Patent mapping and patent statistics

Davide Lingua
Johannes Schaaf 4.5.3

Overview

• PATSTAT as a relational database
• case study
Pre-computer age data storage

Relational Databases

A Relational Database is a collection of tables:
A table is a collection of columns and rows

For instance table `dbo_tls201_appln` from PATSTAT:

<table>
<thead>
<tr>
<th>appln_id</th>
<th>appln_auth</th>
<th>appln_nr</th>
<th>appln_kind</th>
<th>appln_filing_date</th>
<th>ipr_type</th>
</tr>
</thead>
<tbody>
<tr>
<td>54198991 EP</td>
<td>710</td>
<td>F</td>
<td></td>
<td>01-04-2003 PI</td>
<td></td>
</tr>
<tr>
<td>54198992 EP</td>
<td>35274</td>
<td>F</td>
<td></td>
<td>30-05-2003 PI</td>
<td></td>
</tr>
<tr>
<td>54198994 EP</td>
<td>38093</td>
<td>F</td>
<td></td>
<td>11-06-2003 PI</td>
<td></td>
</tr>
<tr>
<td>54198995 EP</td>
<td>39095</td>
<td>F</td>
<td></td>
<td>06-06-2003 PI</td>
<td></td>
</tr>
<tr>
<td>54198995 EP</td>
<td>42718</td>
<td>F</td>
<td></td>
<td>06-06-2003 PI</td>
<td></td>
</tr>
<tr>
<td>54198997 EP</td>
<td>42908</td>
<td>F</td>
<td></td>
<td>20-06-2003 PI</td>
<td></td>
</tr>
<tr>
<td>54198995 EP</td>
<td>48095</td>
<td>F</td>
<td></td>
<td>03-07-2003 PI</td>
<td></td>
</tr>
<tr>
<td>54198999 EP</td>
<td>49481</td>
<td>F</td>
<td></td>
<td>05-06-2003 PI</td>
<td></td>
</tr>
<tr>
<td>54199000 EP</td>
<td>51370</td>
<td>F</td>
<td></td>
<td>07-07-2003 PI</td>
<td></td>
</tr>
<tr>
<td>54199001 EP</td>
<td>69606</td>
<td>F</td>
<td></td>
<td>05-04-2003 PI</td>
<td></td>
</tr>
<tr>
<td>54199002 EP</td>
<td>72277</td>
<td>F</td>
<td></td>
<td>05-09-2003 PI</td>
<td></td>
</tr>
<tr>
<td>54199003 EP</td>
<td>77060</td>
<td>F</td>
<td></td>
<td>18-09-2003 PI</td>
<td></td>
</tr>
<tr>
<td>54199004 EP</td>
<td>91707</td>
<td>F</td>
<td></td>
<td>16-10-2003 PI</td>
<td></td>
</tr>
</tbody>
</table>

Each column contains one element

Table `dbo_tls201_appln` has following elements:

- **appln_id**: Application identification
- **appln_auth**: Application authority
- **appln_nr**: Application number
- **appln_kind**: Kind of application
- **appln_filing_date**: Filing date of application
- **ipr_type**: Type of intellectual property right
- **appln_title_lg**: Language of title of application
- **appln_abstract_lg**: Language of abstract of application
- **internat_appln_id**: Application identification of designating international application
SQL: Structured Query Language

- Allows to talk to the relational database
- Necessary to extract information from any table
- Can connect several tables
- DML: Data Manipulation Language - part of SQL used to make queries

Sample Data Types

- Numeric:
  - INTEGER: whole numbers
  - DECIMAL(m,n): decimal numbers, m=digits, n=decimal digits
- Alphanumeric:
  - CHARACTER(n): string (fixed length)
  - VARCHAR(n): string (variable length)
- Date Data: DATE YYYYMMDD or variations like dd.mm.yyyy
- Time Data: TIME HHMMSS or variations like hh:mm:ss
Structure of a typical SQL Query:
SELECT Statement

Clauses:
- SELECT → Defines the result columns
- FROM → Table names
- WHERE → Conditions (which rows?) - optional
- ORDER BY → Sort the result rows - optional

An example

SELECT dbo_tls201_appln.appln_auth, dbo_tls201_appln.appln_nr
FROM dbo_tls201_appln
WHERE ((dbo_tls201_appln.appln_auth)="EP")
ORDER BY dbo_tls201_appln.appln_nr;

