

# Aplicando rede de IPCs para identificar a formação de clusters de tecnologias para produção de bioetanol

Carolina da Silveira Bueno \*

José Maria Ferreira Jardim da Silveira \*

Antonio Marcio Buainain \*

*\*Núcleo de Economia Agrícola, Instituto de Economia – Universidade Estadual de Campinas*

## ABSTRACT

This paper uses the network approach in order to identify the presence of clusters of technologies in biofuels. The initial methodological step identifies the variables obtained from the patent, through a data mining treatment. Next is formed IPCs network (International Patents Classification). The 263 patents sample information has shown that the 1st generation bioethanol technology classes are directed intra-clusters in the exploratory stage, for 2nd generation bioethanol technology. The investigation has concluded that the more than 500 identified technologies, specialized clusters are IPCs: C12P7 / 06 C12P7 / 08 and C12P / 10. The technologies C12P7 / 06 "... · Ethanol, ie not to drink" are inter-cluster, i.e. directed to the formation of intra-cluster C12P7 / 08 "... produced as by-product or from substrate waste or cellulosic material "and C12P / 10" ... substrate containing cellulosic material."

**Keywords:** innovation networks; bioethanol; technologies; IPCs, second-generation.

## RESUMO

Este artigo usa a abordagem de redes com o objetivo de identificar a presença de clusters de tecnologias em biocombustíveis. A primeira etapa metodológica identifica as variáveis obtidas a partir de patentes, através de um tratamento de *data mining*. Em seguida é formada a rede de IPCs (*International Patents Classification*). A informação de 263 patentes da amostra permitiu verificar que as classes de tecnologias de bioetanol de 1ª geração estão direcionando intra-clusters, em fase exploratória, para tecnologias de bioetanol de 2ª geração. A investigação permitiu concluir que das mais de 500 tecnologias identificadas, os clusters especializados são de IPCs: C12P7/06, C12P7/08 e C12P/10. As tecnologias C12P7/06 "...·Etanol, i.e. não para bebida" são o inter-cluster, isto é, direcionou a formação dos intra-clusters de C12P7/08 "... produzido como subproduto ou a partir de substrato de resíduo ou de material celulósico" e C12P/10 "... substrato contendo material celulósico".

**Palavras-chave:** redes de inovação; bioetanol; tecnologias; IPCs, segunda-geração.

**Área ANPEC:** Área 9 - Economia Industrial e da Tecnologia

**Classificação JEL:** O33; Q16

## 1. Introdução

Os biocombustíveis têm sido considerados promissores para a substituição dos combustíveis fósseis com impactos decisivos no controle do clima (HLPE, 2013). Diferentemente do petróleo, os biocombustíveis são produzidos a partir de fontes renováveis de energia, como por exemplo, as matérias-primas agrícolas, entre elas, a cana-de-açúcar e o milho. Embora muitos tipos de materiais podem ser usados para produzir biocombustíveis, a cana-de-açúcar e o milho, respectivamente, Brasil e Estados Unidos, foram os únicos países que conseguiram viabilizar em larga escala com o uso desses materiais, abastecendo além do mercado interno, também o mercado externo (Sims et al., 2008). A capacidade produtiva combinada com a capacidade de especialização tornam esses países entre os poucos com aptidões para transferir conhecimentos nessa área de fronteira tecnológica, uma vez que foram os dois países que construíram o paradigma desde meados de 1970. Atualmente, diversos países vêm acometido esforços para produzir biocombustíveis, pois que o paradigma está em sua infância, e ainda não está definida uma trajetória tecnológica (Dal Poz et al, 2015). Esse fato decorre do crescente debate em relação aos níveis de emissão dos gases de efeito estufa (*greenhouse gases* – *GHG*) combinada com os acordos internacionais dos países para reduzir a emissão de CO<sup>2</sup> na atmosfera (HLPE, 2013).

A ideia de trajetórias tecnológicas (TT) desenvolvidas por Dosi (1982) contribuiu para esclarecer como um grupo específico de tecnologias evolui ao longo de um caminho, combinando novas ideias originais com processos de aprendizagem, uma combinação de conhecimento tácito e resultados da seleção do mercado. No entanto, existem muitas situações relevantes, onde é difícil identificar trajetórias tecnológicas, com muitas alternativas que competem por um lugar nos mercados futuros, o que significa que a concorrência tinha começado muitos anos antes do primeiro produto ou processo ter chegado no mercado. Este é o caso da biotecnologia (Orsenigo, 1989) e, certamente, é a situação encontrada em bioenergia como mostra os trabalhos de Dal Poz; Silveira; Masago (2013).

Entretanto, o problema é que a quantidade de pesquisas vem aumentando tão rapidamente que entender exatamente como as fronteiras do conhecimento vêm progredindo é um desafio. Uma solução encontrada foi estudar os seguintes indicadores: a) informações de patentes são utilizadas para identificar as classes tecnológicas de especialização e, b) a aplicação de redes para estudar as propriedades desse conhecimento. As informações contidas nas patentes podem ser utilizadas como importantes insumos na investigação de fenômenos relacionados à economia da inovação. A divulgação de informações técnicas relacionadas ao conhecimento novo e o caráter cumulativo das patentes permite que esse tipo de investigação seja realizado (Krafft et al., 2009; Epicoco, 2013; Zilberman, 2013; Souza, 2013). Com isso espera-se identificar a presença de especialização em biocombustíveis em relação à existência de futuros clusters tecnológicos em fase de experimentação, analisado a partir das variáveis obtidas a partir de patentes, em correspondência com as dimensões relacionadas de ligações inter-clusters e intra-clusters.

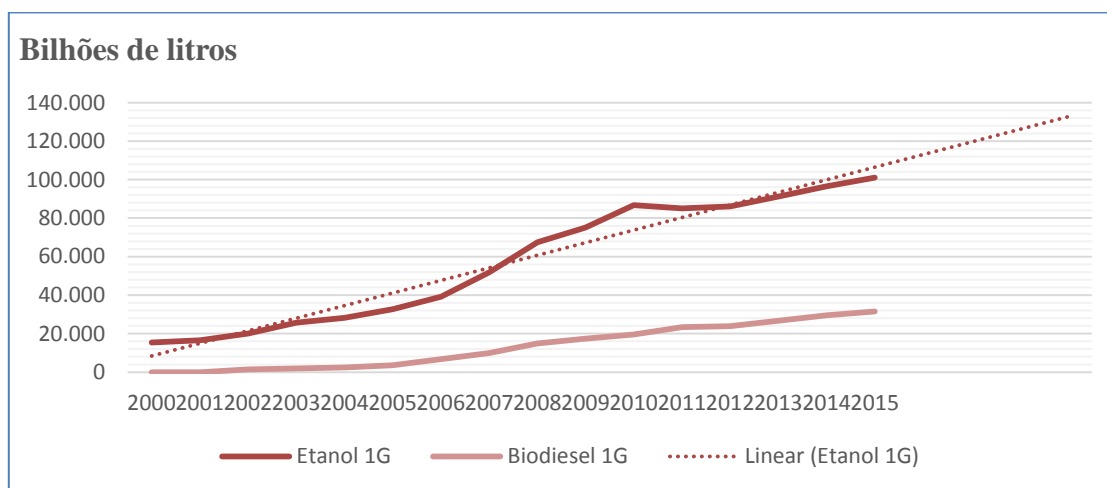
Na próxima seção, o trabalho apresenta os principais conceitos metodológicos e a aplicação da rede e os resultados, e os comentários finais na seção 3.

## 2. Cenários de biocombustíveis e a escolha para bioetanol de cana-de-açúcar

### 2.1 Caracterização dos biocombustíveis

Desde meados da década de 1970, com as duas crises internacionais do petróleo, que diversos países têm induzido esforços para desenvolver, testar e viabilizar a produção de combustíveis alternativos aos derivados do petróleo. Nesse período, o propósito em promover a substituição dos combustíveis fósseis estava no interesse da maioria dos países em diminuir a dependência dos países exportadores de petróleo. Atualmente, o principal fator de fortalecimento do paradigma dos biocombustíveis, além da incerteza do preço do petróleo no mercado internacional (HLPE, 2013), têm sido o papel dos acordos internacionais entre os governos no que tange à preocupação com os níveis de emissão dos gases de efeito estufa (*greenhouse gases* – GHG), combinadas ao imenso impacto (simbólico inclusive) que a percepção do problema do aquecimento global acarretou sobre as visões sobre o futuro da humanidade.

