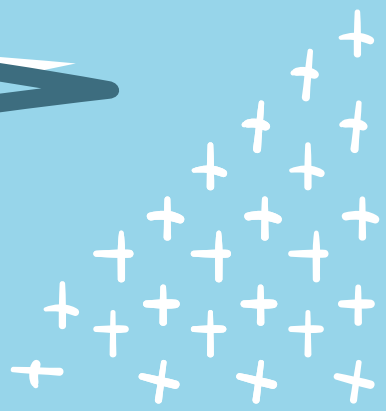
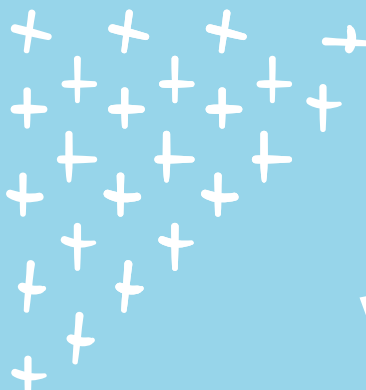
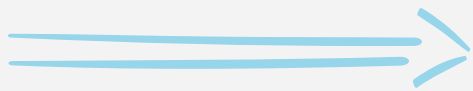




Quality monitoring in Wikipedia: A computational perspective

Animesh Mukherjee

IIT Kharagpur



Our efforts

WikiM (JCDL 2017)

WikiRef (COLING 2018)

StRE (ACL 2019)

NwQM (EMNLP 2020)

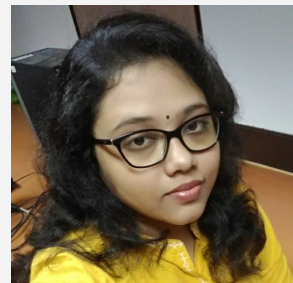
When expertise gone missing (ICADL 2021)

Quality Change: norm or exception? (CSCW 2022)

Soumya



Paramita



Sasi



Bhanu



table of contents

01

Missing Wikipedians

Who are they?

02

Missing Wikipedians

How to detect?

03

Quality change

What are the indicators?

04

Quality change

How to detect?



-01-
Missing
Wikipedians

Who are they?

Who are they?

Prolific Wikipedians: > 1K edits

Active Wikipedians: Prolific Wikipedians actively contributing

Missing Wikipedians ([WP:MISS](#)): Prolific Wikipedians inactive for ≥ 3 months

-A missing Wikipedian-

Contents [hide]

- 1 Retired
- 2 Licensing
- 3 Timeline
- 4 Contact Info
- 5 Pictures I've taken

Retired

I have left Wikipedia. I do not see it as acceptable to have advertisements, whether they be for brand identity or for a product, on Wikipedia. It appears that this line has [1] been breached, and those involved, including several on the foundation board and employees, many of whom I respect, see things differently. I understand their position, but cannot accept it as the future of a project that I will continue to contribute my time to. **Do not leave notes on my talk page.** Email me if you want to keep in touch.

(I am not back, but apparently the mailing list software renumbered those messages some months after I left, and I sometimes get questions about this and prefer to just point them at my userpage, so I'm correcting them below and adding some context...)



Missing Wikipedians

- No edits in the calendar year 2020
- #1146 extracted from WP:MISS

Active Wikipedians

- Similar activity levels as missing Wikipedians
 - Still editing (pages that had been once co-edited by missing Wikipedians)
 - #2569 extracted
- 



-02- Missing Wikipedians

How to detect those at risk?

Research Questions

-Hypothesis-

Early signs of retirement → last trail of their activities

-RQ1-

Missing vs Active?