<table>
<thead>
<tr>
<th>appln_auth</th>
<th>appln_nr</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP</td>
<td>716</td>
</tr>
<tr>
<td>EP</td>
<td>35274</td>
</tr>
<tr>
<td>EP</td>
<td>37304</td>
</tr>
<tr>
<td>EP</td>
<td>38609</td>
</tr>
<tr>
<td>EP</td>
<td>39805</td>
</tr>
<tr>
<td>EP</td>
<td>42718</td>
</tr>
<tr>
<td>EP</td>
<td>42908</td>
</tr>
<tr>
<td>EP</td>
<td>48095</td>
</tr>
<tr>
<td>EP</td>
<td>49481</td>
</tr>
<tr>
<td>EP</td>
<td>51370</td>
</tr>
</tbody>
</table>
Operators and Wildcards

- Boolean Operators: AND, OR, NOT
- Comparison Operators:
  - equal to: =
  - not equal to: <>
  - greater than: >
  - greater than or equal to: >=
  - less than: <
  - less than or equal to: <=
  - BETWEEN...
- Truncation "%" and "LIKE" to search character strings: LIKE "univ%"
- Underscore "_" and "LIKE": LIKE "organi_e"

### Boolean operators

<table>
<thead>
<tr>
<th>Boolean Operator</th>
<th>Expressions</th>
<th>Explanation</th>
<th>Usage</th>
<th>Illustration</th>
<th>Illustration</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>A OR B</td>
<td>Retrieves records in which AT LEAST ONE of the search terms is present</td>
<td>- Search for synonyms - Include different spellings - Retrieve more records</td>
<td>![Illustration of OR operator]</td>
<td>![Illustration of OR operator]</td>
</tr>
<tr>
<td></td>
<td>A OR B OR C</td>
<td></td>
<td></td>
<td>![Illustration of OR operator]</td>
<td>![Illustration of OR operator]</td>
</tr>
<tr>
<td>AND</td>
<td>A AND B</td>
<td>Retrieves records in which ALL of the search terms are present</td>
<td>- Reduce number of records found</td>
<td>![Illustration of AND operator]</td>
<td>![Illustration of AND operator]</td>
</tr>
<tr>
<td></td>
<td>A AND B AND C</td>
<td></td>
<td></td>
<td>![Illustration of AND operator]</td>
<td>![Illustration of AND operator]</td>
</tr>
<tr>
<td>NOT</td>
<td>A AND NOT B (A OR B) AND NOT C</td>
<td>Retrieves records in which a certain search term in NOT present</td>
<td>- Exclude search terms from search results</td>
<td>![Illustration of NOT operator]</td>
<td>![Illustration of NOT operator]</td>
</tr>
</tbody>
</table>
An example:

SELECT dbo_tls201_appln.appln_auth, dbo_tls201_appln.appln_filing_date
FROM dbo_tls201_appln
WHERE (((dbo_tls201_appln.appln_auth)="DK") AND ((dbo_tls201_appln.appln_filing_date) BETWEEN #1/1/2004# AND #12/31/2004#));

The COUNT function:

For instance, I want to know the number of all EP direct filings:

- in a specific year (e.g. filed in 2004)
- only direct filings, that is, excluding:
  - applications of PCT origin
  - any other artificial applications (code "D2")
The SQL code:

SELECT Count(*) AS Expr1
FROM dbo_tls201_appln
WHERE ((dbo_tls201_appln.appln_auth)="EP")
   AND ((dbo_tls201_appln.appln_filing_date) Between #1/1/2004#
        And #12/31/2004#)
   AND ((dbo_tls201_appln.appln_kind)="A ")
   AND ((dbo_tls201_appln.internat_appln_id)=0));

Result of the query: 54,557

Case study n.1

Environmental Technology in Automotive Industry:

Hybrid vehicle technology
A look at the IPC-8 (2006.01)

B SECTION B — PERFORMING OPERATIONS; TRANSPORTING

B60 VEHICLES IN GENERAL

B60K ARRANGEMENT OR MOUNTING OF PROPULSION UNITS OR OF TRANSMISSIONS IN VEHICLES; ARRANGEMENT OR MOUNTING OF PLURAL DIVERSE PRIME-MOVERS; AUXILIARY DRIVES; INSTRUMENTATION OR DASHBOARDS FOR VEHICLES; ARRANGEMENTS IN CONNECTION WITH COOLING, AIR INTAKE, GAS EXHAUST, OR FUEL SUPPLY, OF PROPULSION UNITS, IN VEHICLES [1,8]