De maneira geral, o modelo de viabilização dos biocombustíveis deve-se ao papel das políticas públicas. Decisões dos produtores e a demanda do consumidor são determinadas por um conjunto de políticas e incentivos realizados em diversos países e o resultado é uma produção crescente de biocombustíveis em nível global. O volume de produção do bioetanol 1G, em 2011, representou 78% (85 bilhões de litros) do total produzido de biocombustíveis (pouco mais de 108 bilhões de litros), os outros 22% resultaram da produção de biodiesel. Os Gráficos 1 e 2 apresentam a produção mundial de biocombustíveis com tendência linear para o etanol 1G, e a emissão mundial de CO<sub>2</sub> per capita dos últimos anos.



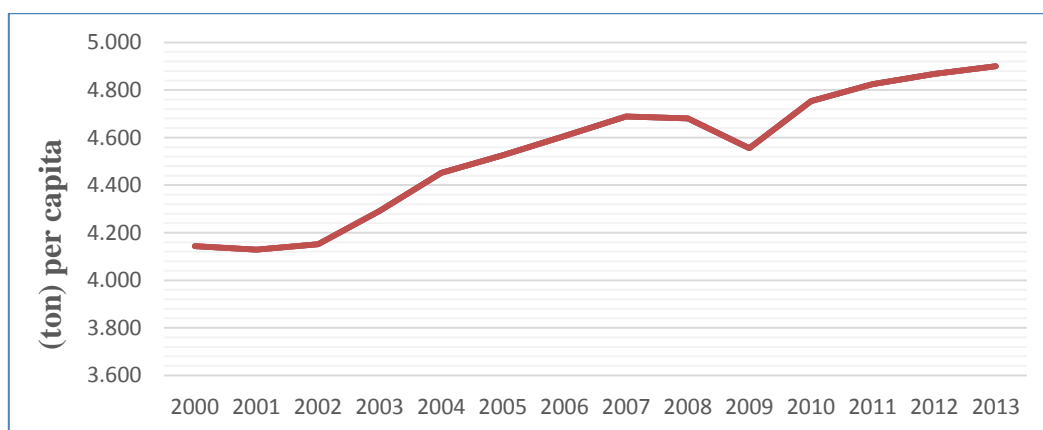
**Gráfico 1. produção mundial de biocombustíveis (2000-2015)**

Fonte: Elaboração própria, a partir dos dados de IEA (2013)

\*2012-2013-2014 e 2015 são projeções.

Da produção global de biocombustíveis, no ano de 2008, cerca de 15% foi utilizado para produzir bioetanol 1G a partir do milho (proporção maior pelos EUA), 18% foram para produzir etanol 1G da cana-de-açúcar (principalmente pelo Brasil), do óleo vegetal mundial produzido (principalmente pela União Europeia), cerca de 10% foram para produção de biodiesel. A elevação da produtividade, como demonstra o Gráfico 1, teve duas implicações importantes.

Concernente à primeira implicação, os biocombustíveis assumiram diferentes perfis para cada país ou região, devido a diversidade de recursos naturais, que por sua vez, deu origem a diferentes medidas de políticas. Por exemplo, na América Latina, o modelo de políticas para os biocombustíveis foi inspirado no programa brasileiro do Proálcool (PróÁlcool). Como característica principal, os países latino-americanos, que são essencialmente urbanos, adotaram políticas específicas para a área de transportes, a maioria com metas para etanol e biodiesel, principalmente, utilizando-se de matérias-primas de origem local, e em menor medida, estímulos para novas culturas.



**Gráfico 2. emissão mundial de CO2 per capita (2000-2013)**

Fonte: Elaboração própria, a partir dos dados de IEA (2013)

Concernente à segunda implicação, a exportação também tem sido igualmente sujeita a estímulos de política. Tais estímulos, e vale mencionar o crescimento da emissão mundial de CO2 como apresenta o Gráfico 2, provocaram a criação de um mercado de combustível cada vez mais globalizado, gerando um fluxo líquido de matérias-primas, etanol e biodiesel. Com os fluxos do mercado internacional e o crescimento da demanda, os países fornecedores de matérias-primas ou fornecedores de biocombustível líquido, tem implementado diversas medidas para aumentar suas exportações.

O fluxo do mercado global e a discussão ambiental estimularam a pesquisa e o desenvolvimento dos biocombustíveis em diversos países, por exemplo, o desenvolvimento de novas plantas ou processos de melhoramento envolvendo áreas como a genética e a biotecnologia. Em detrimento dos avanços das pesquisas, atualmente, os biocombustíveis são denominados em três tipos de geração.

No caso da 1ª geração de biocombustíveis, geralmente se referem ao etanol de culturas convencionais que são ricos em açúcar (como é o caso da beterraba, da cana-de-açúcar, do sorgo) e rico em amido (como milho, trigo, mandioca), além dos ricos em oleaginosas (como soja, girassol, palma), neste exemplo, se encaixa o etanol da cana-de-açúcar e o etanol do milho. Portanto, biocombustíveis de matérias-primas que também podem ser utilizadas para consumo humano e animal. O processo é realizado pela via de fermentação, com uso de leveduras específicas no processo produtivo, como a ‘*Saccharomyces cerevisiae*’.

Biocombustíveis de 2ª geração são produzidos a partir da biomassa lignocelulósica, e tem como base os subprodutos agrícolas (palha e bagaço), resíduos florestais (cascos, restos de talhas) e resíduos de biorrefinarias e lixo doméstico. Podem ser produzidas via duas rotas tecnológicas. As duas rotas para obter etanol de 2ª geração, são a via bioquímica, que inclui a hidrólise dos materiais lignocelulósicos e fermentação dos açúcares obtidos para conversão a etanol, e a rota termoquímica, que envolve o aquecimento da biomassa em altas temperaturas e pressões, em geral mais elevada que o processo bioquímico para obter biomassa líquida (BTL fuels). Rotas termoquímicas são mais flexíveis em termos de matérias-primas utilizadas, bem como na diversidade de combustíveis que podem ser obtidos. O combustível de 3ª geração normalmente se refere aos biocombustíveis que não competem com culturas alimentares nem com terras. Biocombustíveis à base de algas se enquadram nesta categoria (HLPE, 2013).

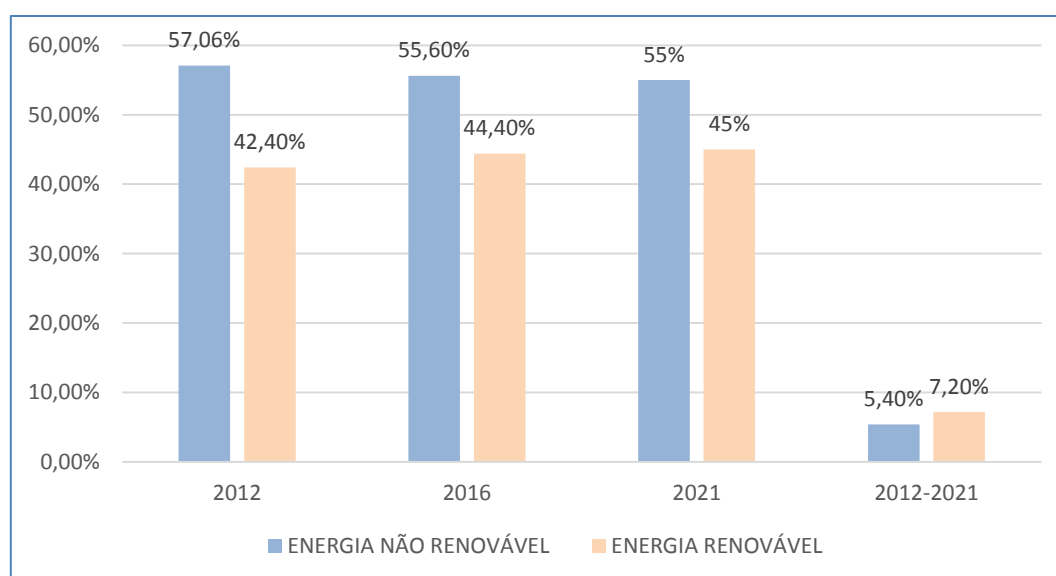
## 2.2 A importância do bioetanol de cana-de-açúcar

Dentre os biocombustíveis de 1ª geração, o bioetanol é o que apresenta maior volume de produção, sendo 80% dele derivado do milho e de cana-de-açúcar, cujos maiores produtores mundiais são, respectivamente, Estados Unidos e Brasil. Nos dois casos, a proeminência na produção do bioetanol deveu-se a políticas governamentais, iniciadas durante a década de 1970, que direcionaram a capacidade

de produção agrícola já existente para aplicações alternativas, devido tanto aos choques do petróleo quanto aos baixos preços dessas *commodities* (milho e cana) no mercado internacional.

