

Abstract

A social network consists of a set of objects connected to each other by social relations. The Swiss Society for Industrial and Applied Mathematics social network contains some of the members of Swiss institutions; they are connected with each other if they have a collaboration in computational science conferences. Some algorithms are used to analyse the Swiss scientific social network: PageRank studies the relationships between institutions members and finds out the most relevant authors and organizations in the area of computational science.

Information Retrieval

1. Crawl online programs of PASC conferences and of SIAM (Parallel Computing, Uncertainty Quantification, Optimization, ...). To complete this task I used Apache Nutch, highly extensible and scalable open source web crawler
2. Extract name, institution and nation for each author
3. a) Create one node for each author and an edge for each collaboration
b) Create one node for each institution and an edge for the collaborations of the organizations members
4. From the nodes and the edges build the adjacency matrix of the social network

Dataset



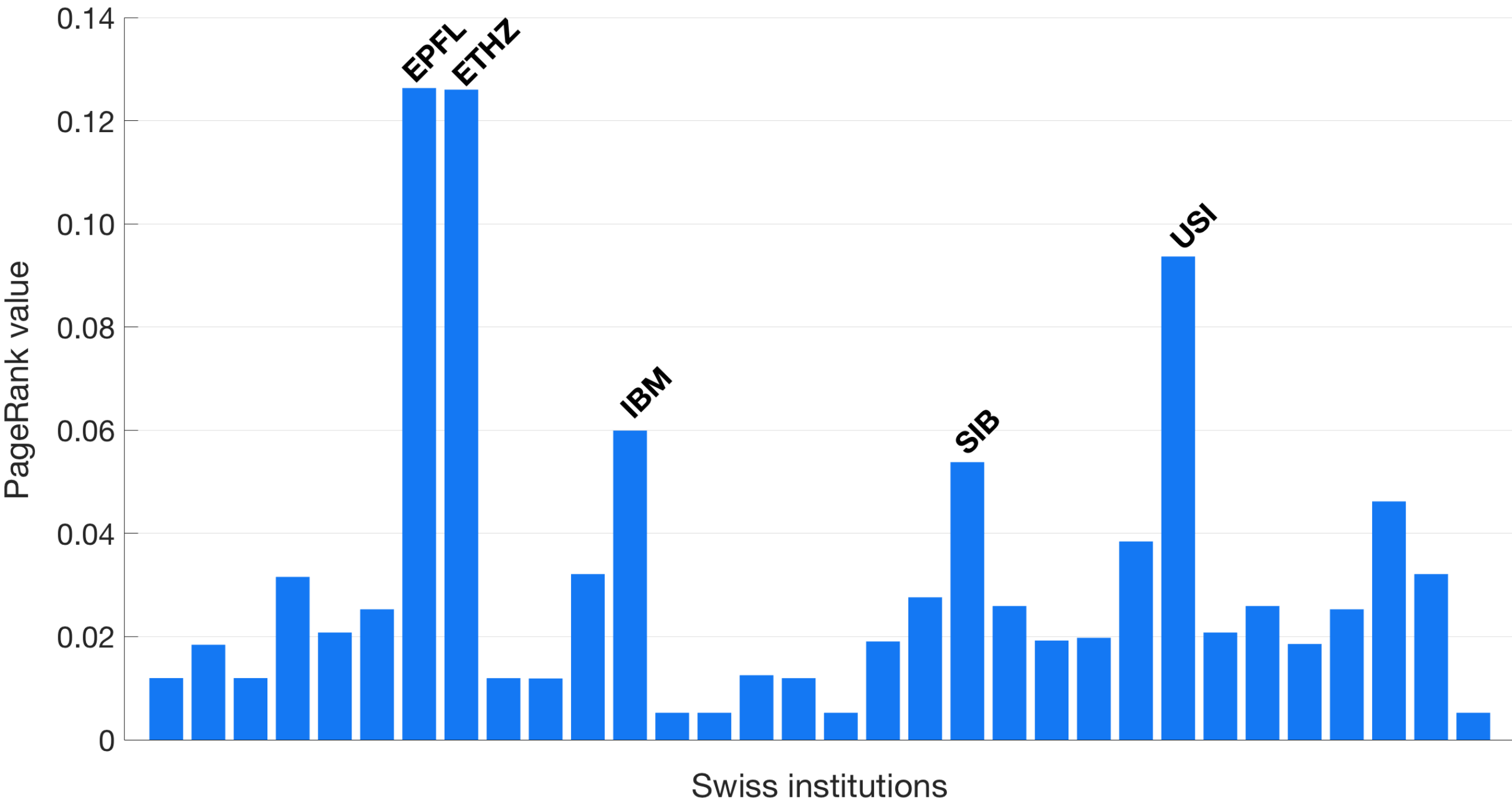
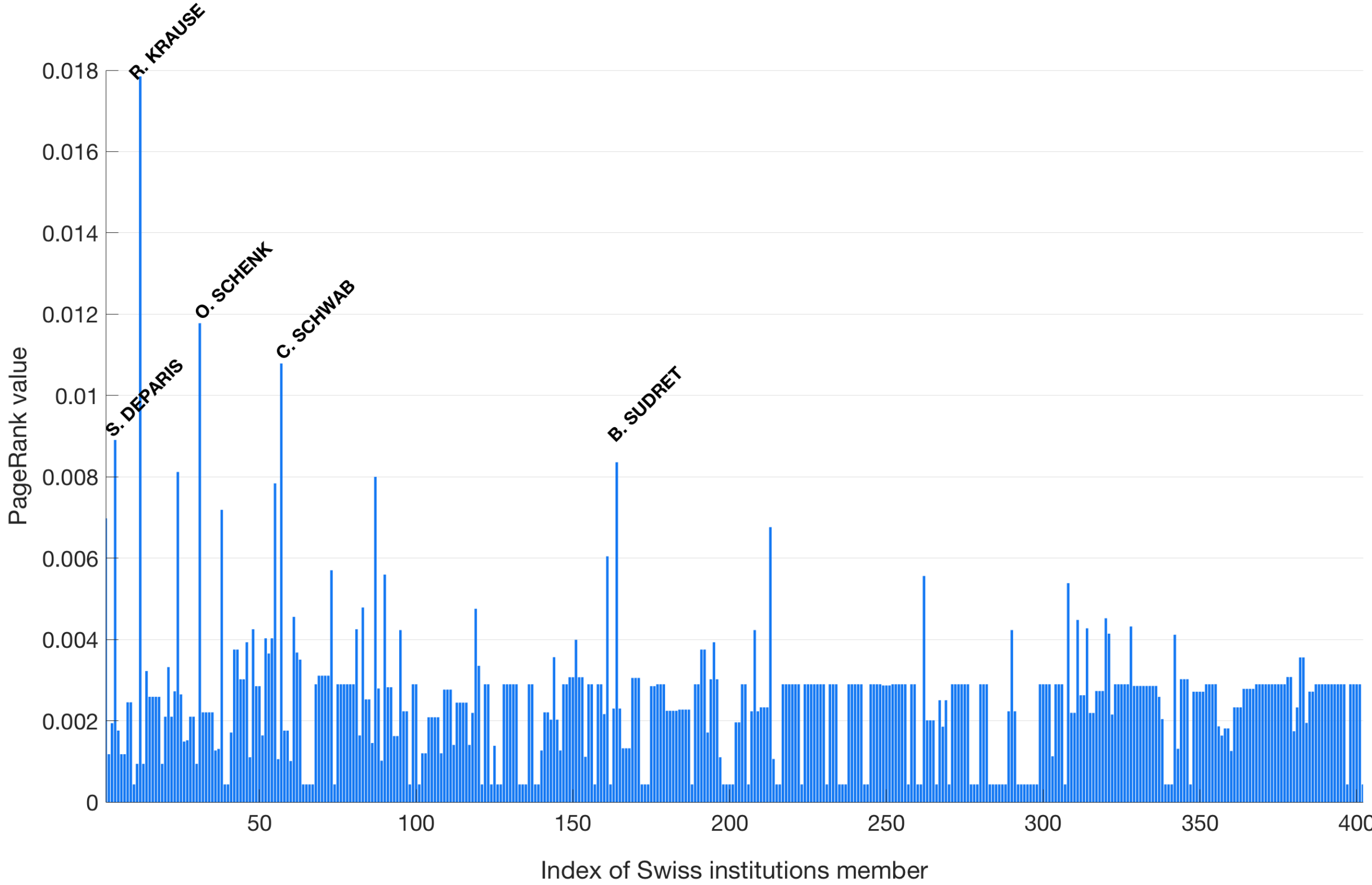
The PageRank algorithm

