans ces brevets.

Pour \tilde{O}_p nous obtenons une répartition des poids technologiques indiquée en figure 6, et une valeur d'environ 0,62 (valeur maximale de 0,75). La somme des poids des technologies est égale à 3, soit le nombre de citations prises en compte sur les cinq totales, deux étant supprimées car n'introduisant

absolument aucune radicalité (celles de la période $T - 1$). Nous nommons cette valeur « originalité relative », relative au contenu technologique de p .

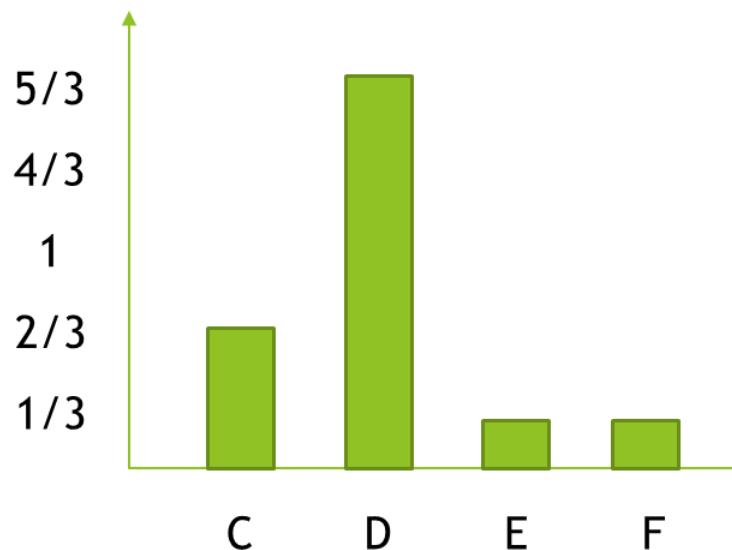


Figure 6. Répartition des contenus technologiques des citations du brevet p de la figure 5

Un deuxième élément additionnel relève d'une analyse classique d'un graphe : le choix d'un niveau de voisinage. Un voisinage direct, ou de niveau 1, est l'ensemble des « nœuds » / des brevets connectés à p . Dans les figures de graphes précédentes, seuls les voisinages directs de p sont dessinés. Un voisinage de niveau 2 prend également en compte les voisins directs des voisins directs de p . Et ainsi de suite pour un voisinage de niveau $r > 2$. Dans notre cas, cela revient à prendre en compte dans les mesures de la qualité technologique d'un brevet les technologies qui ont servi à produire les *inputs* de p . Pour illustrer l'intérêt de ce prolongement, partons d'un sous-graphe de la figure 5 et supposons que l'un des brevets ayant contribué à produire p effectue une citation (figure 7).

