

# Flagged revisions

**Latest results**

# Summary

- Goals
- Methodology
- Results
- Questions and feedback

# Goals

- Study the **impact of flagged revisions** in the editorial work of the **German Wikipedia**.
  - Focus on **anonymous** editors.
  - Focus on **vandalism** and reverts.
- Questions:
  - Is the vandalism from anonymous reduced?
  - Is the number of anonymous blocks reduced?
  - Does it discourage anonymous from editing?
  - Does it decrease number of requests for new accounts? (TODO)

# Methodology (overview)

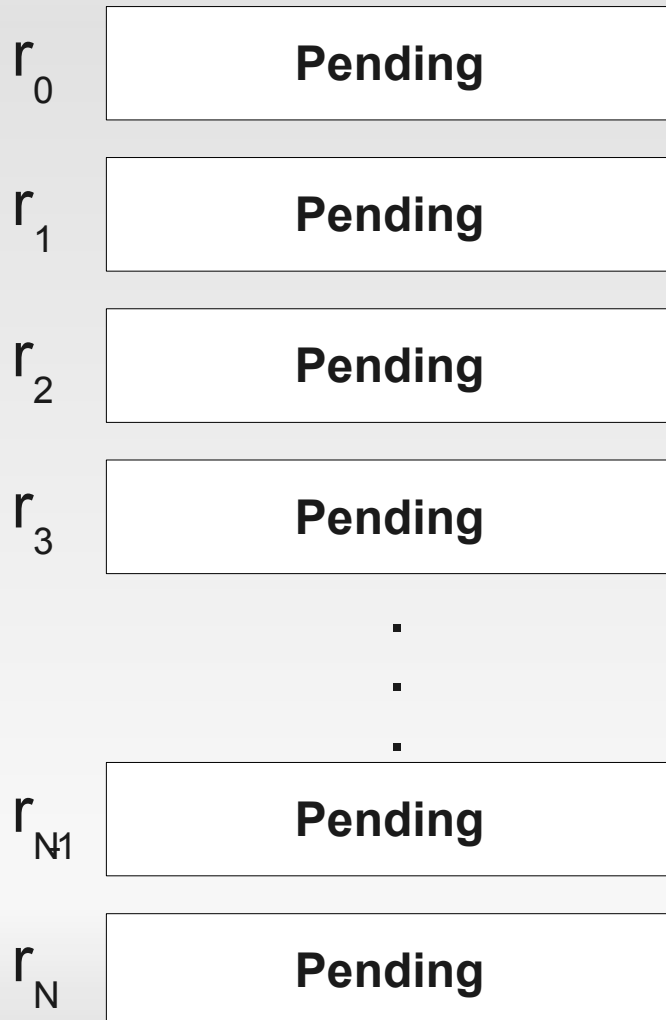
- Parse *page-logging.xml*
- EDA on the data set.
  - Survival analysis on time to review/revert an edit.
  - Numbers and time series:
    - Reverts
    - Blocks /IP-blocks
    - (Semi)-protection
    - **Sighted** anonymous edits
- Comparison with similar versions.
  - **Polish & Russian** (also with flagged-revs).
  - Also WKP comparable size (FR, IT)

# Methodology (concepts)

- We focus on anonymous contributions.
- Sighting actions (manual)
  - **Approved**: Revisions **manually** flagged as **OK**.
  - **Approved-i**: Introduced later to identify first approvals.
- Sighting actions (automated)
  - Approved-a, approved-ia.
  - Automated approval for trusted editors.
  - Filtered-out.

# Methodology (approvals)

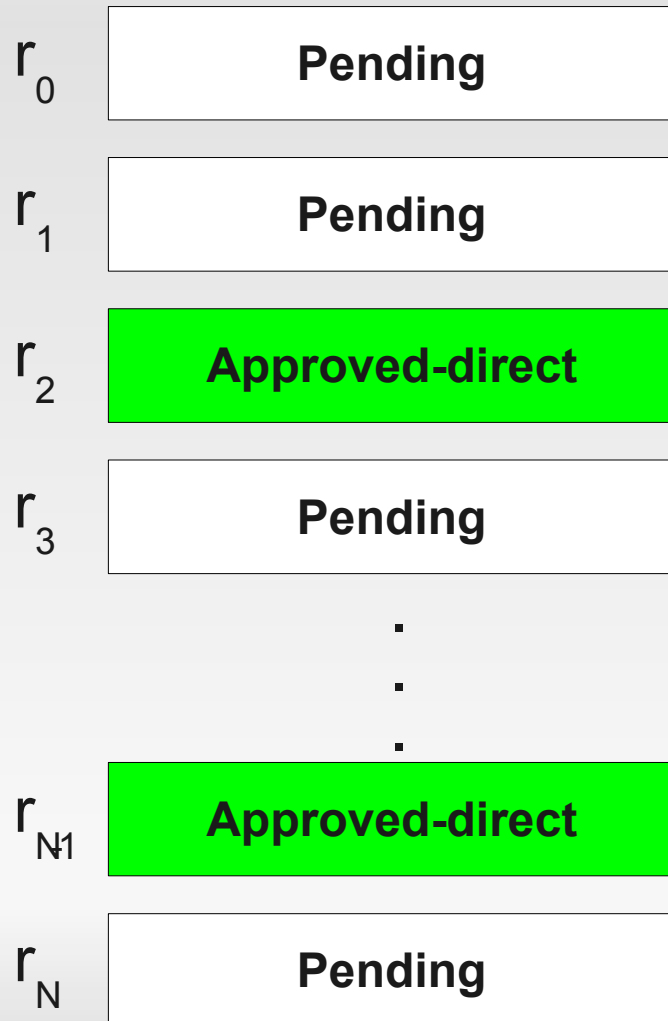
## SET OF CONSECUTIVE ANON REVS (page X)