No Brasil, os canaviais ocupam mais de oito milhões de hectares e geram 500 milhões de toneladas de cana-de-açúcar todos os anos (HLPE, 2013). As pesquisas científicas e tecnológicas nacionais se intensificaram a partir da década de 1970, com uma série de investimentos públicos, entre eles a criação do Planalsucar (Programa Nacional de melhoramento da cana-de-açúcar) e da Copersucar (Cooperativa de Produtores de Cana-de-Açúcar, Açúcar e Alcool do Estado de São Paulo), (Furtado, 2011).

Tais Programas fortaleceram a pesquisa e o desenvolvimento no país, e atualmente, o Brasil ocupa a posição de referência internacional na viabilização do bioetanol a partir da cana-de-açúcar (HLPE, 2013), também se destaca pela participação significativa de fontes renováveis em sua matriz energética, estimada em 42,4% em 2012, com 15,4% proveniente de cana de açúcar como matéria-prima e a formação das redes tecnológicas de pesquisa a partir da colaboração internacional (Bueno, 2016). O Gráfico 3 apresenta a comparação entre o uso dos combustíveis renováveis e não renováveis no Brasil e as projeções futuras para os anos de 2016 e 2021.



**Gráfico 3. demanda de energia – energias renováveis e não renováveis –Brasil (2012-2021)**

Fonte: Elaboração própria, a partir dos Anuários Estatísticos de 2004 e de 2014 da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP).

Em relação a real capacidade dos biocombustíveis em reduzir as mudanças climáticas resultantes da emissão de gases causadores do efeito estufa, o etanol de cana-de-açúcar e o etanol lignocelulósico estão entre os mais significativos. A Tabela 2.1 apresenta as comparações. Os números negativos significam um aumento líquido de emissões de gases de efeito estufa quando comparados aos combustíveis fósseis.

**Tabela 1. Reduções de emissões de GEE dos biocombustíveis em comparação com a gasolina e diesel**

Tipo de produto	Redução de emissões (%)
Etanol de cana	65 / 105
Etanol de trigo	-5 / 90
Etanol de milho	-20 / 55
Etanol de beterraba	30 – 60
Etanol lignocelulósico	45-112
Biodiesel lignocelulósico	5-120

Fonte: Traduzido, adaptado. HLPE (2013).

O biocombustível derivado de cana apresenta redução de emissão de gases de efeito estufa entre 65/105%, enquanto que o etanol de milho apresenta (índice negativo) de -20/55%. Dos que mais contribuem para redução dos gases de efeito estufa está o etanol lignocelulósico (2G) e o biodiesel lignocelulósico, podendo alcançar reduções de até 100% das emissões em comparação aos combustíveis fósseis.

Nota-se a enorme vantagem do etanol brasileiro (derivado da cana) diante não somente dos combustíveis fósseis, mas também do etanol norte-americano, derivado do milho<sup>1</sup>.

### 3. Metodologia

#### 3.1 Delimitando a amostra de patentes

Diversos trabalhos são baseados em informações contidas nos documentos de patentes para investigar fenômenos relacionados com a economia da inovação (Krafft et al, 2009; Ventura et al, 2013; Epicoco, 2013; Dal Poz, Silveira, Masago, 2013; Souza, 2013). A amostra de patentes relacionadas com o tema 'bioetanol' para a investigação de caminhos tecnológicos foi extraído a partir da base de *Derwent Innovations Index* durante o período de janeiro de 1975 e de junho de 2015.

Para levantar as patentes da base de dados da Derwent foi utilizada uma ferramenta chamada *IPC STATS Search*, disponível no site da *World Intellectual Property Organization* (WIPO). A Classificação Internacional de Patentes (IPC<sup>2</sup>), estabelecido pelo Acordo de Estrasburgo 1971, prevê um sistema hierárquico de símbolos independentes da língua para a classificação de patentes e modelos de utilidade, de acordo com as diferentes áreas de tecnologia a que pertençam. Esta ferramenta possibilita que, com a inserção de uma ou de uma combinação de palavras-chave, os IPCs que mais se relacionam com as palavras propostas sejam recuperados. Trata-se de uma ferramenta que se baseia no conteúdo dos documentos do *PATENTSCOPE* (base de dados, disponível no site da WIPO, que fornece acesso às solicitações de patentes, em formato de texto completo, feitas por meio da *Patent Cooperation Treaty* – PCT).

Os termos utilizados com este instrumento foram os seguintes: *ethanol, biofuels, bioenergy, sugarcane, biomass, lignocellulosic, cellulose, hemicellulose, biochemical conversion, bioconversion, enzymatic hydrolysis, hydrolysis, enzyme, fermentation, saccharification*.

Observou-se, com a inclusão destes termos, uma predominância de patentes em subclasses C12P (processos de fermentação ou processos que utilizem enzimas para sintetizar uma composição ou composto químico desejado ou para separar isômeros ópticos de uma mistura racêmica) e C12N (micro-organismos ou enzimas; suas composições; propagação, conservação, ou manutenção de micro-organismos; engenharia genética ou de mutações; meios de cultura).

A investigação dos subgrupos (menor nível de desagregação do IPC, com 8 dígitos) pertencente a classe C12P e C12N, e resultou na seleção de 23 códigos de IPCs de interesse (Tabela 2).

Com os filtros sobre o nível dos subgrupos anteriormente definidos, passou-se para o segundo passo do percurso metodológico: a construção da amostra patente. A base utilizada foi o *Derwent Innovations Index*, de propriedade da *Thomson Reuters* e reúne informações de escritórios e diversos países, incluindo

---

<sup>1</sup> De acordo com Furtado et al. (2011), 'the Brazilian route is more competitive and generates less greenhouse gases than the US route, mainly because the energy required to run the mill comes from biomass in the form of sugarcane bagasse'

<sup>2</sup> Disponível no site da WIPO em < <http://www.wipo.int/classifications/ipc/en/> >

o norte-americano (USPTO), europeu (EPO) e japonês (JPO). A opção por esta base foi devido ao fato de que ele cobre muitos escritórios de patentes.

**Tabela 2. Seleção de IPCs (subgrupos)**

Subclasse	Descrição
C12P7 06	(...) Etanol, não para bebida
C12P7 08	(...) produzido como subproduto ou a partir de substrato de resíduo ou de material celulósico
C12P7 10	(...) substrato contendo material celulósico
C12P7 14	(...) estágios múltiplos de fermentação; utilização de fermentação com diferentes tipos de micro-organismos...
C12P 19/14	(...) produzidos pela ação de carboidrase, p. ex. por alfa-amilase
C12P 39/00	(...) processos envolvendo micro-organismos de diferentes gêneros no mesmo processo, simultaneamente
C12N 1/15	(...) modificados pela introdução de material genético exógeno
C12N 1/16	(...) leveduras; seus meios de cultura
C12N 1/18	(...) levedura de pão; levedura de cerveja
C12N 1/19	(...) modificados pela introdução de material genético exógeno
C12N 1/20	(...) Bactérias; seus meios de cultura
C12N 1/21	(...) modificados pela introdução de material genético exógeno
C12N 1/22	(...) Processos que utilizaram, ou meios de cultura que contêm, celulose ou seus hidrolisados
C12N 9/02	(...) Óxidoredutases, p. ex. Luciferase
C12N 9/04	(...) agindo sobre grupos CHOH como doadores, p. ex. glicose oxidase, lactato desidrogenase
C12N 9/14	(...) Hidrolases
C12N 15/1	(...) Preparação de mutantes sem introdução de materiais genéticos exógenos.
C12N 15/2	(...) Preparação de células híbridas por fusão de duas ou mais células, p. ex. fusão de protoplastos
C12N 15/3	(...) Bactérias
C12N 15/4	(...) Fungos
C12N 15/5	(...) Células vegetais
C12N 15/10	(...) Processos para o isolamento, preparação ou purificação do DNA ou RNA
C12N 15/11	(...) Fragmentos de DNA ou RNA; suas formas modificadas

Fonte: Classificação Internacional de Patentes, disponível em: <<http://ipc.inpi.gov.br>>

Com os filtros sobre o nível dos subgrupos anteriormente definidos, passou-se para o segundo passo do percurso metodológico: a construção da amostra patente. A base utilizada foi o *Derwent Innovations Index*, de propriedade da *Thomson Reuters* e reúne informações de escritórios e diversos países, incluindo o norte-americano (USPTO), europeu (EPO) e japonês (JPO). A opção por esta base foi devido ao fato de que ele cobre muitos escritórios de patentes.