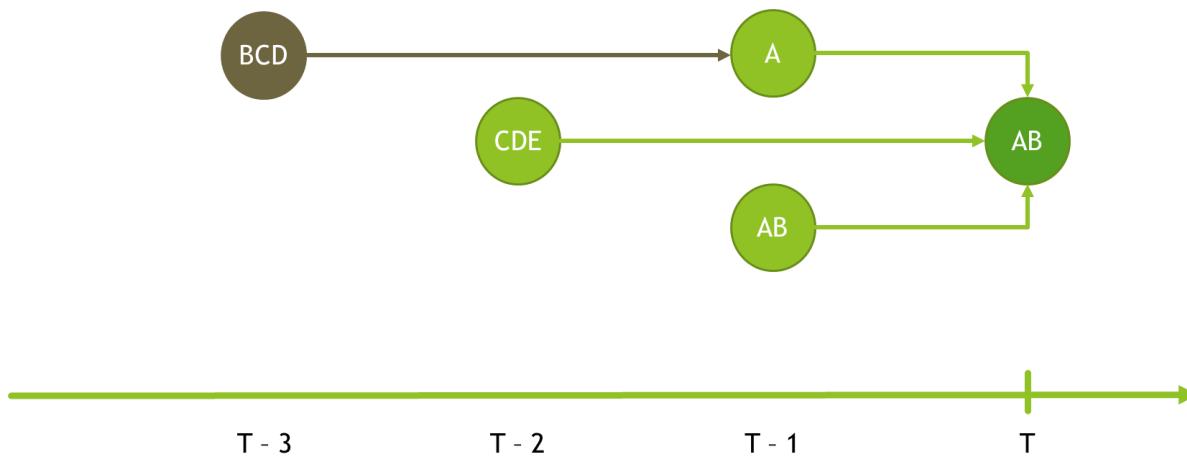


Figure 7. Originalité au voisinage de niveau 2

Un brevet composé des technologies B, C et D est dans l'état de l'art d'un brevet alloué à la technologie A (ce dernier est donc parfaitement radical) qui, lui-même, sert à produire p qui combine les technologies A et B. Si l'on prend en compte de manière paramétrique les pondérations, la radicalité et les rangs de voisinage, alors plusieurs périmètres de l'originalité peuvent être définis :

- l'originalité standard, qui intègre toutes les technologies des brevets de niveau 1 pour p , soit A, B, C, D et E ;

- l'originalité relative de niveau 1, qui ne comprend plus que les technologies C, D et E ;
- l'originalité relative de niveau 2, qui ne comprend plus que la technologie E. L'originalité est ici doublement relative : au contenu technologique de *p* et du brevet qu'il cite et qui comprend lui-même une citation.

Nous n'irons pas plus loin ici dans les prolongements. L'objectif n'est pas pour nous d'être exhaustif, mais simplement de montrer qu'une lecture en termes de graphes peut conduire à proposer des mesures aménagées, paramétrables, de la qualité technologique d'un brevet, en maintenant le cadre de mesure initial privilégié par [TRA 97] : un indice de dispersion intégrant, par opposition, un Herfindahl. Au total, nous aboutissons à deux conclusions fortes :

- la vision graphique des indices de qualité des brevets peut conduire à proposer des ajustements aux indices traditionnels ;
- le contenu technologique de *p* et son état de l'art à différents niveaux (nombre de citations, combinaisons technologiques, historicité / niveau, radicalité) comptent pour apprécier sa qualité « relative ».

3. De nouveaux indices de qualité à partir de matrices de flux technologiques

La figure 4 est la source graphique d'une matrice de flux technologiques (MFT) aux pondérations près. Une MFT exprime un processus de production d'inventions par l'agrégation de flux pondérés entre technologies à partir de données de citations. Les MFT sont utilisées pour identifier des similarités entre les processus de production technologique [LEB 17]. L'objectif de cette section est, par le biais d'une MFT, de définir la qualité technologique d'un brevet en fonction de la qualité des technologies qui le composent. Conséquence marquante de la procédure que nous présentons ici : comme l'originalité, la généralité peut s'appréhender en *backward*, c'est-à-dire à partir d'un graphe de citations type figure 4 et non de son équivalent figure 2 (*forward*). Autrement dit, originalité et généralité se mesurent à partir d'une même MFT qui contient l'ensemble de l'état de l'art à la date de l'évaluation. Le praticien souhaitant mesurer la qualité technologique d'un brevet par sa généralité n'aura pas besoin d'avoir plusieurs années de recul pour pouvoir l'estimer.

Une MFT est une matrice carrée, non symétrique et réflexive (qui autorise les relations d'un élément avec lui-même) qui reporte les pondérations allant des technologies citées (en ligne) aux technologies citant (en colonne). Dans la figure 4, il n'y a que deux technologies citant, A et B, et six technologies citées, de A à F. Le comptage fractionné est utilisé pour définir le poids des relations entre technologies, de sorte que chaque citation vaut pour 1. La somme des poids d'une MFT est donc égale au nombre de citations entre brevets. Si le brevet citant est alloué, comme dans cet exemple, à deux technologies, les relations qu'elles entretiennent avec les technologies citées sont partagées de manière égale.

Prenons un exemple de mesure de pondération, pour la relation réflexive de A. Cette technologie est présente dans trois citations, soit seule, soit associée à une autre, ou encore à deux autres technologies. Le poids de A dans les citations de *p* correspond donc à la somme de 1, 1/2 et 1/3. Comme ce flux de A se répartit de manière égale vers toutes les composantes de *p*, il faut diviser cette somme par deux. Ce qui donne une pondération totale de 0,917 pour le poids de A vers A.