Activity

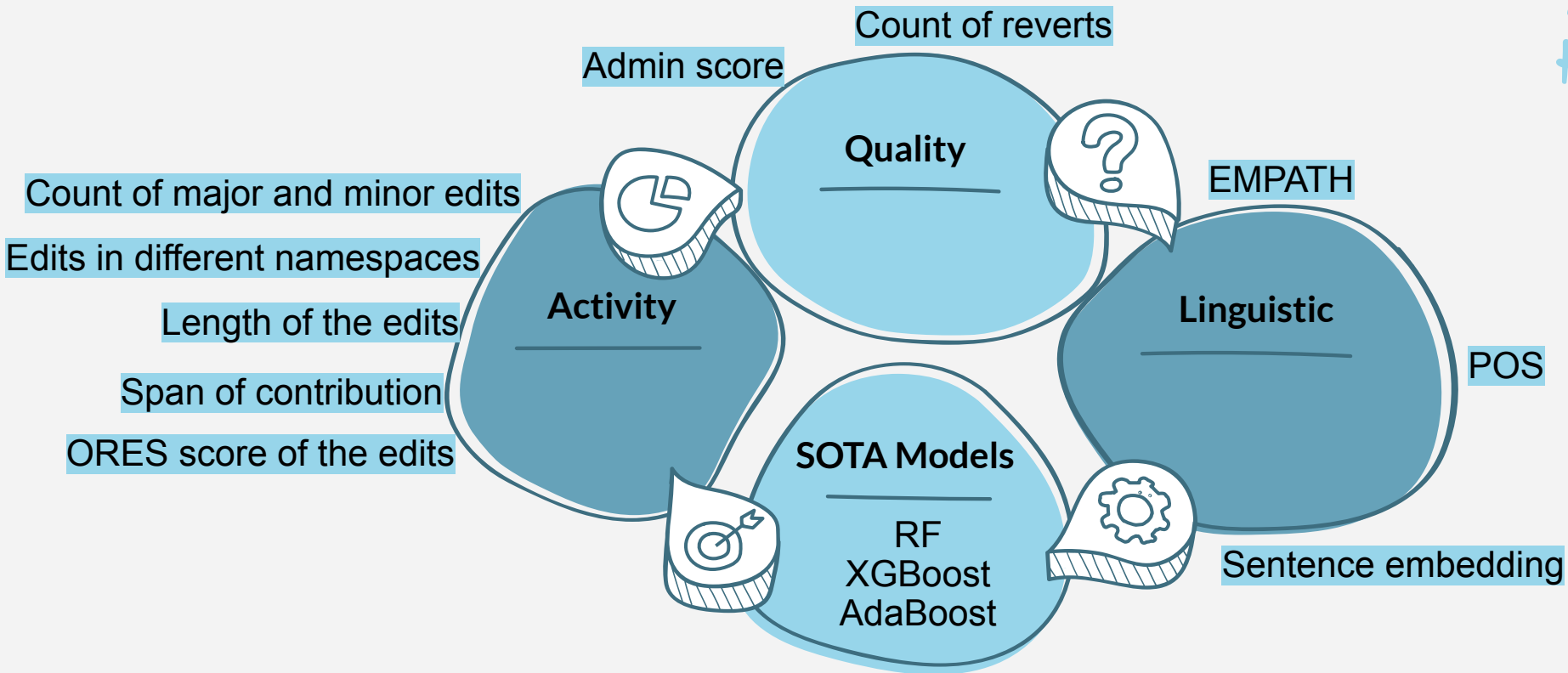
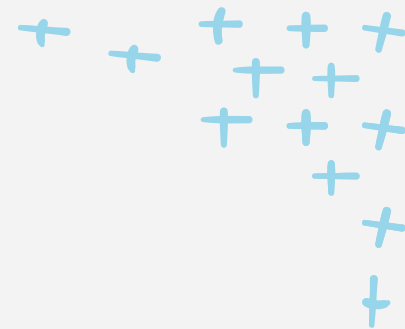
Linguistic

Quality

-RQ2-

Predict: active Wikipedians at a
potential risk to leave the platform soon?