At first, all revisions  
in a given series  
(same page)  
are *pending*

# Methodology (approvals)

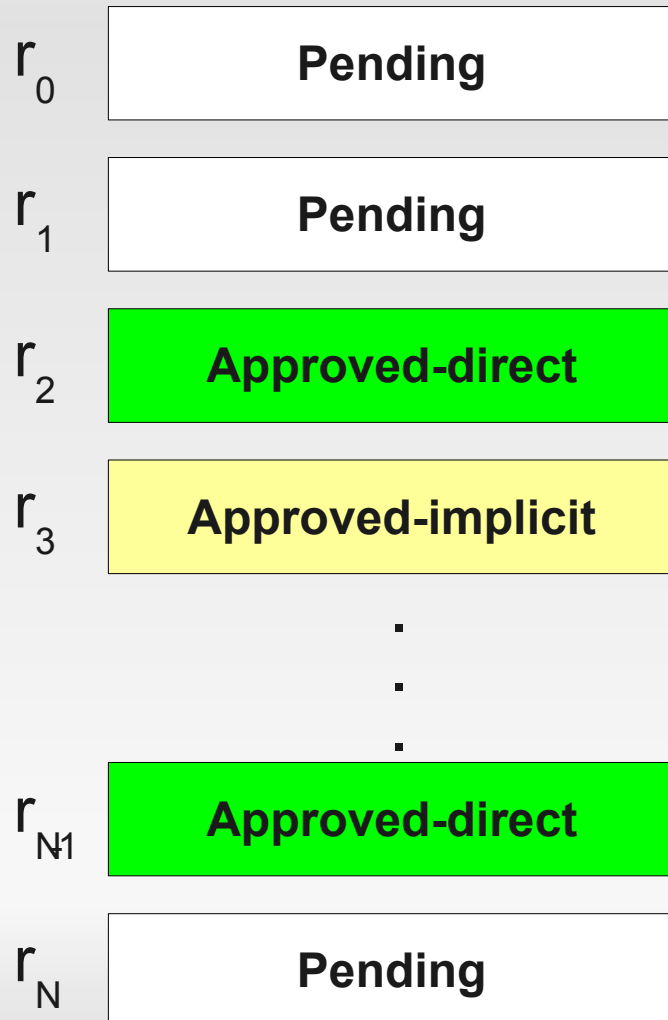
## SET OF CONSECUTIVE ANON REVS (page X)