Contrairement à la représentation de la figure 4, une MFT relie des technologies entre elles, et non des brevets. Elle est donc plus riche, en ce sens qu'elle décompose les brevets en leurs composantes premières, et moins riche en même temps car elle dilue les sources « brevets » de ces relations. Au final, cependant, la MFT s'associe à un graphe spécifique de relations inter-technologiques où les nœuds sont les technologies en question, où les liens sont dirigés des technologies citées vers les

technologies citant, et où ces mêmes liens sont pondérés par la valeur des éléments correspondants de la MFT.

	A	B	C	D	E	F
A	0,917	0,917	0	0	0	0
B	0,542	0,542	0	0	0	0
C	0,292	0,292	0	0	0	0
D	0,458	0,458	0	0	0	0
E	0,167	0,167	0	0	0	0
F	0,125	0,125	0	0	0	0

Tableau 1. MFT sourcée par les citations de la figure 4

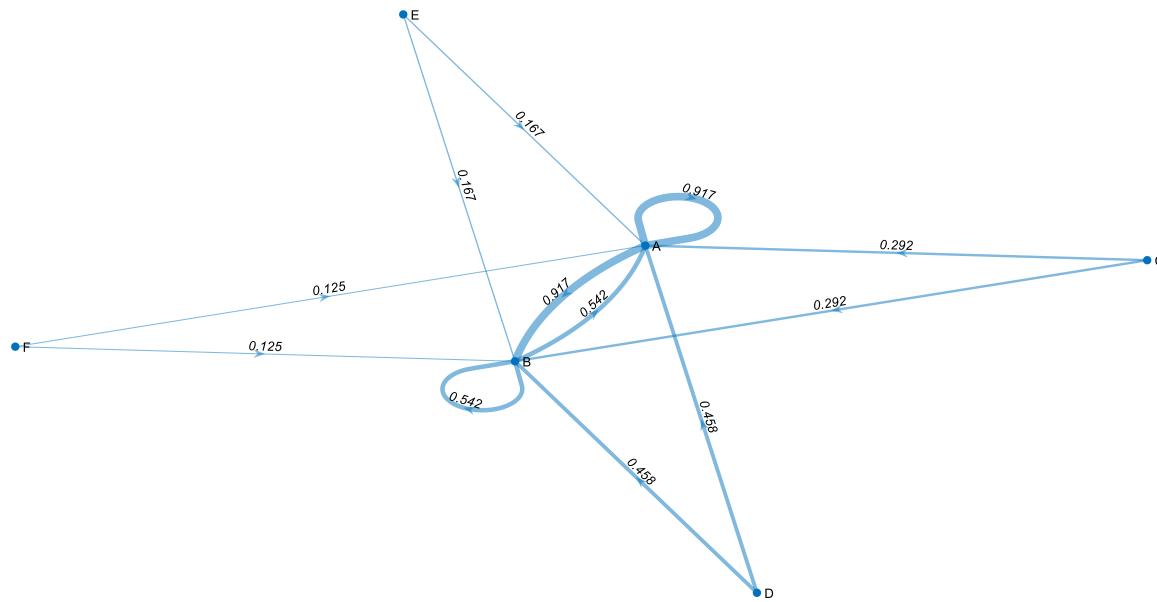


Figure 8. Graphe associé à la MFT du tableau 1

Estimer la qualité d'une technologie à partir d'une MFT revient à apprécier la position de cette technologie dans le graphe qui lui est associé. Il y a deux manières d'envisager cette position :

- en termes de concentration / dispersion, comme dans la section 2 de l'article ;
- en termes de centralité, comme l'envisage l'analyse des réseaux sociaux (ARS) [WAS 94].