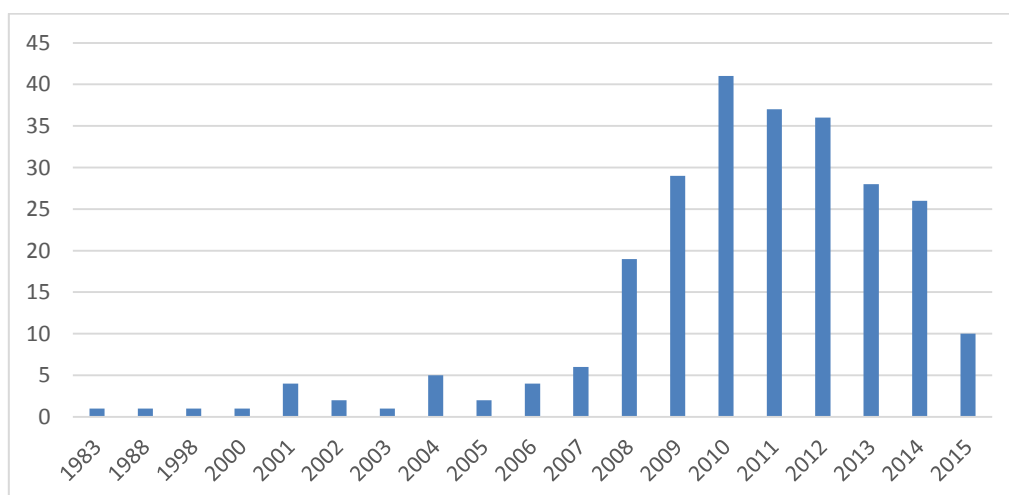
A combinação dos termos foram os seguintes:

- ((C12P-007/06 **OR** C12P-007/08 **OR** C12P-007/10 **OR** C12P-007/14 C12P-019/14 **OR** C12P-039/00 **OR** C12N-001/15 **OR** C12N-001/16 **OR** C12N-001-18 **OR** C12N-001/19 **OR** C12N-001/20 **OR** C12N-001/21 **OR** C12N-001/22 **OR** C12N-009/02 **OR** C12N-009/04 **OR** C12N-009/14 **OR** C12N-015/01 **OR** C12N-015/02 **OR** C12N-015/03 **OR** C12N-015/04 **OR** C12N-015/05 **OR** C12N-015/10 **OR** C12N-015/11) **AND** (sugarcane\* **AND** (ethanol\* **or** biofuels\* **or** biomass **or** bioenergy))

Este procedimento é chamado de data mining, cujo princípio maior é tratar grandes quantidades de dados. Porém, esse método não pode ser confundido com a extração de dados, mas sim, a extração de padrões de conhecimentos (coleta, extração, armazenagem, análise e estatísticas), (WU, 2014).

Para a construção da amostra, foram usadas as 23 classes de IPCs e os códigos selecionados pelo filtro de pesquisa. A combinação de palavras-chave foi cruzada com as classes de IPCs, resultando num total de 263 patentes que compõe a amostra (Graf. 4), a agregação foi feita com o uso do operador booleano (OR) e (AND), é muito provável que a pesquisa feita na base da *Derwent* recuperou informação de patentes que

têm pelo menos um dos termos aplicados no assunto abordado neste artigo. Elas foram combinadas para uma busca precisa para tecnologias de bioetanol, especialmente, a partir da cana-de-açúcar.

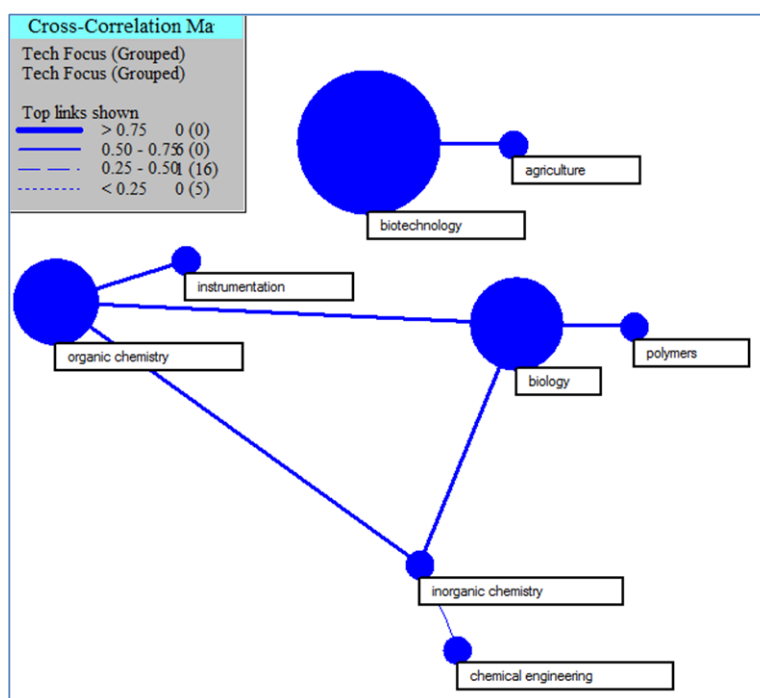


**Gráfico 4. Patentes da amostra. Distribuição por ano.**

Fonte: base Derwent Innovations Index, disponível em <<http://www.periodicos.capes.gov.br/>>.

A opção para compor a amostra de patentes para identificar as especializações tecnológicas para bioetanol é, inicialmente, entender que as inovações tecnológicas devem ser pensadas como sistemas complexos contendo vários elementos que são interdependentes e interagem de várias maneiras. Estas interdependências e interações entre os vários elementos é o que cria os *trade-offs* e, portanto, abre caminhos para diferentes rotas tecnológicas (Dosi, 1982).

Observe a Fig. 1, a amostra que compõe as 263 patentes na rota dos biocombustíveis mostra oito grandes áreas de especialização, com interdependência entre as áreas, principalmente, entre a biotecnologia, biologia e química orgânica. A agregação e *data mining* foi feita com o uso do software *VantagePoint*<sup>3</sup>.



**Figura 1. Grandes áreas de especialização que compõe a amostra de patentes.**

Fonte: base Derwent Innovations Index, disponível em <<http://www.periodicos.capes.gov.br/>>.

\*Ferramenta cross-correlation map. –The VantagePoint.

<sup>3</sup> O programa TI pode ser acessado através do link <<https://www.thevantagepoint.com/>>.



As grandes áreas de especialização que compõe a amostra de patentes são biotecnologia e agricultura, sendo as tecnologias na área de biotecnologia bem maior em comparado a tecnologias na agricultura. Biologia e química apresentam concentração de tecnologias semelhantes, se relacionando em menor grau com tecnologias nas áreas de polímeros, instrumentação, química orgânica e engenharia química. Essa dimensão ajuda a orientar o próximo passo da construção metodológica.

### 3.2 Aplicando a rede de IPCs em bioetanol

A matriz de coocorrência tem a função de mostrar a relação entre dois elementos ou mais, um chamado de referência e o outro de vizinho. Em outras palavras, a matriz (obrigatoriamente quadrada) representa em cada elemento o número de vezes que ocorreu uma transição do vizinho considerando a distância e a direção. O programa VantagePoint, através de um comando operacional, gera a matriz de acordo com as variáveis escolhidas pelo usuário. O mesmo programa gera a visualização da matriz em uma imagem de rede. Essa rede está representada por grafos, que é uma ilustração da rede representada da seguinte maneira: um grafo direcionado consiste de um conjunto  $V$  de vértices (ou nós), um conjunto  $E$  de arestas (links ou Edges),  $E \rightarrow V$ , onde  $(e)$  é a fonte e  $(v)$  é o alvo da aresta direcionada (Goyal, 2007). Então, qualquer conjunto de elementos no qual alguns deles estão conectados através de links será representado.