B60K 6/00: Arrangement or mounting of plural diverse prime-movers for mutual or common propulsion, e.g. hybrid propulsion systems comprising electrical and internal combustion motors

Country analysis for B60K 6/00

We would like to analyse the applications filed at the EPO:

• Countries of applicants: Which are the dominant countries?
• Countries of inventors: Are they the same?
• Cooperation of applicants: How much research cooperation?
• What can we learn in terms of technology acquisition and workforce distribution?
Let's build our query

Steps
• Select IPC core symbol: B60K6/00
• Select application authority: EP
• Select the elements to be displayed:
  – application details
  – inventor and applicant details

Tables
⇒ dbo_tls209_appln_ipc  
⇒ dbo_tls201_appln
⇒ dbo_tls201_appln
⇒ dbo_tls207_pers_appln and
dbo_tls206_person

Look for IPC core symbol B60K6/00

SELECT dbo_tls209_appln_ipc.ipc_class_symbol,  
FROM dbo_tls209_appln_ipc  
WHERE (dbo_tls209_appln_ipc.ipc_class_symbol)="B60K  6/00"

Note the three empty spaces before the "6"!
The query in MS Access "Design View"

Result of query: 16,671

Look for application authority: EP

```
SELECT dbo_tls201_appln.appln_auth
FROM dbo_tls201_appln
WHERE (dbo_tls201_appln.appln_auth)="EP"
```
The query in MS Access "Design View"

Result of query: 2,231,594

Combining IPC class and application authority

SELECT dbo_tls209_appln_ipc.ipc_class_symbol,
dbo_tls201_appln.appln_auth,
FROM (dbo_tls209_appln_ipc INNER JOIN
dbo_tls201_appln ON
dbo_tls209_appln_ipc.appln_id = dbo_tls201_appln.appln_id)
WHERE ((dbo_tls209_appln_ipc.ipc_class_symbol)="B60K 6/00")
AND ((dbo_tls201_appln.appln_auth)="EP");
The query in MS Access "Design View"

Result of query: 1,411

Show inventor and applicant details

- Inventor and applicant names are in the same table: `dbo_tls206_person`
- Table `dbo_tls207_pers_appln` will give us roles:
  - `applt_seq_nr`
  - `invnt_seq_nr`
- We need to display these values to be able to work with names
### How to "Join" the two applicant/inventor tables to the previous query

One join is done on surrogate key `appln_id` from table `dbo_tls201_appln` to `dbo_tls207_pers_appln`.

The other is done on surrogate key `person_id` from table `dbo_tls206_person` to `dbo_tls207_person`.

#### The query in SQL

```sql
SELECT dbo_tls209_appln_ipc.ipc_class_symbol, 
       dbo_tls201_appln.appln_auth, 
       dbo_tls201_appln.appln_nr, 
       dbo_tls201_appln.appln_kind, 
       dbo_tls201_appln.appln_filing_date, 
       dbo_tls207_pers_appln.applt_seq_nr, 
       dbo_tls207_pers_appln.invt_seq_nr, 
       dbo_tls206_person.person_ctry_code, 
       dbo_tls206_person.person_name  
FROM (dbo_tls209_appln_ipc INNER JOIN dbo_tls201_appln ON 
       dbo_tls209_appln_ipc.appln_id = dbo_tls201_appln.appln_id) 
INNER JOIN (dbo_tls206_person INNER JOIN dbo_tls207_pers_appln ON 
            dbo_tls206_person.person_id = dbo_tls207_pers_appln.person_id) ON 
            dbo_tls201_appln.appln_id = dbo_tls207_pers_appln.appln_id  
WHERE (((dbo_tls209_appln_ipc.ipc_class_symbol)="B60K 6/00") AND  
       ((dbo_tls201_appln.appln_auth)="EP");
```
The query in MS Access "Design View"

Result of query: 5,380

Result lines: can be exported (e.g. MS Excel)
Have a look at the data!

PCT application data format

- PCT application data as published (inid 21): PCT/EP2004/051252, PCT/IT02/00396, etc.