Commençons par l'approche en termes de concentration / dispersion (figure 9). Dans une MFT, une lecture en ligne indique les usages d'une technologie *t* par elle-même (autocitations) et par les autres comme *input* de l'invention. La manière dont sont conçues ces matrices est dite *backward* : la date de référence est le brevet citant auquel on associe tous les cités indépendamment de leurs dates. Pour la construire de manière *forward*, il faudrait fixer la date des brevets cités et indiquer en colonne les citant indépendamment de leurs dates. La lecture en ligne d'une matrice *backward* est une adaptation pratique de la généralité : on se pose la question « comment les inventions d'hier se sont-elles diffusées à la date d'aujourd'hui ? » au lieu de « comment les inventions d'une date passée se sont-elles

diffusées jusqu'à aujourd'hui ? ». La logique est la même, mais la date de référence change, allant des cités vers les citant. On utilise la lecture en colonne pour mesurer « l'originalité » d'une technologie t . L'originalité de t est d'autant plus importante qu'elle mobilise pour être produite un nombre élevé de technologies de manière équilibrée. Dans ces mesures, nous supprimons conventionnellement les autocitations.

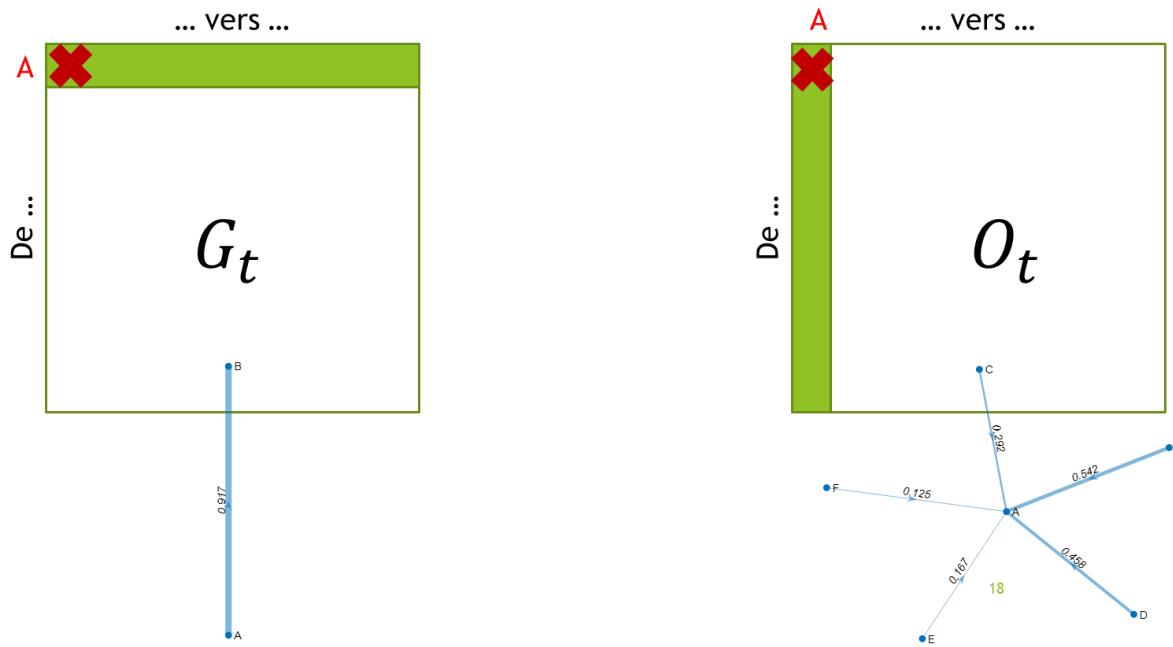


Figure 9. Architecture de mesure de l'originalité et la généralité par une MFT pour la technologie A

Au final, la qualité d'une technologie t peut s'apprécier par deux indices.

$$O_t = 1 - \sum_{j \neq t} s_{tj}^2 \quad [5]$$

où $s_{tj} = M_{tj} / \sum_{l \neq t} M_{tl}$, avec M la MFT. O_t est une mesure de la complexité du processus de production de la technologie t .

$$G_t = 1 - \sum_{k \neq t}^N s_{kt}^2 \quad [6]$$

où $s_{kt} = M_{kt} / \sum_{l \neq t} M_{lt}$. G_t est une mesure de la diversité des usages de la technologie t .