Esse procedimento será aplicado para investigar a composição das classes de IPCs encontradas na amostra de 263 patentes. A partir dos métodos anteriormente citados para a construção da amostra, formamos o conjunto de amostra de IPCs das patentes. Foram identificadas 93 classes de IPCs (4 dígitos) formando um conjunto total de 544 sub-classes de IPCs (8 dígitos). Isso significa que destas 93 classes (Tab. 3) identificadas na amostra de patentes, cujas áreas tecnológicas podem estar classificadas de A a H, possuem 544 áreas tecnológicas diferentes.

**Tabela 3. Classes tecnológicas de IPCs (4 dígitos) identificadas na amostra de patentes.**

Classes de IPCs	Descrição de IPCs
Classes de IPCs (4dígitos)	C12P; C12N; C12R; A01H; C07H; C13K; C07K; C12M; C10L; C08B; D21C; A01N; C12Q; C07C; A23K; B01D; A23L; B09B; A01P; A61K; C08H; C08G; C13B; B01J; C12S; C13D; A01G; C02F; C07B; C07G; C08F; C08J; C08L; C10G; C11D; C12F; C40B; D06M; G01N; A01K; A21D; A23C; A61L; A61P; B65D; C09K; C10J; C11C; C12G; D01F; D21H; F23G; G06F; A01C; A01D; A23D; A46D; A61Q; A62D; B04B; B04C; B05B; B05D; B09C; B29B; B30B; B32B; B82Y; C01B; C01C; C01F; C05D; C05F; C05G; C08K; C09B; C09J; C10M; C12C; C12J; C13C; C13F; D06P; D07B; D21B; D21F; D21J; F02B; F16L; G06K; G09C; H01H; H01M.

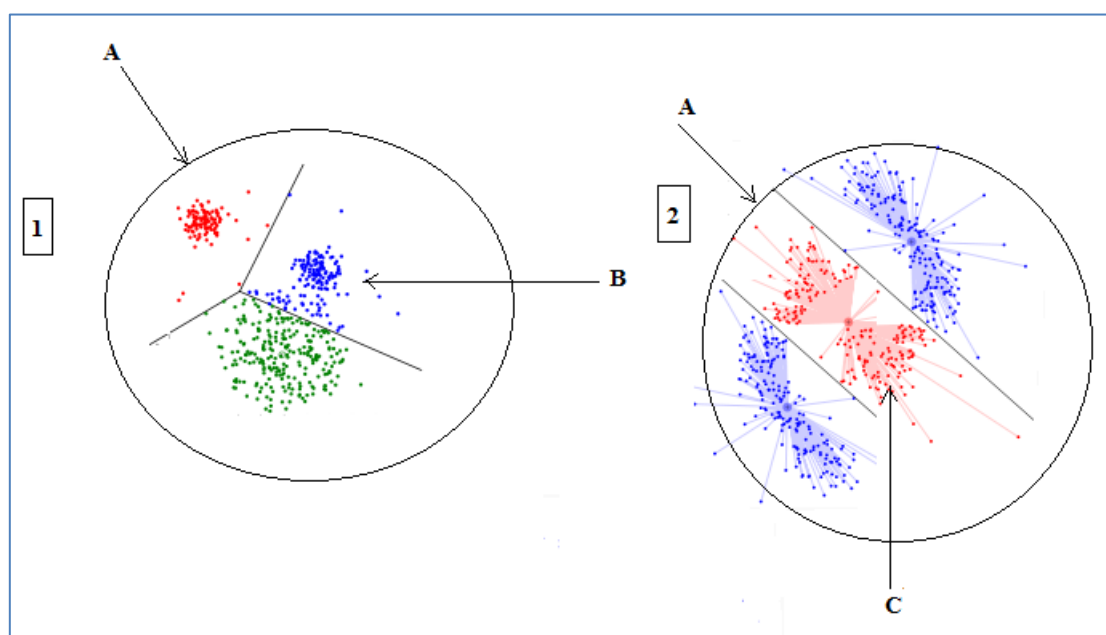
Fonte: Preparação dos dados, a partir da base *Derwent Innovations Index*, disponível em <<http://www.periodicos.capes.gov.br/>>.

Uma vez que a quantidade de classes é alta, o número de todas as sub-classes possíveis de "conteúdo de biocombustíveis" também é grande. Espera-se que, com o desenvolvimento da rede de IPCs, algumas dessas classes sejam dirigidas para um grupo de cluster de tecnologias que potencialmente sintetizam um conjunto de especializações tecnológicas.

Para avaliar o grau de especialização, ou seja, se os caminhos surgiram relativos às tríades específicas de especialização, foi aplicado a matriz de coocorrência a partir do programa VantagePoint. O algoritmo do programa TI aplica a rede para um esquema iterativo orientado por uma árvore de extensão mínima da rede, a fim de gerar coordenadas para os nós expressando ligações inter-cluster e intra-cluster. Este indicador mostra a especialização por área tecnológica e a dependência mútua entre os caminhos tecnológicos de interesse.

Um cluster (ou agrupamento) representa um conjunto de propriedades de que os objetos no mesmo grupo são mais semelhantes (em algum sentido ou outro) entre si do que com os de outros grupos (ou outros clusters). O agrupamento de propriedades com base na conectividade, também conhecido como agrupamento hierárquico, representa um conjunto de agrupamento que pode ser descrito, pela distância máxima necessária para ligar as partes do conjunto de propriedades. Isto é, as distâncias diferentes, os grupos diferentes, irão se formar de acordo com as conexões entre eles. Esse procedimento ocorre porque em clusters hierárquicos de conexão, os algoritmos não fornecem uma única partição do conjunto de dados, mas sim uma extensa hierarquia de clusters que se fundem uns com os outros em determinadas distâncias. Ou seja, este método não irá produzir somente um único conjunto de dados, mas também o sentido de valores extremos, como aglomerados adicionais ou até mesmo outros clusters, conhecido como "fenômeno de encadeamento" (Achtert et al., 2006).

Analisar os fenômenos de encadeamento, intra-cluster e inter-cluster, é importante, uma vez que permitem investigar exatamente esse encadeamento, de ligações externas e internas no conjunto de um aglomerado de dados. Exemplificando, note a Figura 2, em **A** temos um cluster, **B** temos um intra-cluster e **C** um inter-cluster. Isto é, **B** e **C** tem algum nível de interdependência de **A**, explicada por duas possíveis causas: 1) **A** tem ligações diretas que formam o cluster **B**, (imagem 1), e 2) o cluster **A** originou o cluster **C**, este cluster está entre clusters (imagem 2). O inter-cluster apresenta características de agrupamento ao intra-cluster, ou seja, pode ser o cluster que originou um grande conjunto de dados que formaram intra-clusters, ou pode estar entre clusters, intermediário ou não.



**Figura 2. Ilustração de clusters**

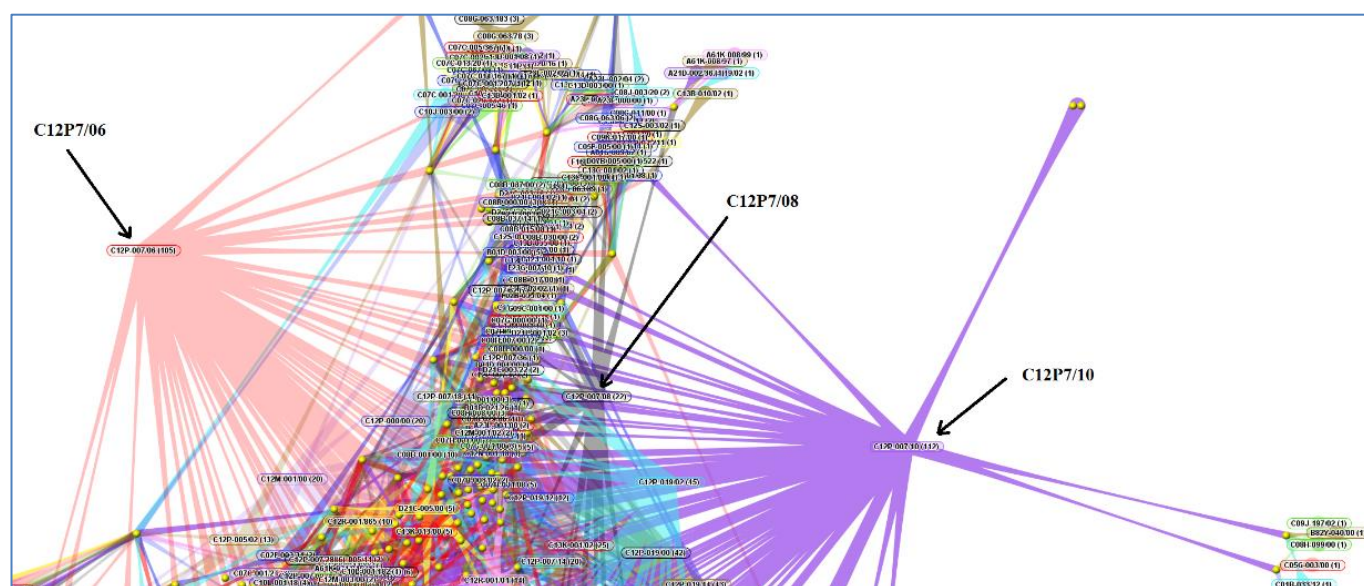
Fonte: Adaptado, a partir de Achtert et al (2006).