Feature engineering



Prediction outcomes

Features	Classifier	Precision	Recall	F-score	Accuracy
G1	XGBoost	0.74	0.74	0.74	0.75
G2	XGBoost	0.63	0.68	0.64	0.68
G3	AdaBoost	0.63	0.67	0.63	0.67
G1 $\oplus$ G2	Random Forest	0.77	0.78	0.76	0.78
G1 $\oplus$ G3	XGBoost	0.75	0.76	0.75	0.76
G1 $\oplus$ G4	AdaBoost	0.78	0.79	0.79	0.78
G1 $\oplus$ G5	AdaBoost	0.78	0.79	0.79	0.77
G1 $\oplus$ G2 $\oplus$ G4	XGBoost	0.77	0.78	0.77	0.78
G1 $\oplus$ G3 $\oplus$ G4	XGBoost	0.81	0.82	0.81	0.82
G1 $\oplus$ G4 $\oplus$ G5	AdaBoost	0.81	0.82	0.81	0.82
G1 $\oplus$ G3 $\oplus$ G5	XGBoost	0.78	0.79	0.78	0.79
G1 $\oplus$ G2 $\oplus$ G4 $\oplus$ G5	Random Forest	0.82	0.82	0.82	0.81
G1 $\oplus$ G3 $\oplus$ G4 $\oplus$ G5	XGBoost	0.80	0.81	0.80	0.81

Feature combination

Alias

Activity features

G1

POS Tags + Empath

G2

Sentence vector

G3

Admin score

G4

Revert count

G5

What do we learn?

So: we can indeed find Wikipedians at risk with 82% accuracy

Most predictive: Activity + Quality

Some appreciation received by the editor at the early stage:

- Thanks for you help in reverting
- I've noted you've had to deal with a lot of vandalism.
- Thanks for helping out with the cryptography categorization!

Several warnings for the editor before inactivity :

- The page appears to have no meaningful content or history, and the text is unsalvageably incoherent
- Do you remember you're an admin in C***? It seems that it won't last long....
- WHAT THE HELL? The B*** article is really stupid..

Comments showing disagreement with the community:

- I'm done with this place. It's now officially stupid.
- I understand their position, but cannot accept it as the future of a project that I will continue to contribute my time to. Do not leave notes on my talk page.
- Too many editors out there drinking the politically correct Kool-Aid and trying to make everybody worship at the altar of the Almighty Wikipedia Rule Book.



-03- Quality change

What are the indicators?

What is article quality?

An indication: how good an Wikipedia article is?

- Comprehensible
- Well-organized
- Readable
- Well-referenced

-Quality class-

FA, A, GA, B, C, Start, Stub

Talk:Social networking service

From Wikipedia, the free encyclopedia

This is the [talk page](#) for discussing improvements to the **Social networking service** article.
This is **not** a forum for general discussion of the article's subject.

- Put new text under old text. Click here to start a new topic. - Please sign and date your posts by typing four tildes (~~~~). - New to Wikipedia? Welcome! Ask questions, get answers.	- Be polite and welcoming to new users - Assume good faith - Avoid personal attacks - For disputes, seek dispute resolution	Article policies - No original research - Neutral point of view - Verifiability
--	---	--

Material from Social networking service was split to Issues relating to social networking services on 26 June 2020 from this version. The former page's history now serves to [provide attribution](#) for that content in the latter page, and it must not be deleted so long as the latter page exists. Please leave this template in place to link the article histories and preserve this attribution.

 **Social networking service** has been listed as a level-5 **vital article** in Technology. If you can improve it, please do. This article has been rated as B-Class.

 This article is of interest to the following **WikiProjects**: [hide]

WikiProject Computing	(Rated B-class, High-importance)	[show]
WikiProject Internet culture	(Rated B-class, Mid-importance)	[show]
WikiProject Sociology	(Rated B-class, Mid-importance)	[show]
WikiProject Internet	(Rated B-class, Low-importance)	[show]

Whatz the problem?