Appliquons ces mesures au graphe de la figure 8, pour obtenir les résultats du tableau 2. A apparaît légèrement plus originale que B, alors que les quatre autres technologies apparaissent aussi générales les unes que les autres.

Intéressons-nous maintenant à la *centralité* des technologies. Dans le graphe de la figure 8, une technologie peut être centrale de multiples manières. Les deux notions de centralité – d'importance relative – les plus utilisées dans la littérature ARS sont le *degré* et la *force*. Le degré correspond au nombre de connexions d'un nœud. La force correspond à la somme des intensités de connexions d'un nœud. Comme nous avons affaire à un graphe dirigé, les notions de degré et de force peuvent se distinguer selon qu'il s'agit de flux entrants ou de flux sortants du nœud. On parle alors de in/out degré et de in/out force. Nous supprimons encore les autocitations dans le cas de ces mesures de centralité. Une technologie est plus ou moins centrale selon les relations qu'elle entretient avec les autres technologies, et non avec elle-même.

	Originalité	Généralité
A	0,748	0,000
B	0,693	0,000
C	0,000	0,500
D	0,000	0,500
E	0,000	0,500
F	0,000	0,500

Tableau 2. Originalité et généralité des technologies

Dans le même temps, rappelons-nous que dans une MFT l'originalité est affaire de flux entrants, alors que la généralité est affaire de flux sortants. De sorte que l'on peut croiser deux dimensions pour produire une typologie : degré *vs* force, originalité *vs* généralité (figure 10).