Then, we mark  
all revisions  
*directly approved*

# Methodology (approvals)

## SET OF CONSECUTIVE ANON REVS (page X)

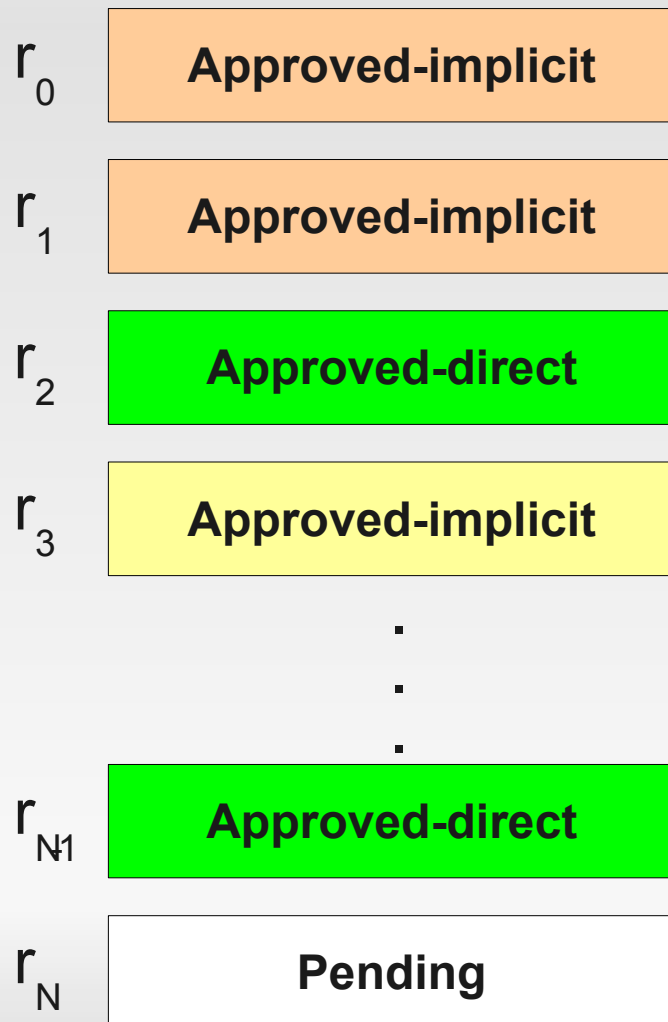


All revs in between  
2 app-direct are  
also *approved  
implicitly*



# Methodology (approvals)

## SET OF CONSECUTIVE ANON REVS (page X)



All revs before  
direct approval  
are also *approved  
implicitly*

# Methodology (reverts)

- **Vandal** reverts
  - Identified by regexps.
  - Does not include standard admin reverts.
- **Other** reverts.
  - Reverts without explicit reference to vandalism
  - Includes admin reverts.

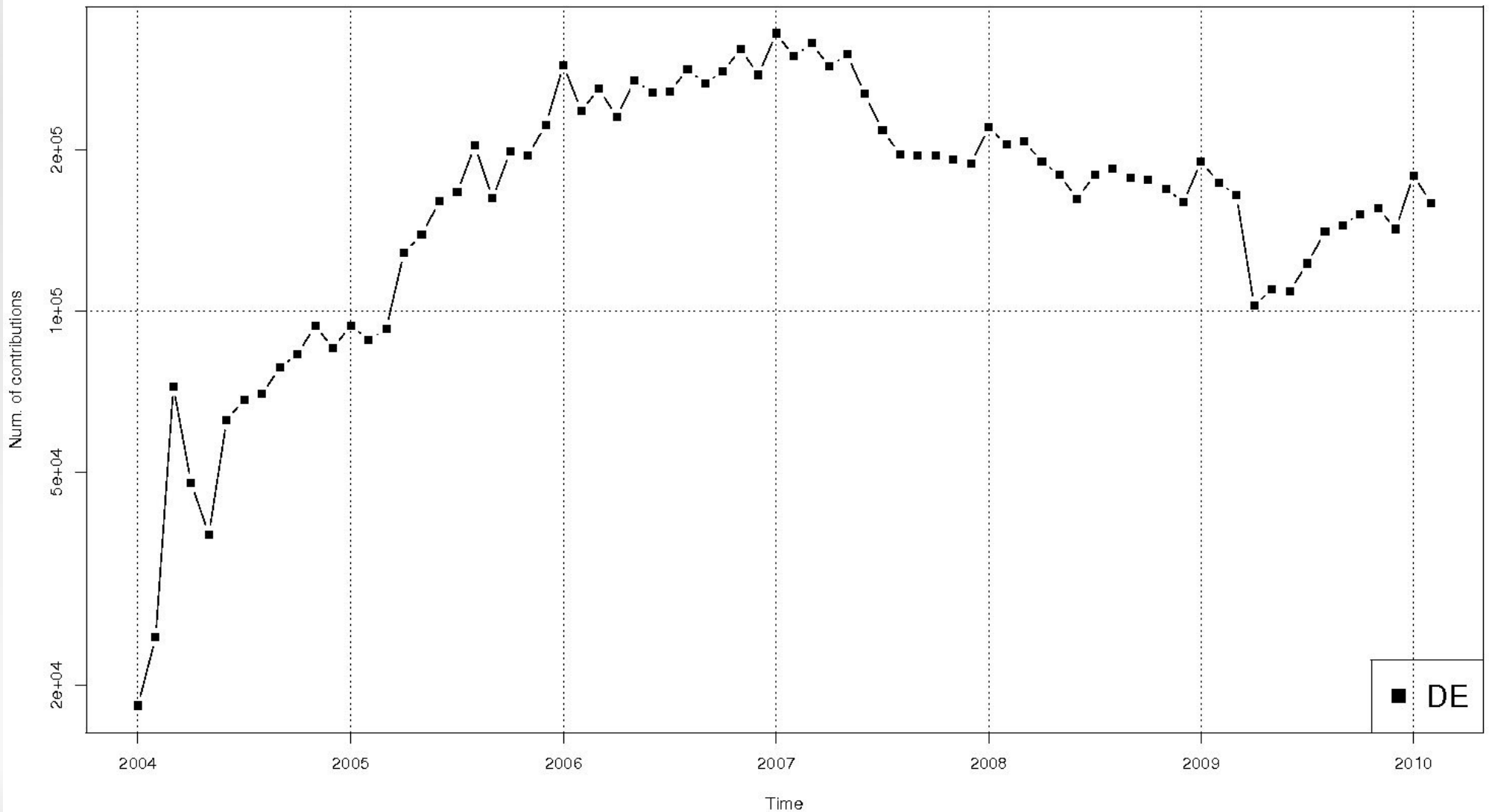
# Methodology (reverts)

- Detecting reverted revisions
  - First, look for the newest revision with the same size (after revert action).
  - From that on, mark as reverted all consecutive revisions performed by same IP.
  - If that fails, look for IP info in the comment field, and look for newest revision performed by that IP.
  - Then, same procedure to mark all consecutive revisions.
- Feedback?