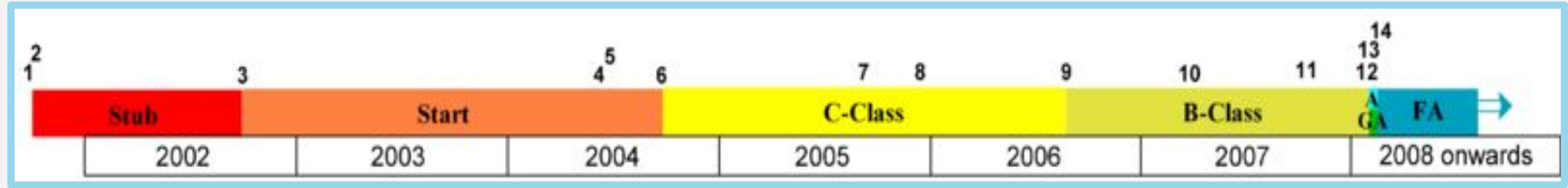
-Very few "good" quality articles-

-Close to 50% articles are start/stub-

All rated articles by quality and importance						
Quality	Importance					Total
	Top	High	Mid	Low	???	
★ FA	1,362	2,164	2,093	1,459	170	7,248
★ FL	160	595	662	594	104	2,115
ⓘ A	283	597	747	485	88	2,200
⊕ GA	2,638	6,038	12,049	14,358	1,732	36,815
B	14,058	27,145	43,147	41,692	16,992	143,034
C	13,764	42,391	101,294	173,468	66,097	397,014
Start	18,339	86,331	366,119	1,157,719	372,552	2,001,060
Stub	4,477	32,416	273,720	2,398,576	869,913	3,579,102
List	4,005	14,480	44,041	136,791	76,034	275,351
Assessed	59,086	212,157	843,872	3,925,142	1,403,682	6,443,939
Unassessed	117	522	2,060	16,401	435,250	454,350
Total	59,203	212,679	845,932	3,941,543	1,838,932	6,898,289

Lifecycle

https://en.wikipedia.org/wiki/Wikipedia:Content_assessment



-RQ1-

How do the Wikipedia articles transition through different quality states over time?

-RQ2-

How to detect the dynamic change in article quality?

Data

Quality class	Count
FA	3536
A	511
GA	5780
B	5335
C	4884
Start	5469
Stub	5321
Total	30826



Old Class	New Class	Count
FA	FA	3536
A, GA	AGA	6291
B, C	BC	10219
Start, Stub	SS	10780



Evolution - I

Type	Number of hops	Count	Avg time (in days)	SD (in days)
SS $\Rightarrow$ BC	1	8594	1253.70	1165.91
BC $\Rightarrow$ AGA	1	6144	390.58	728.73
AGA $\Rightarrow$ FA	1	2283	294.18	476.34
SS $\Rightarrow$ AGA	2	1384	1198.25	1206.46
BC $\Rightarrow$ FA	2	487	535.28	881.65
SS $\Rightarrow$ FA	3	97	1873.44	1380.87

-Only promotion-

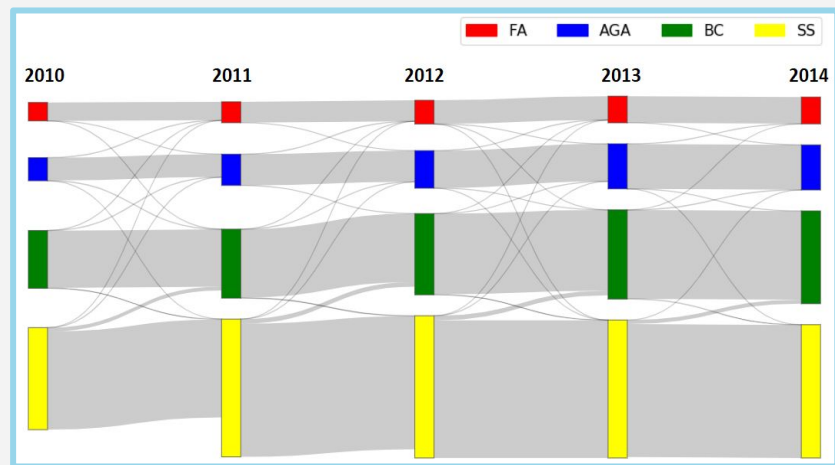
Type	Number of hops	Count	Avg time (in days)	SD (in days)
BC $\Rightarrow$ SS	1	83	542.86	655.51
AGA $\Rightarrow$ BC	1	105	400.33	401.32
FA $\Rightarrow$ AGA	1	2	469.88	294.94
AGA $\Rightarrow$ SS	2	1	21.92	0
FA $\Rightarrow$ BC	2	32	753.54	694.44
FA $\Rightarrow$ SS	3	0	0	0