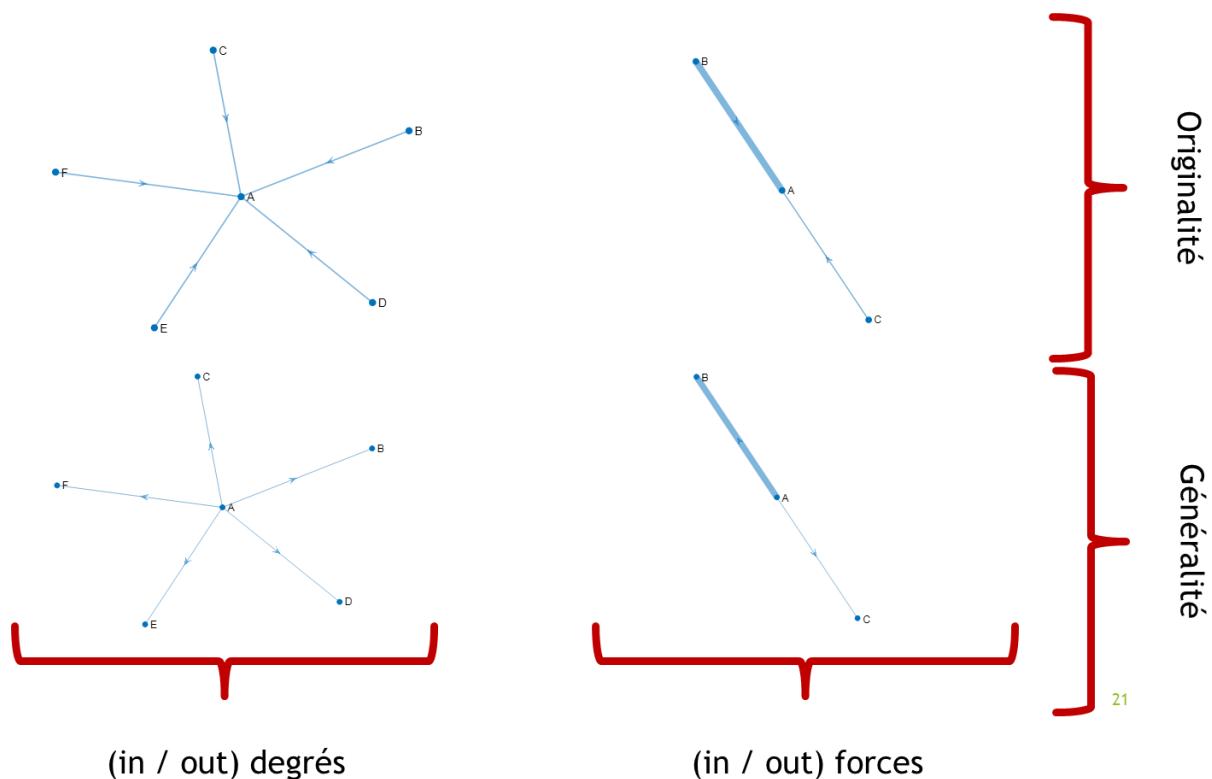


Figure 10. Typologie des relations entre technologies

On peut donc apprécier l'originalité et la généralité d'un nœud par sa position relative soit en termes de degré, soit en termes de force. [OPS 10] proposent un indice composite qui combine degré et force, en distinguant les flux entrants des flux sortants.

$$C_{t, in/out}^{[\alpha \geq 0]} = x_{t, in/out}^{\alpha} y_{t, in/out}^{1-\alpha} \quad [7]$$

avec x_t le degré de la technologie t , et y_t sa force.

[OPS 10] jouent sur le paramètre α pour donner plus ou moins d'importance au degré ($\alpha \gg 0$) ou bien à la force (α proche de 0). Ils autorisent même des valeurs de α strictement supérieures à 1 ce qui, à notre sens, n'apporte pas d'information utile voire conduit à des contre-intuitions, et ce qui devrait autoriser, pour assurer une symétrie, des valeurs strictement négatives de ce paramètre.

Lorsque l'on applique cette formule pour différentes valeurs de α à partir du sous-graphe de la figure 11 (qui correspond au graphe complet de la figure 8 excluant les autocitations), nous obtenons les résultats des tableaux 3 et 4, selon la direction des liens.

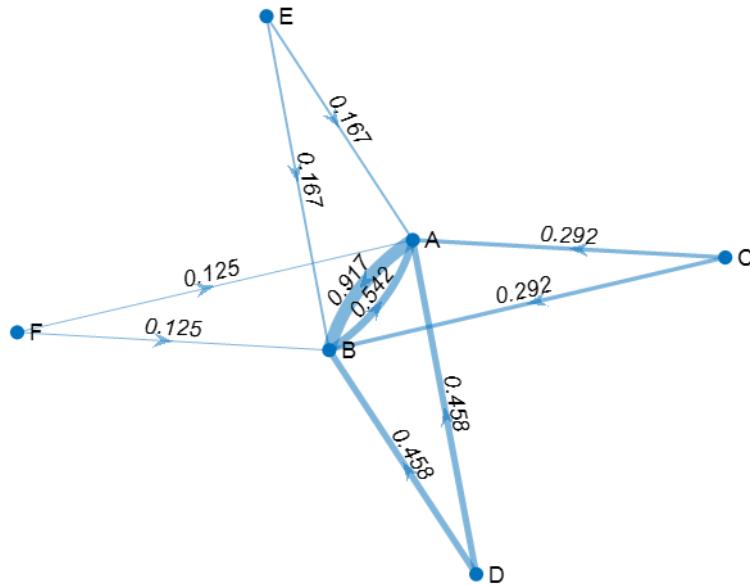


Figure 11. Sous-graphe du graphe de la figure 8 (suppression des autocitations)

α	0,00	0,25	0,50	0,75	1,00	1,25	1,50	1,75	2,00
A	1,58	2,11	2,81	3,75	5,00	6,66	8,88	11,84	15,78
B	1,96	2,48	3,13	3,96	5,00	6,32	7,99	10,10	12,76
C	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
D	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
E	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
F	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Tableau 3. Originalité pour différentes valeurs de α

α	0,00	0,25	0,50	0,75	1,00	1,25	1,50	1,75	2,00
A	0,92	0,94	0,96	0,98	1,00	1,02	1,04	1,07	1,09
B	0,54	0,63	0,74	0,86	1,00	1,17	1,36	1,58	1,85
C	0,58	0,79	1,08	1,47	2,00	2,72	3,70	5,03	6,85
D	0,92	1,11	1,35	1,65	2,00	2,43	2,96	3,59	4,37
E	0,33	0,52	0,82	1,28	2,00	3,13	4,89	7,66	11,98
F	0,25	0,42	0,71	1,19	2,00	3,36	5,66	9,51	16,00

Tableau 4. Généralité pour différentes valeurs de α

Prenons une valeur équilibrée de α (0,5) : la technologie B est la plus centrale en matière de complexité du processus de production, alors que la technologie D l'est pour la diversité de ses usages.