# RESULTS

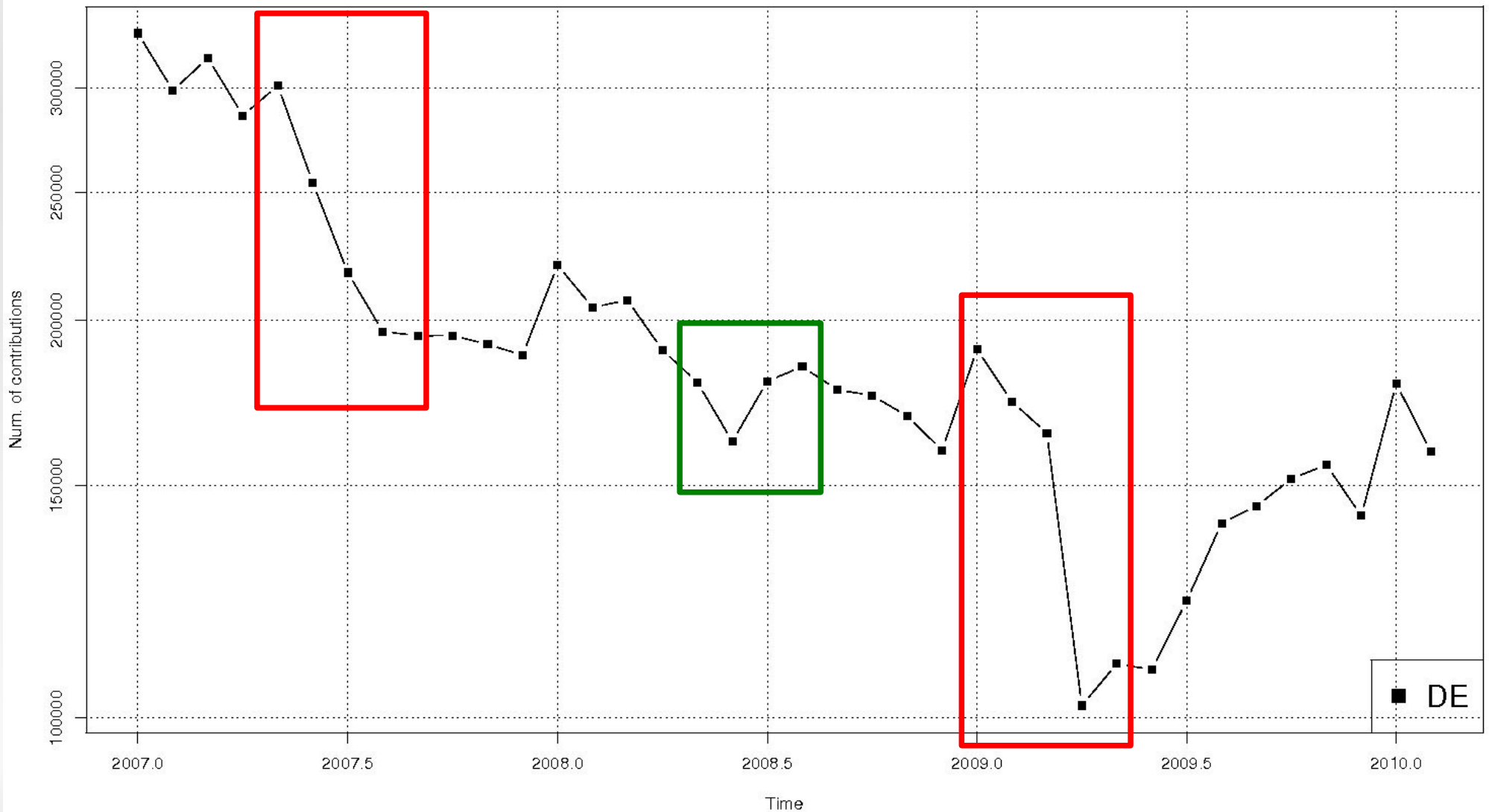
# Evolution anonymous edits

Number of contributions from anonymous users