#### 4. Resultados

A fim de aplicar a rede, para identificar os clusters, foram inseridas as 544 classes de IPCs da amostra de patentes para produção de bioetanol a partir dos métodos anteriormente citados. Finalmente, verificou-se a formação de três clusters principais do agrupamento que forma o cluster total. São as classificações tecnológicas das classes: "C12P7/06", "C12P7/08" e "C12P7/10".

Note a rede a seguir (Fig. 3), a rede destaca os três principais clusters de tecnologias encontrados na amostra de IPCs das patentes. São os clusters das classes de IPCs: C12P7/06 à esquerda (representada pela cor rosa) "...*Etanol, i.e. não para bebida*", outro cluster a direita (representada pela cor roxa) que corresponde à classe de IPC C12P7/10 "...*substrato contendo material celulósico*", e no centro um outro

cluster (representado pela cor cinza escuro) da classe C12P7/08 “...produzido com sub produto ou a partir de substrato de resíduo ou de material celulósico”.



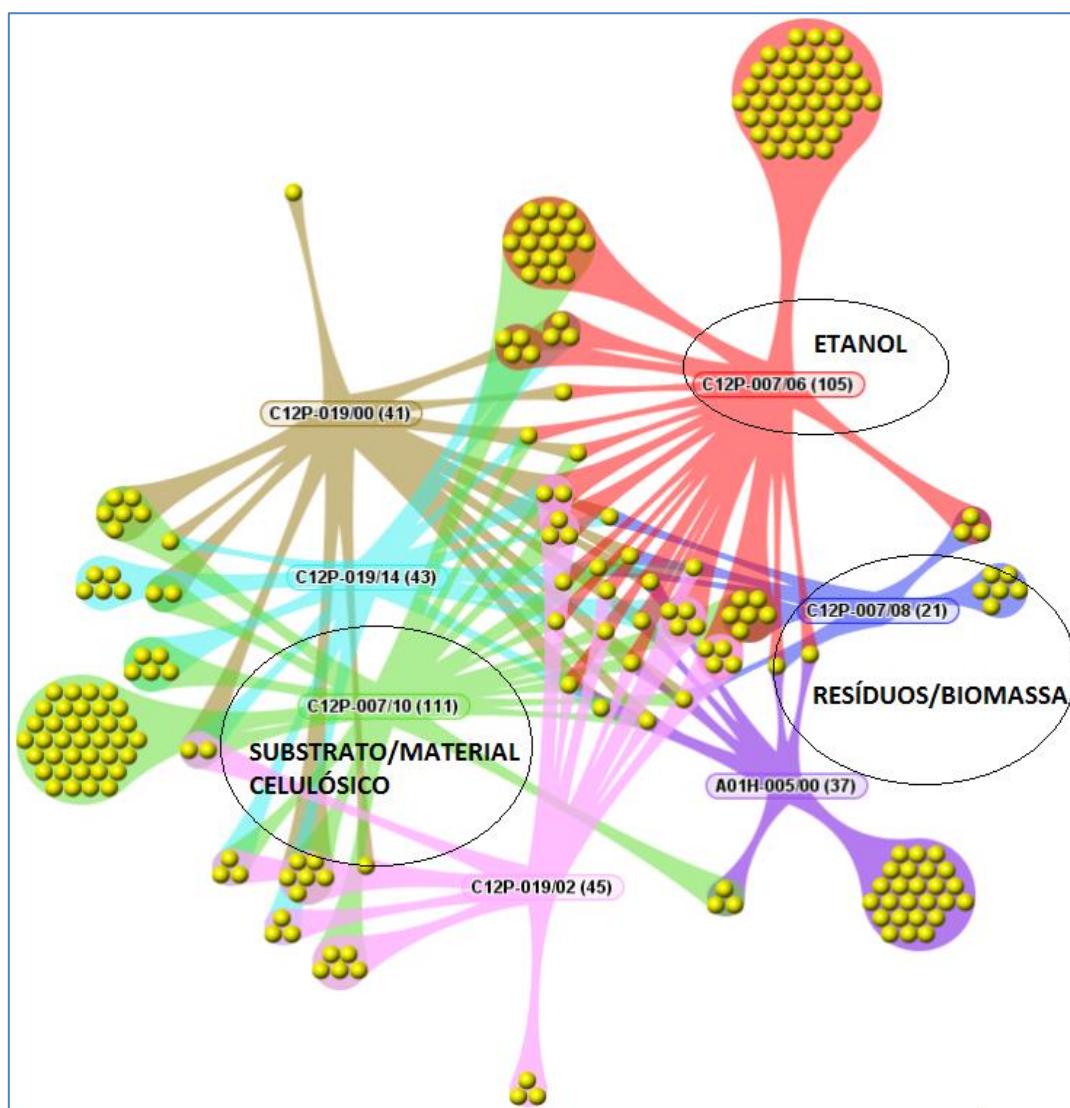
**Figura 3. Detalhe dos clusters identificados na Rede de IPCs.**

Fonte: Dados da pesquisa, a partir da base Derwent Innovations Index, disponível em <<http://www.periodicos.capes.gov.br/>>.

\*Ferramenta Aduna cluster map. –The VantagePoint.

No detalhe da rede, quando agrupamos as patentes somente para estes três clusters, o cluster 06 é o mais desenvolvido, além das tecnologias tradicionais de etanol, também há tecnologias de áreas de avanço, como as que pertencem ao cluster 08 e 10. As tecnologias do cluster 08 e 10 captam as trajetórias que estão dentro do cluster 06, por outro lado, o cluster 08 não capta do 10 e vice versa, isto significa que o cluster 06 é o cluster principal de tecnologias e a formação dos outros clusters são feitas a partir das trajetórias do 06.

Note o detalhe na rede (Fig. 4), o cluster 06/ETANOL é o mais desenvolvido, e sua área de especialização está em fase de expansão, e suas trajetórias tecnológicas estão desdobrando a formação de outros clusters tecnológicos, como o 08 e 10.



**Figura 4. Detalhe do clusters, etanol/resíduo e substrato. Rede de IPCs.**

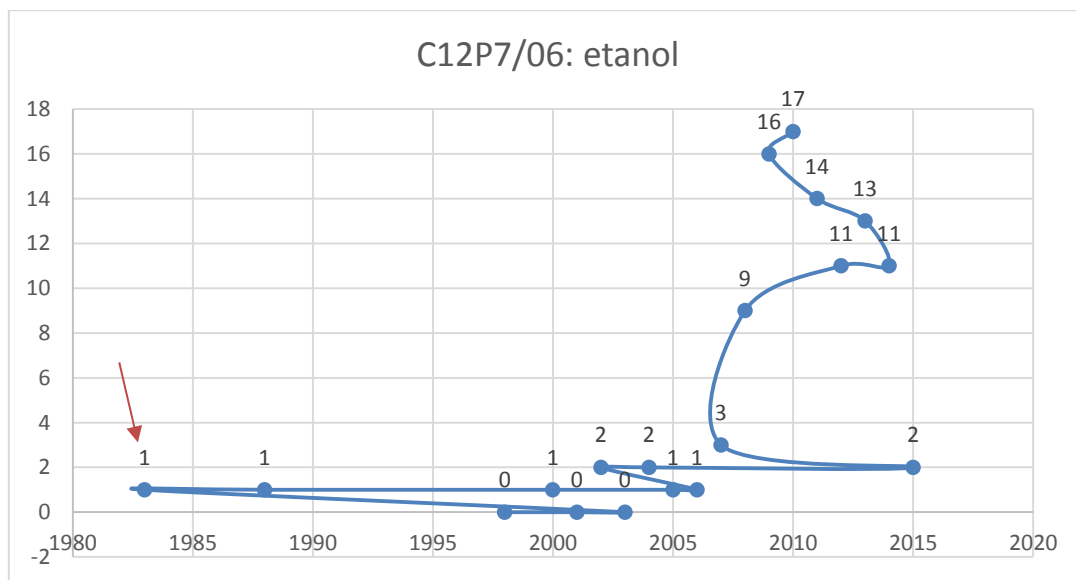
Fonte: Dados da pesquisa, a partir da base Derwent Innovations Index, disponível em <<http://www.periodicos.capes.gov.br/>>.