- This number is stored in PATSTAT ("DODCB format"):
  - country code (appln_auth): EP (or IT, ...)
  - application number (appln_nr): 2004051253
  - application kind code (appln_kind): W
**Pitfalls**

EP = Application authority contains both:
- EP Patents and
- WO patents filed with the EPO

Check:
Kind code or publication authority

**Kind code**

Total 5356

W : 448

A : 4908

**Pitfalls**

No double counting of inventions due to multiple applicants/inventors

Count only: first applicant

**Kind code**

Total 5356

W : 448

A : 4908

**Kind code (cleaned)**

Total 1396

W : 127

A : 1269
Case study B60K6/00 and application EP

Country of Applicant

Country of applicant assigned to first applicant

- JP: 663
- US: 152
- DE: 79
- FR: 32
- IT: 25
- GB: 17
- SE: 75
- Others: 226

Distribution of applications over the years - absolute

- JP
- US
- DE
- FR
- IT
- GB
- SE
- Others
Case study B60K6/00 and application EP

Distribution of applications over the years - relative

Country of Applicant - more precisely

According to first applicant (inner circle)

Weighted applicant (outer circle)

<table>
<thead>
<tr>
<th>Country</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>JP</td>
<td>663</td>
</tr>
<tr>
<td>US</td>
<td>655.5</td>
</tr>
<tr>
<td>DE</td>
<td>224.5</td>
</tr>
<tr>
<td>FR</td>
<td>226</td>
</tr>
<tr>
<td>IT</td>
<td>152</td>
</tr>
<tr>
<td>GB</td>
<td>152.5</td>
</tr>
<tr>
<td>SE</td>
<td>75</td>
</tr>
<tr>
<td>Others</td>
<td>79</td>
</tr>
</tbody>
</table>

10% difference!
### Weighted Applicant/Inventor

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>applicant</td>
<td>appln</td>
<td>inv</td>
<td>inv2</td>
<td>person</td>
<td>city</td>
<td>code</td>
<td>person</td>
<td>name</td>
<td>weighted appln</td>
</tr>
<tr>
<td>3375</td>
<td>EP 0020088</td>
<td>A</td>
<td>0</td>
<td>1</td>
<td>JP</td>
<td>Toyota</td>
<td>0</td>
<td>0</td>
<td>10% difference!</td>
<td></td>
</tr>
<tr>
<td>3376</td>
<td>EP 0020088</td>
<td>A</td>
<td>0</td>
<td>2</td>
<td>FR</td>
<td>Kanger</td>
<td>0</td>
<td>0</td>
<td>10% difference!</td>
<td></td>
</tr>
<tr>
<td>3377</td>
<td>EP 0020088</td>
<td>A</td>
<td>0</td>
<td>3</td>
<td>FR</td>
<td>Puppin</td>
<td>0</td>
<td>0</td>
<td>10% difference!</td>
<td></td>
</tr>
<tr>
<td>3378</td>
<td>EP 0020088</td>
<td>A</td>
<td>0</td>
<td>4</td>
<td>JP</td>
<td>Linzha</td>
<td>0</td>
<td>0</td>
<td>10% difference!</td>
<td></td>
</tr>
<tr>
<td>3379</td>
<td>EP 0020088</td>
<td>A</td>
<td>0</td>
<td>5</td>
<td>JP</td>
<td>Yamauchi</td>
<td>0</td>
<td>0</td>
<td>10% difference!</td>
<td></td>
</tr>
<tr>
<td>3380</td>
<td>EP 0020088</td>
<td>A</td>
<td>0</td>
<td>6</td>
<td>JP</td>
<td>Takahiro</td>
<td>0</td>
<td>0</td>
<td>10% difference!</td>
<td></td>
</tr>
<tr>
<td>3381</td>
<td>EP 0020088</td>
<td>A</td>
<td>0</td>
<td>7</td>
<td>JP</td>
<td>NISSAN MOTOR COMPANY LIMITED</td>
<td>0.5</td>
<td>0</td>
<td>10% difference!</td>
<td></td>
</tr>
<tr>
<td>3382</td>
<td>EP 0020088</td>
<td>A</td>
<td>0</td>
<td>8</td>
<td>FR</td>
<td>Renaults</td>
<td>0.5</td>
<td>0</td>
<td>10% difference!</td>
<td></td>
</tr>
</tbody>
</table>

### Country of Inventor

According to first inventor (inner circle) and Weighted inventor (outer circle) with 10% difference.
Comparison of country - Applicant vs Inventor

Technology acquisition or inventors from non domestic subsidiaries file through headquarter

Number of applicants and inventors per application

Case study B60K6/00 and application EP
International co-operation

Weighted EP applications

with more than two applicants from different countries

Thank you for your attention

For more information please contact
dlingua@epo.org
jschaaf@epo.org