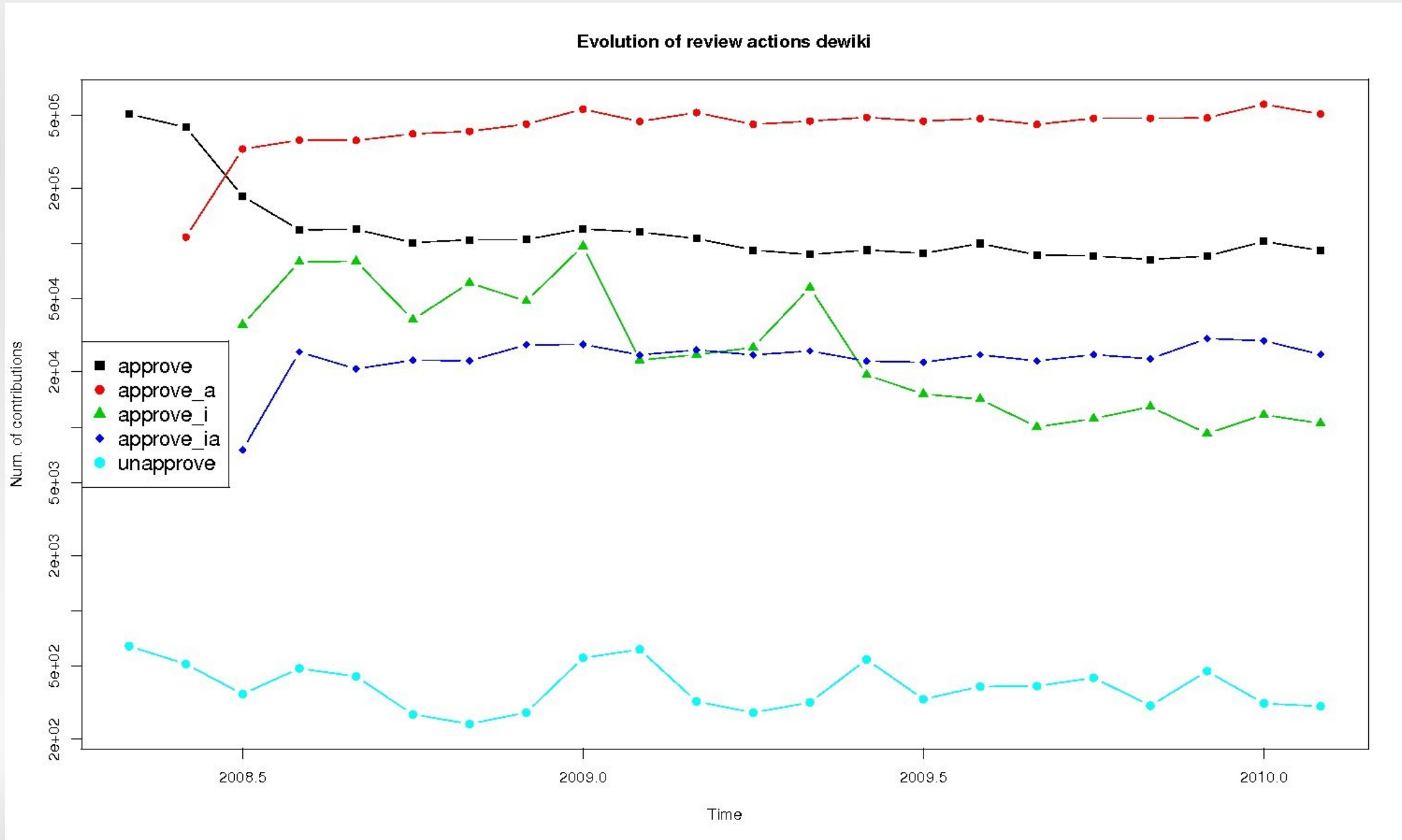


# Evol anonymous edits (focus)

Number of contributions from anonymous users

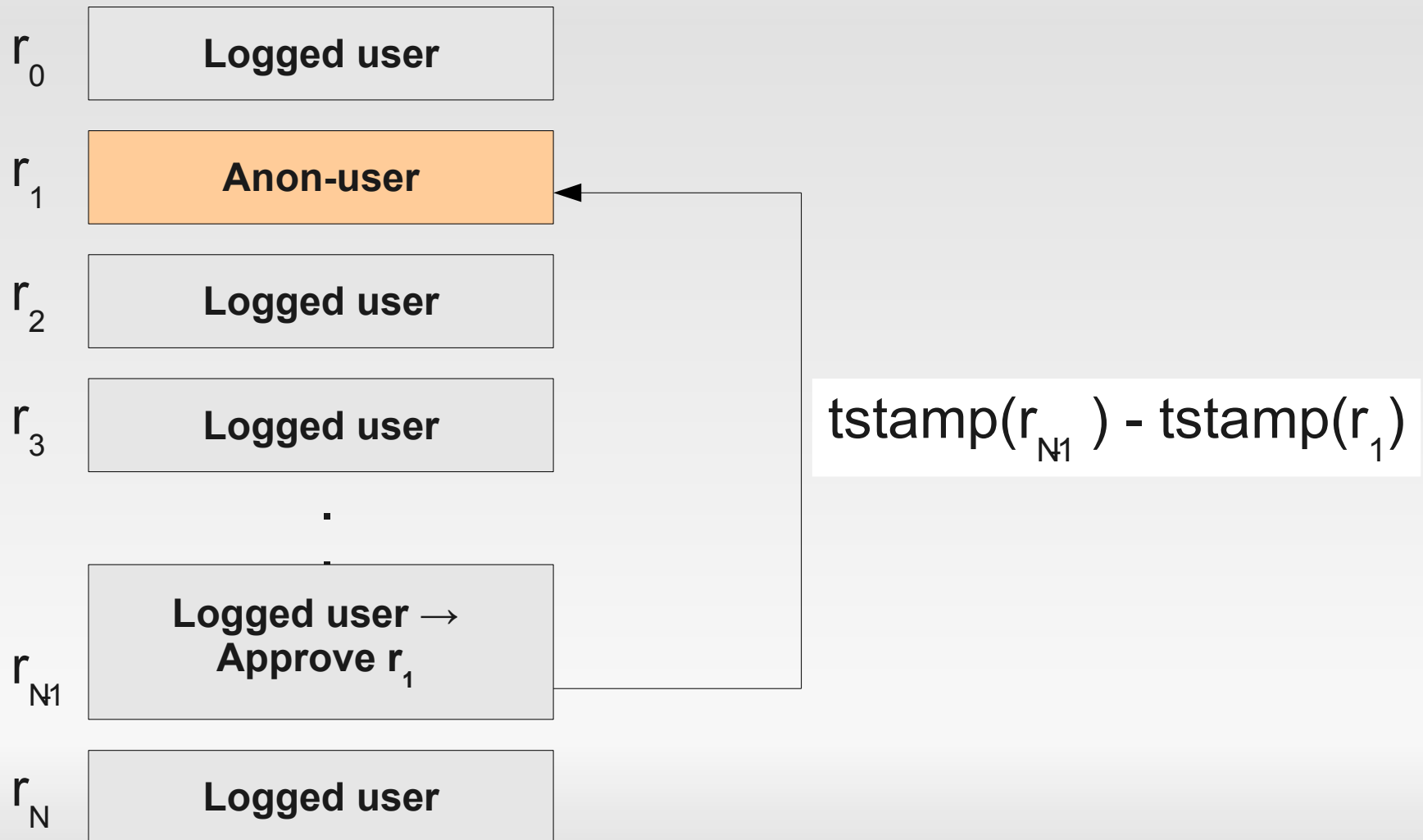