\*Ferramenta Aduna cluster map. –The VantagePoint.

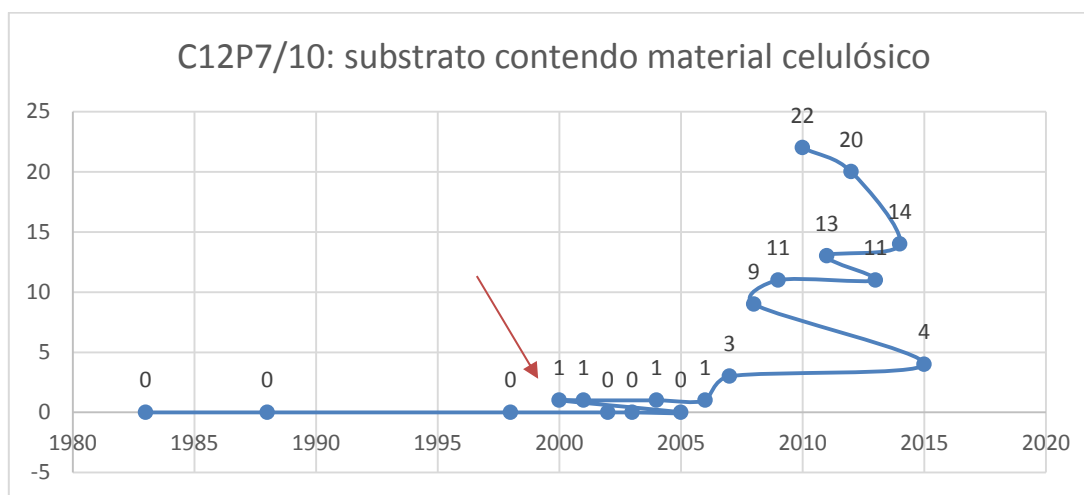
Para responder a partir de qual cluster o grau de especialização está se manifestando no caminho de formação dos intra-clusters (08 e 10), foi necessário examinar a extensão da rede de IPCs a partir das informações das patentes de cada cluster. Este indicador revela a presença (ou não) de dependência de caminho de interesse e os seus respectivos clusters.

A análise seguiu o mesmo método anterior para formar a rede de IPCs, a diferença é que os dados foram separados por grupos de clusters (com base na classificação de IPCs da amostra). O primeiro cluster a ser aplicado o método foi o C12P7/06, todas as patentes deste cluster foram agregadas por ano.

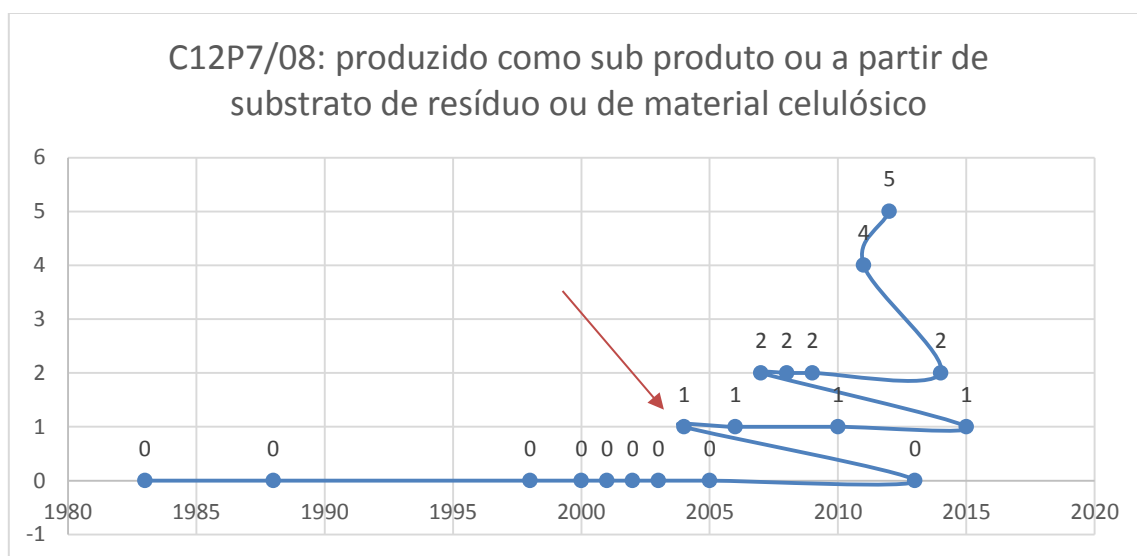
Observe os gráficos a seguir (5, 6 e 7), as classes de IPCs presentes nas patentes desse cluster iniciam-se em 1983 e percorre um caminho até 2015. Ao contrário ocorre com os clusters C12P7/10 (Graf. 6), em que o caminho tecnológico inicia a partir dos anos 2000, e C12P7/08 (Graf. 7) a partir de 2005. Para identificar as especializações tecnológicas de cada cluster, as patentes foram analisadas uma a uma. Os gráficos são apresentados, e em seguida a descrição da análise.



**Gráfico 5. timeline: Cluster C12P7/06.**  
 Fonte: Elaboração própria.



**Gráfico 6. timeline: Cluster C12P7/10.**  
 Fonte: Elaboração própria.



**Gráfico 7. timeline: Cluster C12P7/08.**  
 Fonte: Elaboração própria.

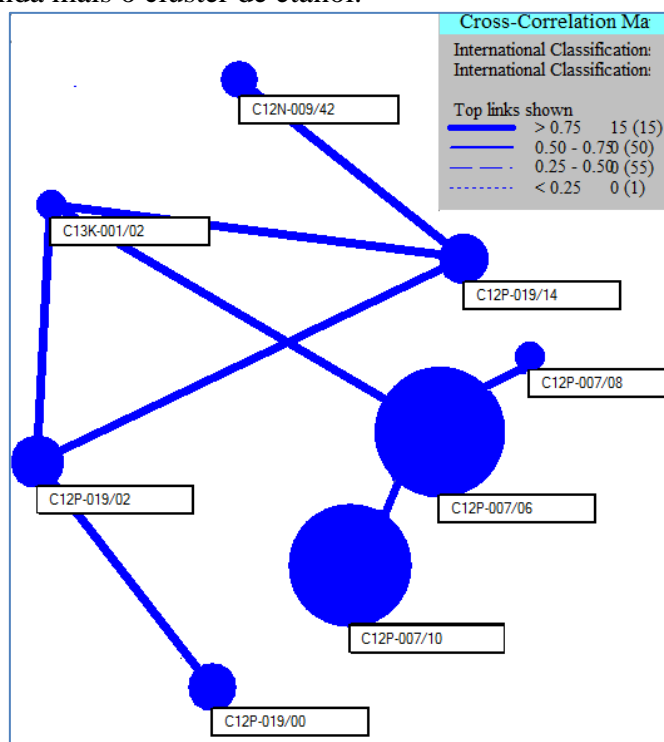
O cluster C12P7/06 é o cluster de tecnologias de etanol. As tecnologias são desde processos de hidrólise enzimática, *mix* de fermentação por aquecimento, modificação genética, biotecnologia, sacarificação, microrganismos e enzimas, até a engenharia genética. É um cluster de áreas genéricas. Desde tecnologias tradicionais até as mais avançadas. A partir de 2008 as patentes começam a apresentar forte crescimento, tendo uma concentração nesse período, ou seja, este cluster continua a se desenvolver.

