

# DETECTION OF THE TYPE OF PUBLICATION FOR THE URAP SYSTEM

## 1 Motivation

The automated detection of the type of publication is not yet possible in URAP system.

However, even manually there are algorithmic steps that can be followed to identify the type.

The aim is to implement a code that helps the user of URAP to detect the data that are tedious.

The main target is to implement a code that works for computer science and provide it as open-source.

The code should be extendable for other domains than computer science.

## 2 Basic information

The publications can be of the following types:

1. journal paper
2. conference paper
3. workshop paper
4. book chapter
5. book
6. poster
7. others

According to the national rules (CNATDCU/CNFIS), the publications can be classified according to the publication venue in (ISI means WOS in the next sections):

1. NATURE
2. ISI ROSU
3. ISI GALBEN
4. ISI ALB
5. ISI ANH
6. ISI ESCI

7. ERIH INT1
8. ERIH INT2
9. ERIH PLUS
10. ISI PROC
11. IEEE PROC
12. BREVET INT
13. BREVET NAT
14. unclassified

The CNATDCU commission for Computer Science, classifies the publication in 5 categories:

1. A\*
2. A
3. B
4. C
5. D
6. E

The information related to the three categories mentioned above and marked with boxes is provided manually in the local URAP system (marked in boxes in Figure ??).

Research Activity Platform

Caută în citări

colò M. Calcavechia, Bogdan A. Caprarescu, Elisabetta Di Nitto, Daniel J. Dubois, Dana Petcu. DEPAS: a decentralized probabilistic algorithm for auto-scaling. Computing, 2012/06/22

Articol

An

Încadrare

CINATDCU

Încadrare

Domeniu

<div> <div></div> <div> <div>Towards Bio-inspired Auto-Scaling Algorithms: An Elasticity Approach for Container Orchestration Platforms</div> <div>Jose Herrera, German Molito</div> <div>IEEE Access</div> </div> </div>	2020	ISI	A*
<div> <div></div> <div> <div>Cloud-based machine learning architecture for big data analysis</div> <div>R Pakdel</div> <div>PhD Thesis</div> </div> </div>	2019		D
<div> <div></div> <div> <div>Design and Evaluation of Decentralized Scaling Mechanisms for Stream Processing</div> <div>Mehdi Mokhtar Belkhiria, Cedric Tedeschi</div> <div>2019 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)</div> </div> </div>	2019	IEEE	C
<div> <div></div> <div> <div>Autonomic Decentralized Microservices: The Gru Approach and Its Evaluation</div> <div>Elisabetta Di Nitto, Luca Florio, Damian A. Tamburri</div> <div>Microservices</div> </div> </div>	2019		D
<div> <div></div> <div> <div>Elasticity in Cloud Computing: State of the Art and Research Challenges</div> <div>Yahya Al-Dhuraibi, Fawaz Paraiso, Nabil Djarallah, Philippe Merle</div> <div>IEEE Transactions on Services Computing</div> </div> </div>	2018	ISI	A
<div> <div></div> <div> <div>Auto-Scaling Web Applications in Clouds</div> <div>Chenhao Gu, Rodrigo N. Calheiros, Rajkumar Buyya</div> <div>ACM Computing Surveys</div> </div> </div>	2018	ISI	A*
<div> <div></div> <div> <div>Social Auto-Scaling</div> <div>Peter Smith, Horacio Gonzalez-Velez, Simon Caton</div> <div>2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)</div> </div> </div>	2018	ISI	C
<div> <div></div> <div> <div>A self-scalable distributed network simulation environment based on cloud computing</div> <div>Sergio Serrano-Iglesias, Eduardo Gómez-Sánchez, Miguel L. Bote-Lorenzo, Juan I. Asensio-Pérez, Manuel Rodríguez-Cayetano</div> <div>Cluster Computing</div> </div> </div>	2018	ISI	B
<div> <div></div> <div> <div>Workload characterization and autoscaling in Cloud environments</div> <div>Momin I.M. Tabash</div> <div>PhD Thesis</div> </div> </div>	2017		D
<div> <div></div> <div> <div>A Runtime Auto-Scaling Mechanism for Multi-Tenant Software</div> <div>A. Ghafori, M. Sharifi</div> <div>Journal Of Electronical and Cyber Defence</div> </div> </div>	2017		D
<div> <div></div> <div> <div>Provisionamento Automático de Recursos em Nuvem IaaS: eficiência e limitações de abordagens reativas</div> <div>F. Moraes, R. Lopes, F. Brasileiro</div> <div>C Grande-PB-Brasil, 35th The Brazilian Symposium on Computer Networks and Distributed Systems (SBRC)</div> </div> </div>	2017		D
<div> <div></div> <div> <div>SHDF-A Scalable Hierarchical Distributed Framework for Data Centre Management</div> <div>Abdul Rahman Hummaida, Norman W Paton, Rizos Sakellariou</div> <div>2017 16th International Symposium on Parallel and Distributed Computing (ISPDC)</div> </div> </div>	2017	ISI	C
<div> <div></div> <div> <div>Online machine learning for cloud resource provisioning of microservice backend systems</div> <div>Haniieh Alipour, Yan Liu</div> <div>2017 IEEE International Conference on Big Data (Big Data)</div> </div> </div>	2017	IEEE	D
<div> <div></div> <div> <div>A statistical approach to virtual server resource management</div> <div>Dimitris Kontoudis, Panayotis Fouliras</div> <div>Concurrency and Computation: Practice and Experience</div> </div> </div>	2017	ISI	C
<div> <div></div> <div> <div>Scaling framework for querving</div> <div>Mohit SIKRI</div> <div>US Patents</div> </div> </div>	2016		D
<div> <div></div> <div> <div>Comprehensive formalization and modeling of the network and system virtualization process</div> <div>Kontoudis, Dimitrios</div> <div>PhD thesis</div> </div> </div>	2016		D
<div> <div></div> <div> <div>Resource Provisioning for Web Applications under Time-varying Traffic</div> <div>Ali Rajabi</div> <div>PhD thesis</div> </div> </div>	2016		D
<div> <div></div> <div> <div>Auto-scaling and deployment of web applications in distributed computing clouds</div> <div>Chenhao Gu</div> <div>PhD thesis</div> </div> </div>	2016		D
<div> <div></div> <div> <div>Functional Programming Languages in Computing Clouds</div> <div></div> <div></div> </div> </div>	2016		D