# Evolution review actions



# Time to approve/revert revisions

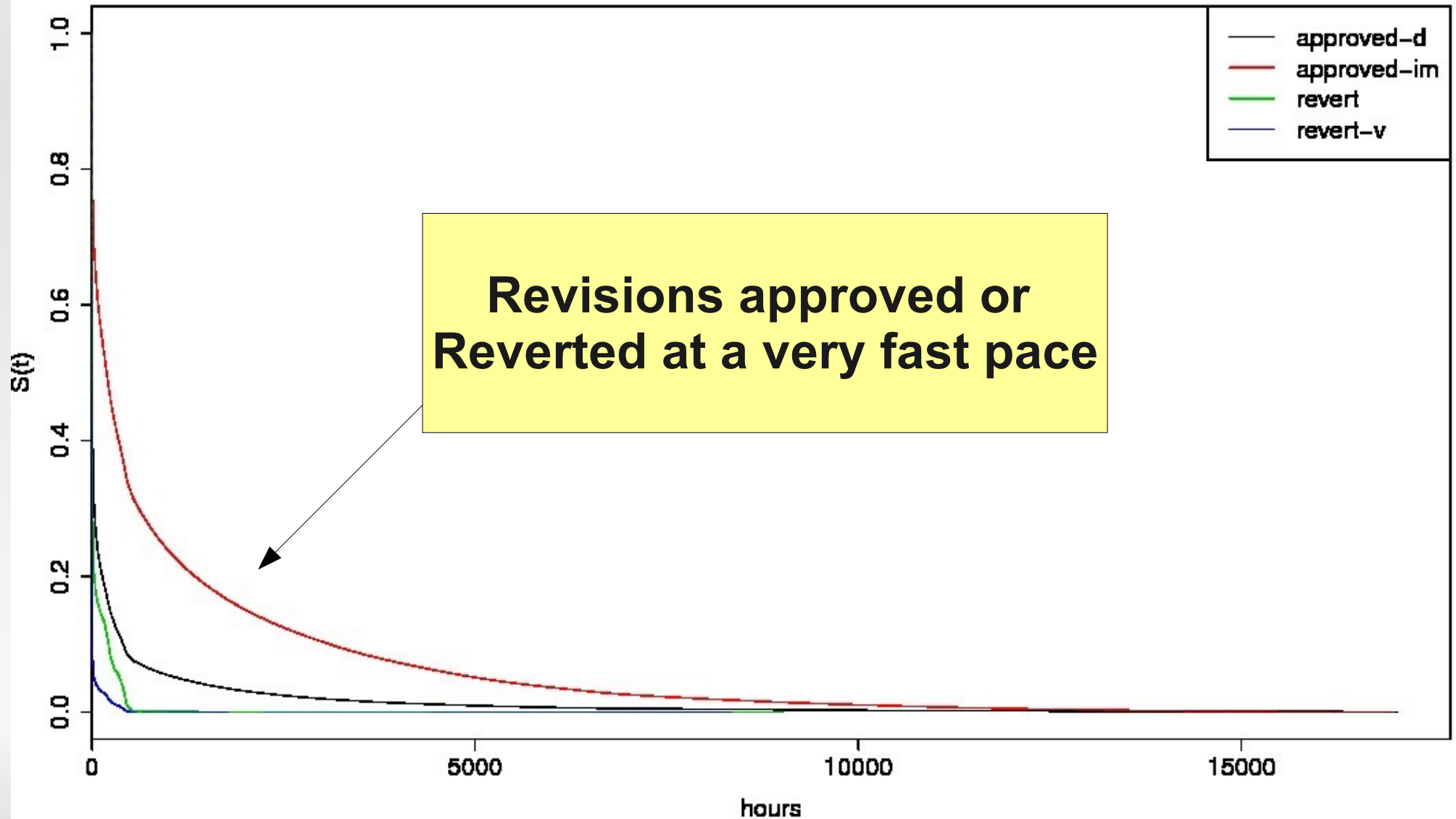
## SET OF CONSECUTIVE ANON REVS (page X)