-Only demotion-

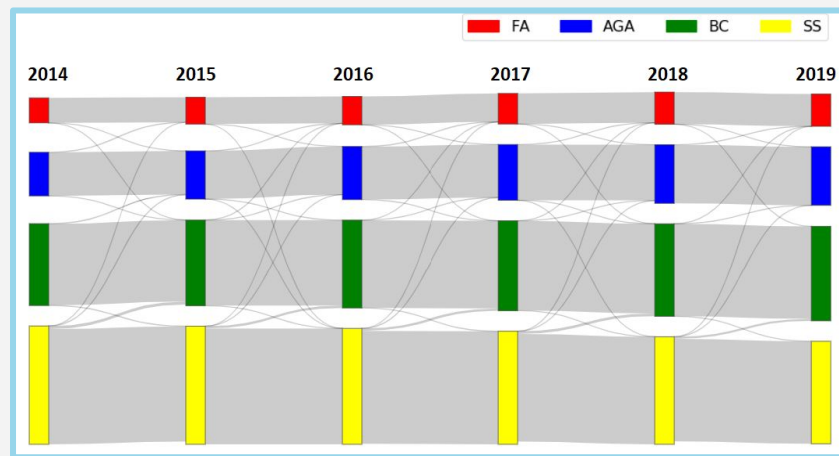
Quality Class	Count	Mean time (in days)	SD (in days)
FA	341	2039.14	1750.22
AGA	375	1165.19	1610.10
BC	4586	935.45	1075.24
SS	10647	316.02	581.87

-No change (51.73%)-

Evolution - II

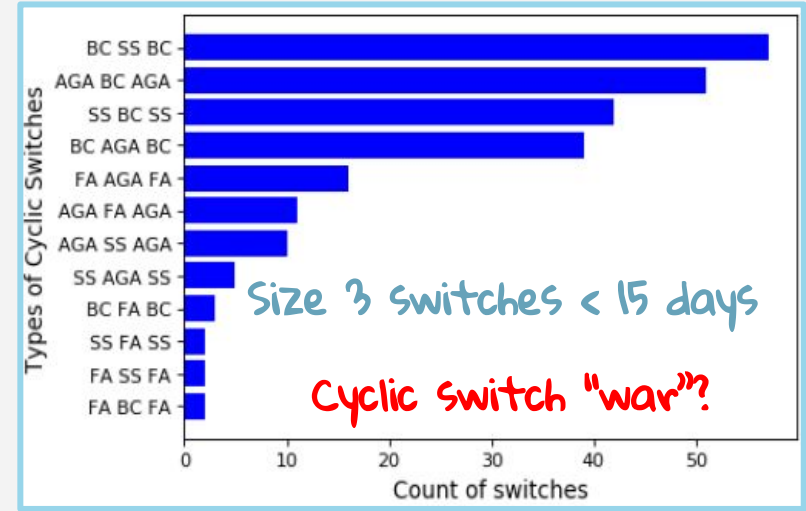
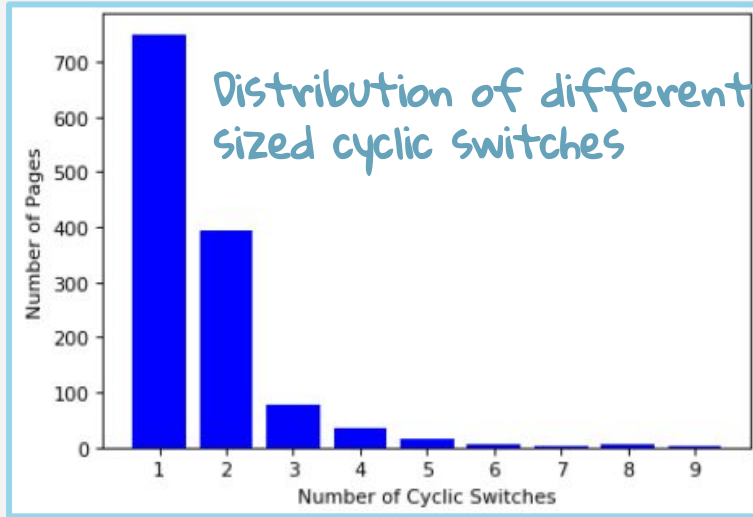