DOI articol

10.1109/tsc.2017.2711009

Introduceți DOI articol și apăsați butonul Căutare din dreapta.

WOS Accession Number

WOS:000429798800017

Tip publicație

Jurnal

Titlu articol \*

Elasticity in Cloud Computing: State of the Art and Research Challenges

Autori articol

Yahya Al-Dhuraibi, Fawaz Paraiso, Nabil Djarallah, Philippe Merle

Denumire Jurnal/volum

IEEE Transactions on Services Computing

Editori jurnal/volum

Număr volum

11

Număr apariție

2

Cod ISSN Online

Cod ISSN Print

1939-1374

Cod ISBN

Data publicării (AAAA/LL/ZZ) \*

2018/03/01

Clasificare

A

Clasificare articol CINATDCU

ISI ROSU

Figure 1: Snapshot from URAP system

### 3 Input data from the user

1. Digital Object Identifier (doi) of the publication
2. Answer to the question if it is computed for Computer Science with yes/no

### 4 Files

1. SCIE classification of the journals for each year
2. SSCI classification of the journals for each year
3. CORE for several years (for computer science conferences)
4. SENSE for book categories (for computer science) for all years

Notes related to the usage of SCIE and SSCI:

- a journal can be found either in SCIE either in SSCI, but can appear in the same file several times in different categories
- beyond the issn field, the journalImpactFactor and articleInfluenceScore fields are important
- in the following algorithm the ‘best’ position of the journal is identified, being relative to the position in the category according to the sorting on decreasing order of the impact factor or article influence score (best of the two classification and best in different categories).
- the first 25% journals in a category are ‘ISI ROSU’, next 25% are ‘ISI YELLOW’, the rest from SCE si SSCI are ‘ISI ALB’;
- the first 20% journals from the ‘ISI ROSU’ are ‘A\*’ (let x be their number), the rest of ‘ISI ROSU’ + next x are ‘A’, the rest of ‘ISI GALBEN’ + next x are ‘B’, the rest of ‘ISI ALB’ are ‘C’

The files are available in a ZIP file (password: master; accesible only with e-uvt.ro address, not gmail or others) available at:

<https://drive.google.com/file/d/1C2wVP9GVUHqP2Zxlo6i3c3FLLFkeDvZq/view?usp=sharing>

Other files in the same ZIP file:

- More details about the 5 categories for Computer Science are in the file CS-Criteria (in Romanian) – only for supplementary information.
- DOIs for testing are available in the file Test-Data

### 5 Output data

1. Type of publication
2. CNATDCU/CNFIS classification
3. If is in WOS then WOS number
4. If INFO then INFO classification

## 6 (Minimal) Algorithm

The following algorithm identifies NATURE, ISI ROSU, ISI GALBEN, ISI ALB, ISI ESCI, ISI PROC, IEEE PROC and INFO classification. Note that curl works under Linux.

```
read doi in form x/y
read boolean INFO

if curl -s -I 'http://ws.isiknowledge.com/cps/openurl/service?url_ver=Z39.88-2004&rft_id=info:doi/x/y' | grep WOS:
  returns a line
  then
    isInWoS=true
    WOS= id of 16 digits from line after 'WOS:'
  esle
    isInWoS=false

get output-text from
  curl -L -H 'Accept: application/citeproc+json, application/unixref+xml' 'http://dx.doi.org/x/y'
extract from output-text
  extract year
  extract type
  switch (type)
    'article-journal':
      extract ISSN
      extract container-title
    'paper-conference':
      extract event
      extraxt acronym from event /*in brackets*/
      extract event_title from event /*without brackets and without numbers or year*/
    'chapter':
    'book':
      extract publisher
      extract publisher-location

if type is 'article-journal'
  classCNATDCU=' '; classINFO='D'
  repeat
    search ISSN in the files SCIE and SSCI corresponding to the year of the publication or the closest year
    /* ISSN can be in present in several categories in one file */
    if ISSN is found then
      extract journalImpactFactor, articleInfluenceScore, categoryName
      let n be the number oj journals from the file within the same categoryName
      load a matrix n x 3 with all journals with the same categoryName
        with each line having (issn, journalImpactFactor,articleInfluenceScore)
```

```

    sort the lines of the matrix according decreasing journalImpactFactor

    /*marker*/
    let rank be the journal's line in the sorted matrix line
    if rank<=ceil(0.25*n) then
        classCNATDCU='ISI ROSU'
        if container-title includes NATURE then classCNATDCU='NATURE'
    else
        if rank<=ceil(0.5*n) then
            if not classCNATDCU='ISI ROSU' then classCNATDCU='ISI GALBEN'
        else
            if not classCNATDCU='ISI ROSU' or 'ISI GALBEN' then classCNATDCU='ISI ALB'
    if INFO then
        x=floor(0.2*ceil(0.25*n))
        if rank<=x then classINFO='A*'
        else if rank<=ceil(0.25*n)+x then
            if classINFO is not 'A*' then classINFO='A'
        else
            if rank<=ceil(0.5*n)+x then
                if classINFO is not 'A*' or 'A' then classINFO='B'
            else
                if classINFO is not 'A*' or 'A' or 'B' then classINFO='C'

    /*end marker*/

    if articleInfluenceScore is positive then
        sort the lines of the matrix according decreasing articleInfluenceScore
        repeat /* from market to end marker*/
until end of both files SCIE and SSCI
if classCNATDCU=' ' and isInWoS then
    classCNATDCU='ISI ESCI'
if INFO and classINFO='D'
    if https://plu.mx/plum/a/?doi=x/y returns a page /*in Scopus */
    then
        classINFO='C'

if type is 'paper-conference'
if isInWoS then clasificCNATDCU='ISI PROC'
    else if x=10.1109 then clasificCNATDCU='IEEE PROC'
    else clasificCNATDCU=' '

if INFO
    classINFO='D'
    while not end of file CORE with closest year and not found
        search exact acronym in CORE with closest year

```

```
        if found
            if match of event_title in 75% then
                classINFO=class from CORE (A*, A, B or C)

if INFO and (type is 'chapter' or 'book')
    while not end of file SENSE and not found
        search publisher & publisher-location in SENSE with exact or closest year
        if found then classINFO=class from SENSE (A, B, C, D or E)

print type
print clasificCNATDCU
if isInWoS then print WOS
if INFO print clasificINFO
```