Algorithm – PageRank

```
1: procedure PAGERANK(authors list, adjacency matrix)
2:    $U \leftarrow$  authors list
3:    $G \leftarrow$  adjacency matrix
4:    $p \leftarrow$  dumping factor
5:    $c \leftarrow \text{sum}(G,1)$ ,  $r \leftarrow \text{sum}(G,2)$ 
6:    $D \leftarrow \text{diag}(1./c)$ 
7:   Solve system  $(I - pGD)x = e$ 
8:    $x \leftarrow x/\text{sum}(x)$ 
9:    $x \leftarrow \text{sort}(x, \text{'descending'})$ 
10: return x
```

Index	PageRank	Collaborations	Name, Institution
12	0.0178	30	R. Krause, USI
31	0.0117	12	O. Schenk, USI
57	0.0107	12	C. Schwab, ETHZ
4	0.0089	13	S. Deparis, EPFL
164	0.0083	8	B. Sudret, ETHZ
24	0.0081	12	P. Arbenz, ETHZ
87	0.0079	7	F. Nobile, EPFL
55	0.0078	11	C. Bekas, IBM
38	0.0071	7	R. Zenklusen, EPFL
1	0.0069	8	A. Quarteroni, EPFL

Index	PageRank	Collaborations	Institution
7	0.1263	16	EPFL
8	0.1260	16	ETHZ
25	0.0936	12	USI
12	0.0599	7	IBM
20	0.0538	7	SIB



Conclusion*

After collecting data I managed to construct and analyse, implementing the PageRank algorithm, authors and institutions graphs. The results provide an interesting picture of the different organizations in Switzerland and how they interact with each other.
* The original goal was to crawl published journal papers but for technical problems I took data from scientific conferences based on unreviewed abstracts submissions, so all the results are questionable.