Já no cluster C12P7/10, as tecnologias apresentam um nível maior de complexidade. Este cluster está especializado em tecnologias que envolvem estudos da lignina. A lignina é uma macromolécula que confere rigidez a parede celular da planta, dificultando o acesso dos açúcares que estão na biomassa (bagaço). As áreas principais deste cluster são a química orgânica e a biotecnologia. O período é recente, as patentes iniciam-se nos anos 2000. Assim como cluster de etanol, este apresenta um elevado grau de desenvolvimento de patentes no período recente. É um cluster novo, e que continua a se desenvolver. Este cluster capturou trajetórias tecnológicas do cluster 06, e não do cluster 08.

A complexidade tecnológica aumenta ainda mais no cluster C12P7/08, ou seja, um cluster de especialização em tecnologias para converter biomassa em etanol. É o menor cluster em comparado aos outros, com patentes iniciadas em 2005. Neste cluster as tecnologias são nas áreas de química orgânica, física orgânica, processos completos para retirar a lignina, genética, microrganismos, enzimas, pré-tratamento, sacarificação para segunda geração, DNA recombinante, biotecnologia. Este cluster capturou trajetórias do cluster 06, mas não do cluster 10.

## 5. Conclusão

A investigação da aplicação da rede de IPCs para identificar a formação de clusters de tecnologias para produção de bioetanol permitiram concluir que existe uma concentração de especialização em três classes tecnológicas, a C12P7/06 como sendo o inter-cluster direcionador do progresso inovativo, e os intra-clusters C12P7/08 e C12P7/10 como clusters em etapa de especialização para bioetanol a partir da biomassa. Isto é, o inter-cluster está a direcionar a formação de trajetórias tecnológicas dos intra-clusters. isto significa que os intra-clusters capturaram as trajetórias do cluster 06, e estão desenvolvendo suas trajetórias, fortalecendo ainda mais o cluster de etanol.



**Figura 8. Inter-cluster e a formação dos intra-clusters.**

Fonte: Dados da pesquisa, a partir da base Derwent Innovations Index, disponível em <<http://www.periodicos.capes.gov.br/>>.

\*Ferramenta Aduna cluster map. –The VantagePoint.



No detalhe da rede (Fig. 8), o cluster mais desenvolvido é o de etanol, em seguida a formação do intra-cluster C12P7/10, também bem desenvolvido. O cluster C12P7/08 é menor em comparado aos outros. Estes dois capturaram tecnologias do inter-cluster de etanol, fortalecendo o cluster.

As pesquisas de tecnologias de hidrólise enzimática de celulose, fermentação, modificação genética para a produção de bioetanol de cana-de-açúcar são patenteados durante os anos 80, baseados em grande parte da classe de IPC, C12P7/06. As pesquisas realizadas a partir de técnicas de segunda geração começou a ser patentado na maior peso a partir dos anos 2000. Precisamente o período em que se destaca o fortalecimento dos dois clusters de bioetanol C12P7/08 e C12P7/10, lignina e biotecnologia de sacarificação.

A contribuição do presente trabalho é mostrar como uma base de conhecimentos gerais vai sendo apropriada (no caso do presente trabalho via patentes) e favorece a configuração de novas trajetórias tecnológicas que amplificam o paradigma tecnológico representado pelos combustíveis renováveis e pelas potencialidades da química verde (por exemplo, lactatos, álcoois superiores, óleos para cosméticos, combustíveis para avião). A concepção de inovação tecnológica torna-se cada vez mais sistêmica, sequencial, configurando sistemas complexos modulares, em que gargalos para a viabilidade da inovação aparecem em pontos específicos que não são apropriados pelo agente que inova em produto.

Ao mesmo tempo, os desdobramentos das classes de IPC mostram que ao longo de trajetórias de inovação em tecnológicas habilitadoras, surgem novos produtos da biotecnologia, produtos de um mercado de pesquisa, cujo consumidor é a empresa inovadora *downstream*, como no caso das enzimas para hidrólise ou as tecnologias de sacarificação. Os processos de natureza incremental estão acoplados aos processos de inovações radicais na forma de redes de pesquisa, que recriam na atualidade as parcerias pré-competitivas e as novas formas de acordos contratuais, em que a divisão dos ganhos se dá precocemente, pela avaliação da potencialidade da contribuição de um determinado parceiro ao conjunto da rede. O presente trabalho contribui para entendimento deste processo e demanda investigações dos agentes envolvidos nos processos, seu tipo de filiação e os acordos de cooperação entre eles.

## 6. Referências

- ACHTERT, E.; BÖHM, C.; KRÖGER, P. (2006). "DeLi-Clu: **Boosting Robustness, Completeness, Usability, and Efficiency of Hierarchical Clustering by a Closest Pair Ranking**". LNCS: Advances in Knowledge Discovery and Data Mining. Lecture Notes in Computer Science 3918: 119–128.
- Base *Derwent Innovations Index*, available at <<http://www.periodicos.capes.gov.br/>>.
- BUENO, C. S.; SILVEIRA, J. M. F. J.; BUAINAIM, A. M. An evaluation of how Brazilian scientific research contributes to the patenting of sugarcane ethanol. **Anais..., 19th ICABR Conference**, Italy, June, 2015. ICABR International Consortium on Applied Bioeconomy Research, 2015.
- DAL POZ, M. E.; SILVEIRA, J. M. F. J. da; MASAGO, F. Innovation networks in biotechnology: lessons from Protil; André Luís Ribeiro Lima. (Org.). **A contribuição da ciência e tecnologia para o desenvolvimento do agronegócio**. 1ed. Viçosa/MG: Suprema Gráfica e Editora, v. 1, 2013, p. 1-656.
- DOSI, G. Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technological change. **Research Policy**, Amsterdam, v. 11, n. 3, p. 147-162, 1982.
- EPICOCO, M. Knowledge patterns and sources of leadership: mapping the semiconductor miniaturization trajectory. **Research Policy**, v. 42, n. 1, p.180-195, Feb. 2013.
- FRENKEN, K. A complexity approach to innovation networks. The case of the aircraft industry (1909 - 1997). **Research Policy**, v. 29, n. 2, p.257-272, Feb. 2000.
- GOYAL, SANJEEV. Connections: An introduction to the Economics of Networks, 2007. Princeton and Oxford: Princeton University Press.
- HLPE. **Biofuels and food security: a report by the high level panel of experts on food security and nutrition**. Committee on World Food Security. Roma. 2013. 132p.

IEA. 2013. Status of advanced biofuels demonstration facilities in 2012. A report to IEA Bioenergy Task 39. D. Bacovsky, N. Ludwiczek, M. Ognissanto, & M. Wörgetter  
 IPC STATS, available at <<http://web2.wipo.int/ipcpub/search/stats/#version=20140101&lang=en>>. Acesso em: Nov. 2015.

KRAFFT, J.; QUATRARO, F.; SAVIOTTI, P. The evolution of knowledge base in knowledge-intensive sectors: social network analysis of biotechnology. **Working Paper Series**, n. 9, 2009.

SIMS, R.; TAYLOR, M.; SADDLER, J.; MABEE, W. From 1<sup>st</sup> to 2<sup>nd</sup> generation biofuel technologies: an overview of current industry and RD&D activities. Paris: **International Energy Agency (IEA)**, Nov. 2008. 124p. Available at <<http://www.ieabioenergy.com/MediaItem.aspx?id=6060>>. Acesso em: Dez. 2015.

SOUZA, Luiz Gustavo Antonio de. Redes de inovação em etanol de segunda geração. 2013. Tese (Doutorado em Economia Aplicada) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2013. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/11/11132/tde-26082013-135850/>>. Acesso em: Dez, 2015.

VENTURA, V.; FRISIO, D.; FERRAZZI, G.; SILETTI, E. Forecasting the evolution of agbiotech innovation: lessons from patent data. **Anais...**, 17th ICABR Conference, Italy, June 2013.

World Intellectual Property Organization (WIPO), available at <<http://www.wipo.int/classifications/ipc/en/>>. Acesso em: Jan. 2016.

ZILBERMAN, D. The Economics of Sustainable Development. *Amer. J. Agr. Econ.* 96(2): 385–396; doi: 10.1093/ajae/aat075, 2013.

WU, Xindong et al. Data mining with big data. **IEEE transactions on knowledge and data engineering**, v. 26, n. 1, p. 97-107, 2014.