Un inconvénient majeur de cette méthode est que α reste un paramètre, et [OPS 10] ne donnent pas de piste pour son endogénéisation, le point de vue de l'expert conduisant à arbitrer sur un intervalle pertinent de valeurs. Nous suggérons ici un critère simple pour attribuer une valeur à ce paramètre : la densité du graphe de la figure 11. Imaginons un graphe où toutes les connexions possibles entre les technologies sont exploitées. Dans ce cas, l'information sur les degrés est de peu d'importance, et ce qui distingue les nœuds entre eux est le différentiel d'intensités de ces connexions, i.e. leurs forces. A l'inverse, si le graphe comprend peu de connexions, les différentiels de degrés peuvent être prépondérants pour estimer la position d'un nœud dans le graphe complet. Si l'on nomme d la densité du graphe, on obtient :

$$d = K/(Q^2 - Q) \quad [8]$$

avec Q le nombre de nœuds dans le graphe, K le nombre de connexions dans ce graphe, et $0 \leq d \leq 1$.

Les indices de qualité technologique s'expriment alors comme suit :

$$C_{t,in/out} = x_{t,in/out}^{1-d} y_{t,in/out}^d \quad [9]$$

Dans le graphe de la figure 11, la valeur de d est de 1/3, ce qui donne les scores de centralité du tableau 5 ($C_{t,in}$ pour originalité, $C_{t,out}$ pour généralité).

	Originalité	Généralité
A	3,41	0,97
B	3,66	0,82
C	0,00	1,33
D	0,00	1,54
E	0,00	1,10
F	0,00	1,00

Tableau 5. Originalité et généralité des technologies avec le paramètre α endogénéisé

Ces scores sont obtenus par technologie, sans référence aux brevets dont elles sont les composantes. Passer d'une « qualité d'une technologie » à une « qualité technologique d'un brevet » nécessite de procéder à un choix d'agrégation des scores des technologies qui composent ce brevet. Dans notre exemple initial, il s'agit d'agréger les scores des technologies A et B pour p en originalité et en généralité. En retenant une moyenne arithmétique, l'originalité de p se fixe à 3,53 et sa généralité à 0,89. Pour p , enregistré aujourd'hui, avec ces scores obtenus à partir d'une MFT *backward*, l'indice de généralité reflète une estimation du succès *futur* du brevet en fonction des usages *actuels* des technologies A et B dans les processus d'invention. C'est une estimation, ce n'est pas un *potentiel* de diffusion car la réalité du succès de p (le nombre de ses usages à plusieurs années) peut largement dépasser ce score. Imaginons cependant un brevet édité aujourd'hui, possédant un fort score de généralité de par les technologies qui le composent et qui, cependant, ne sera pas cité dans le futur. Peut-être que cela peut révéler une forme de sous-exploitation de ce brevet, et qu'il peut être intéressant de le mobiliser dès aujourd'hui pour produire de l'invention future.

4. Conclusion

Il existe une multitude d'indices pour estimer la qualité technologique des brevets [OCD 13]. En dépit de leurs biais, les indices d'originalité et de généralité ([TRA 97], [HAL 01]) sont aujourd'hui largement utilisés, soit en reprenant les formulations de leurs concepteurs, soit en les normalisant. Quoi qu'il en soit, nous voulions montrer dans cet article qu'une interprétation structurale de ces deux indicateurs conduisait à introduire de nouveaux paramètres (radicalité, niveau de voisinage, pondération des technologies) pour mieux en préciser la portée. Et parce que cela revient à manipuler des briques de matrices de flux technologiques, nous avons envisager dans un deuxième temps l'utilisation de ces matrices pour produire de nouveaux indices de la qualité technologique des brevets, indices reposant sur la qualité des technologies elles-mêmes. Nous proposons, en fin d'article, un indice composite de la qualité des technologies reposant sur la notion de centralité en ARS et dont le paramètre clé est endogénéisé selon une propriété de la structure globale des connexions inter-technologiques : la densité de cette structure.