# Time to approve/revert revisions

$S(t)$  for sighted edits from anonymous users



# “Truth in numbers”

	events	*rmean	*se(rmean)	median	0.95LCL	0.95UCL
Status=1	2892231	241.8	0.578	3.369	3.345	3.3933
Status=2	1839583	988.0	1.442	194.082	193.445	194.7781
Status=3	31774	51.0	0.744	0.807	0.756	0.8700
Status=4	119632	10.9	0.168	0.010	0.010	0.0103

1 = Approved-d ; 2 = Approved-im ; 3 = revert ; 4 = revert-v

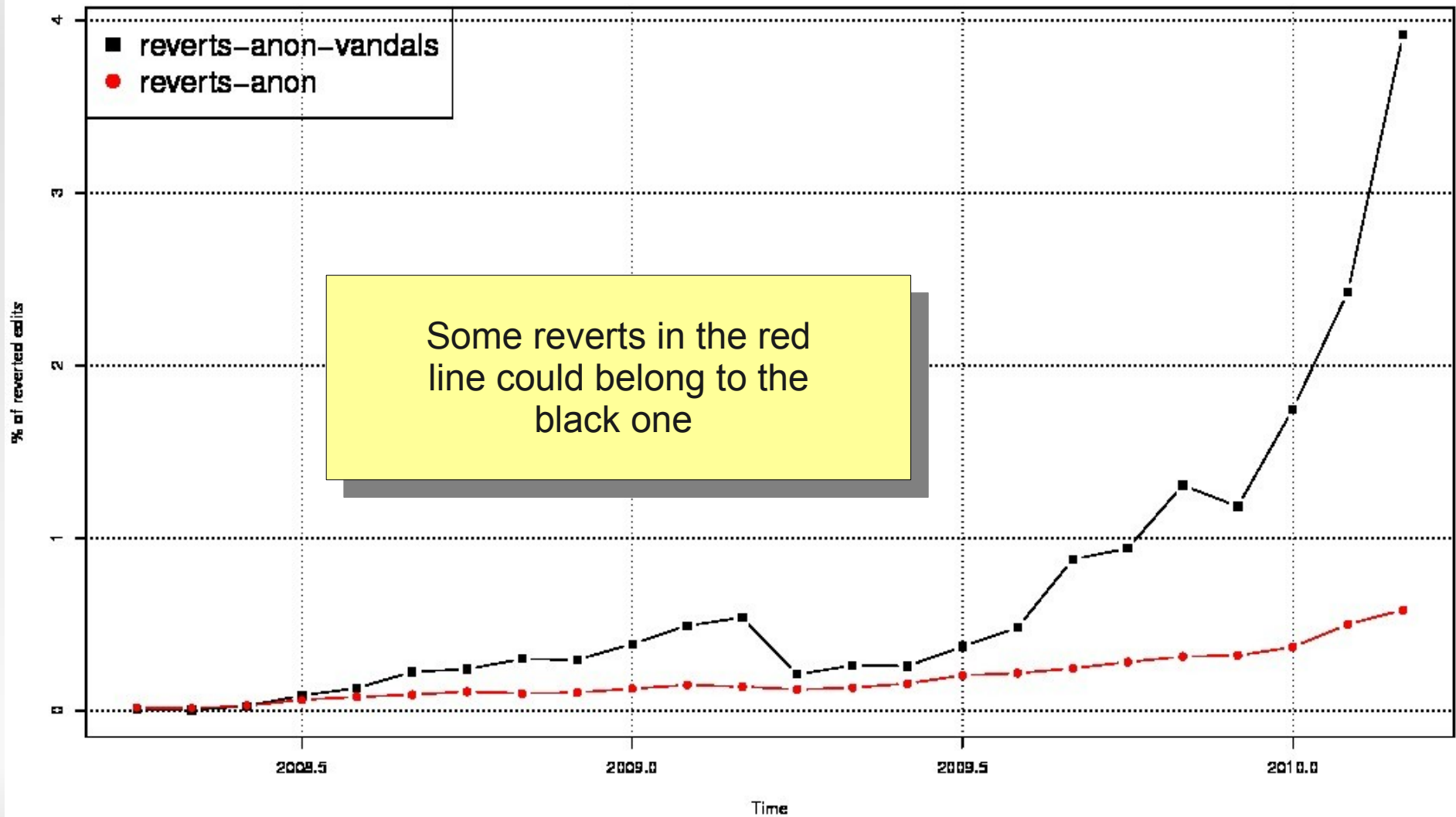
- Revert actions performed at very fast pace.
  - **Revert** (median) : 48 min.
  - **Revert-v** (median) : **36 seg.** (iii)
- High number of actions registered → **accuracy**