2010 - 2014



2014 - 2019

Cyclic switches

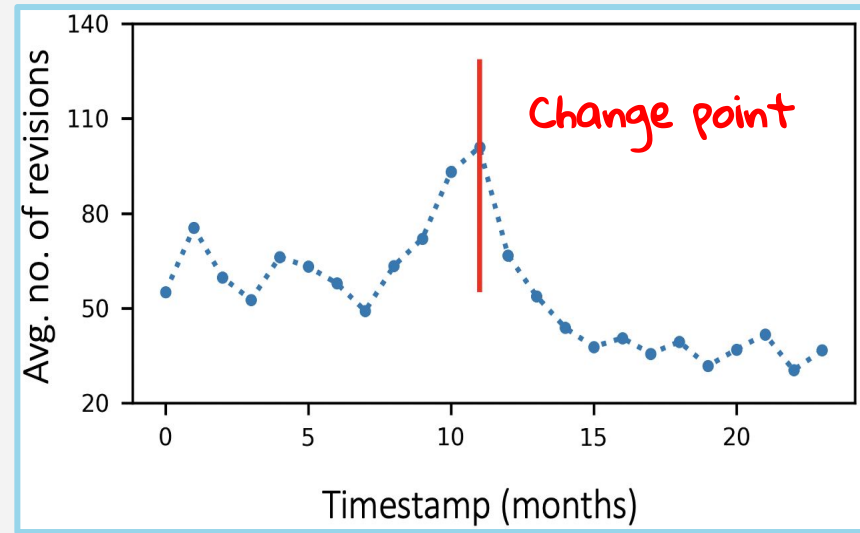
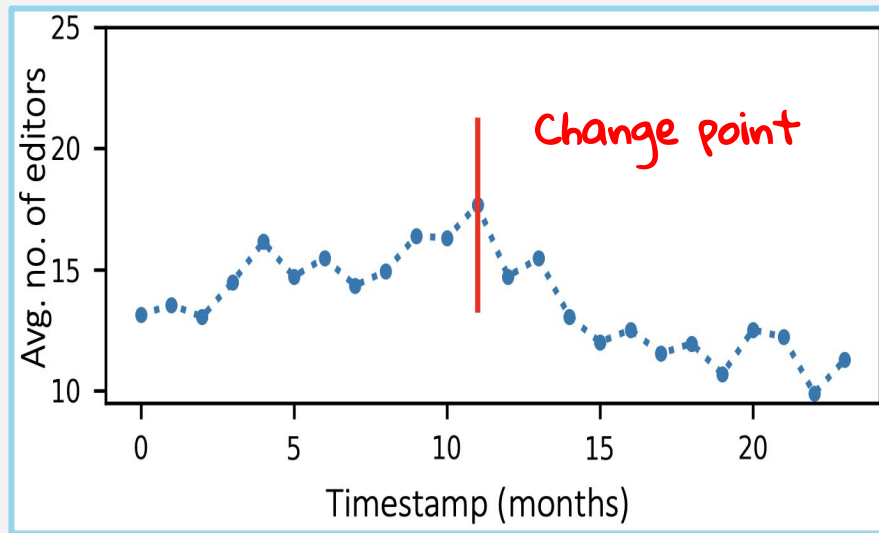




-04- Quality change

How to detect?