Les pistes ouvertes par cette recherche sont nombreuses. Il s'agit dorénavant d'appliquer ces indices à des bases de données de brevets et à procéder, à la manière de [KIM 19], à un examen systématique des propriétés analytiques de ces nouveaux indices.

Bibliographie

- [HAL 01] HALL B.H., JAFFE A.B., TRAJTENBERG M., « The NBER patent citations data file: lessons insights and methodological tools », *NBER Working Paper*, n° 8498, 2001.
- [JAF 93] JAFFE A.B., TRAJTENBERG M., HENDERSON R., « Geographic localization of knowledge spillovers as evidenced by patent citations », *Quarterly Journal of Economics*, vol. 108, p. 577-598, 1993.
- [JAF 99] JAFFE A.B., TRAJTENBERG M., « International knowledge flows: evidence from patent citations », *Economics of Innovation and New Technology*, vol. 8, p. 105-136, 1999.
- [JAF 00] JAFFE A.B., TRAJTENBERG M., FOGARTY M., « Knowledge spillovers and patent citations: evidence from a survey of inventors », *American Economic Review*, vol. 90(2), p. 215-218, 2000.
- [JAF 02] JAFFE A.B., TRAJTENBERG M., *Patents citations and innovations: a window on the knowledge economy*, MIT Press, Cambridge (Mass.), 2002.
- [JOS 06] JOST L., « Entropy and diversity », *Oikos* 10.1111/j.2006.0030-1299.14714.x, 2006.
- [KIM 19] KIM K., HWANG J., JUNG S., KIM E., « Which technology diversification index should be selected? Insights for diversification perspectives », *Cogent Business & Management*, vol. 6, 1643519, 2019.
- [LEB 17] LEBERT D., EL YOUNSI H., *International specialization dynamics* Wiley & ISTE Press, Londres, 2017.
- [MAR 18] MARKU E., « Measuring innovation quality: a patent analysis », *IOSR Journal of Business and Management*, vol. 20(8), p. 51-58, 2018.
- [OCD 13] OCDE, « Measuring patent quality: indicators of technological and economic value », *Science Technology and Industry Working Papers*, n° 2013/03, 2013.
- [OPS 10] OPSAHL T., AGNEESSENS F., SKVORETZ J., « Node centrality in weighted networks: generalizing degree and shortest paths », *Social Networks*, vol. 32, p. 245-251, 2010.
- [PER 05] PERI G., « Determinants of knowledge flows and their effect on innovation », *Review of Economics and Statistics*, vol. 87, p. 308-322, 2005.
- [RAV 13] RAVE T., GOETZKE F., « Climate-friendly technologies in the mobile air-conditioning sector: a patent citation analysis », *Environmental Economics and Policy Studies*, vol. 15, p. 389-422, 2013.
- [SHA 13] SHANE S. (2001), « Technological opportunities and new firm creation », *Management Science*, vol. 47, p. 205-220, 2001.
- [TRA 97] TRAJTENBERG M., JAFFE A.B., HENDERSON R., « University versus corporate patents: a window on the basicness of invention », *Economics of Innovation and New Technology*, vol. 5, p. 19-50, 1997.

- [VER 99] VERSPAGEN B., DE LOO I., « Technology spillovers between sectors and over time », *Technological Forecasting and Social Change*, vol. 60, p. 215-235, 1999.
- [VER 04] VERSPAGEN B., SCHOENMAKERS W., « The spatial dimension of patenting by multinational firms in Europe », *Journal of Economic Geography*, vol. 4, p. 23-42, 2004.
- [WAS 94] WASSERMAN S., FAUST K., *Social network analysis: Methods and applications*, Cambridge University Press, Cambridge (Mass.), 1994.