Change point detection (CPD)



Multivariate change point detection algorithms

Binary segmentation (BinSeg)
Pruned Exact Linear Time (PELT)
Non-parametric CPD (ECP)

Features

Contribution based (G_c)

#registered editors of talk page

#unregistered editors of talk page

#registered editors of article page

#unregistered editors of article page

Content based based (G_p)

article length in bytes

#references in the article

#hyperlinks in the article

#categories in the article text

#citation templates

#non-citation templates

?infobox

#images/article length

#level 2 headings

#level 3+ headings

Readability

noise

Activity based (G_a)

Mean/variance time between two consecutive revisions of talk pages

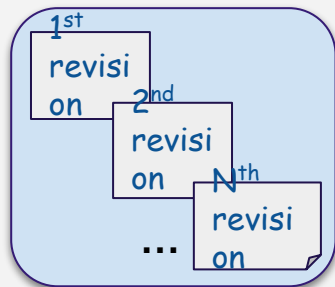
Mean/variance time between two consecutive revisions of article pages

#Revisions of the talk pages

#Revisions of the main pages

Time series features of revisions of a page

Framework



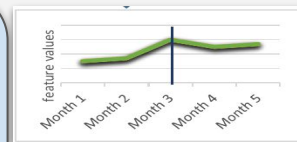
Revision history for page P (month wise)



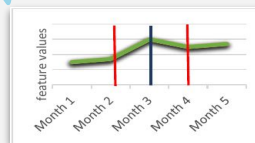
Contribution based (G_c)
Activity based (G_a)
Content based based (G_p)

Calculate Covering, Precision and Recall for page P

Representation of page P as time series



Change Point Detection



Blue line- ground truth
Red lines- predicted change points



Results

Features	BinSeg [n_bkps = 1]			ECP [min_size = 5]			PELT [pen_val = 1]			HYBRID		
	Covering	Precision	Recall	Covering	Precision	Recall	Covering	Precision	Recall	Covering	Precision	Recall
G_c	0.68	0.36	0.29	0.39	0.23	0.27	0.52	0.28	0.47	0.73	0.51	0.60
G_a	0.68	0.37	0.30	0.41	0.27	0.29	0.52	0.29	0.47	0.74	0.53	0.59
G_p	0.76	0.60	0.46	0.60	0.45	0.45	0.60	0.37	0.61	0.79	0.68	0.69
$G_c \oplus G_a$	0.68	0.38	0.30	0.42	0.27	0.31	0.53	0.29	0.45	0.74	0.53	0.59
$G_a \oplus G_p$	0.75	0.56	0.43	0.59	0.42	0.44	0.60	0.37	0.58	0.78	0.66	0.68
$G_p \oplus G_c$	0.74	0.53	0.41	0.56	0.39	0.43	0.60	0.37	0.57	0.78	0.64	0.60
$G_c \oplus G_a \oplus G_p$	0.74	0.53	0.41	0.56	0.39	0.43	0.60	0.37	0.57	0.78	0.64	0.67

#Articles	CPD HYBRID method			ORES		
	Covering	Precision	Recall	Covering	Precision	Recall
Set 1	0.79	0.68	0.69	0.56	0.31	0.60
Set 2	0.82	0.80	0.73	0.63	0.40	0.71

Set 1: at least one change point in the ground truth reality

Set 2: articles that got promoted to the FA class at least once

Key features

mean/ variance of time elapsed between two consecutive revisions of talk pages

#revisions of the talk pages

#revisions of article pages

presence of “difficult words” → ‘xenon’, ‘pipeline’, ‘anole’, ‘touchdown’, ‘epilepsy’, ‘carfilzomib’.

Thank You