# Comments on time to app/revert

- Implications
  - Looks like extremely fast pace for acting on revisions.
  - Community takes this new role very seriously.
  - Provides stronger incentive to watch content even closely.

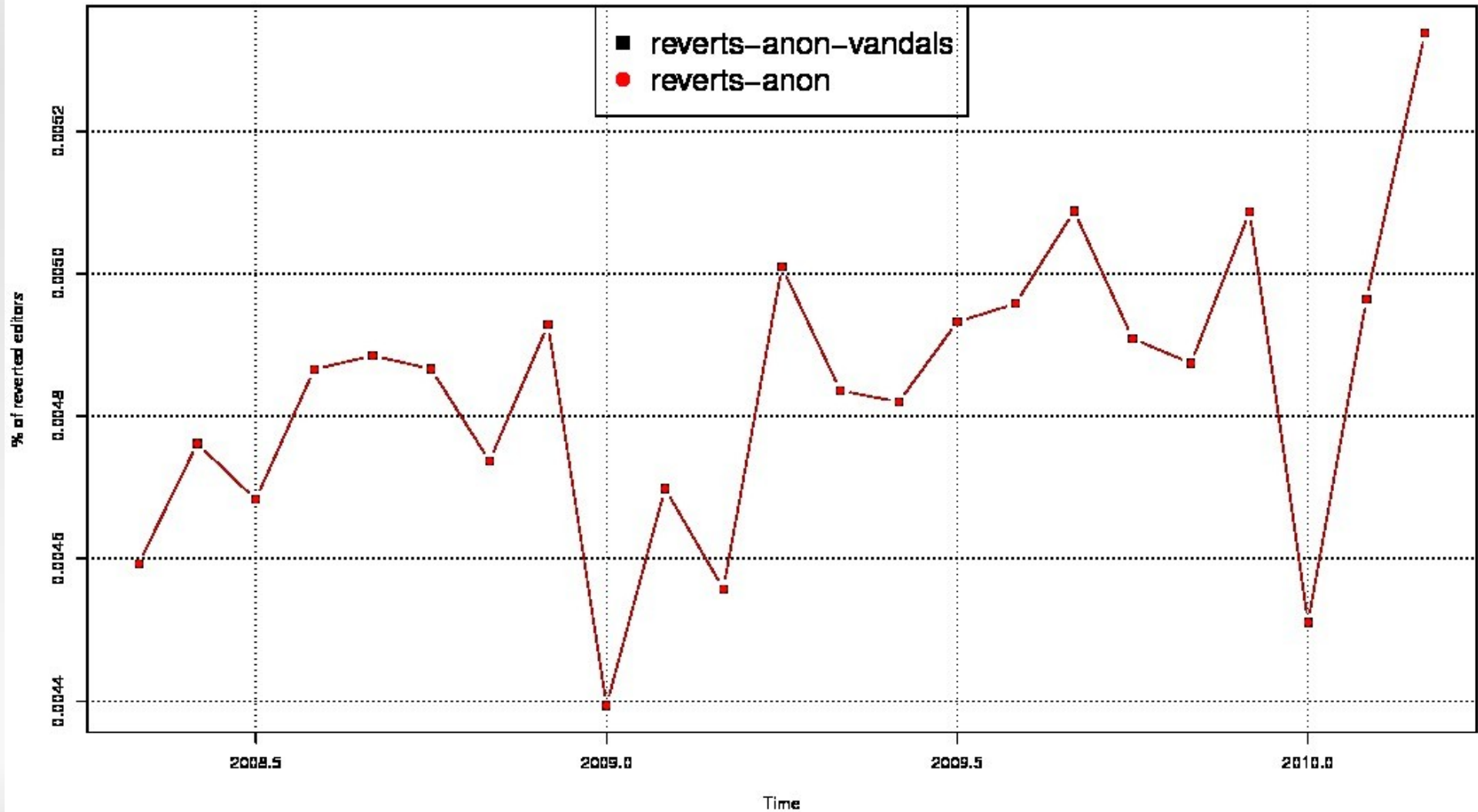
# Evolution % reverted edits

Development of %reverts over monthly anon edits



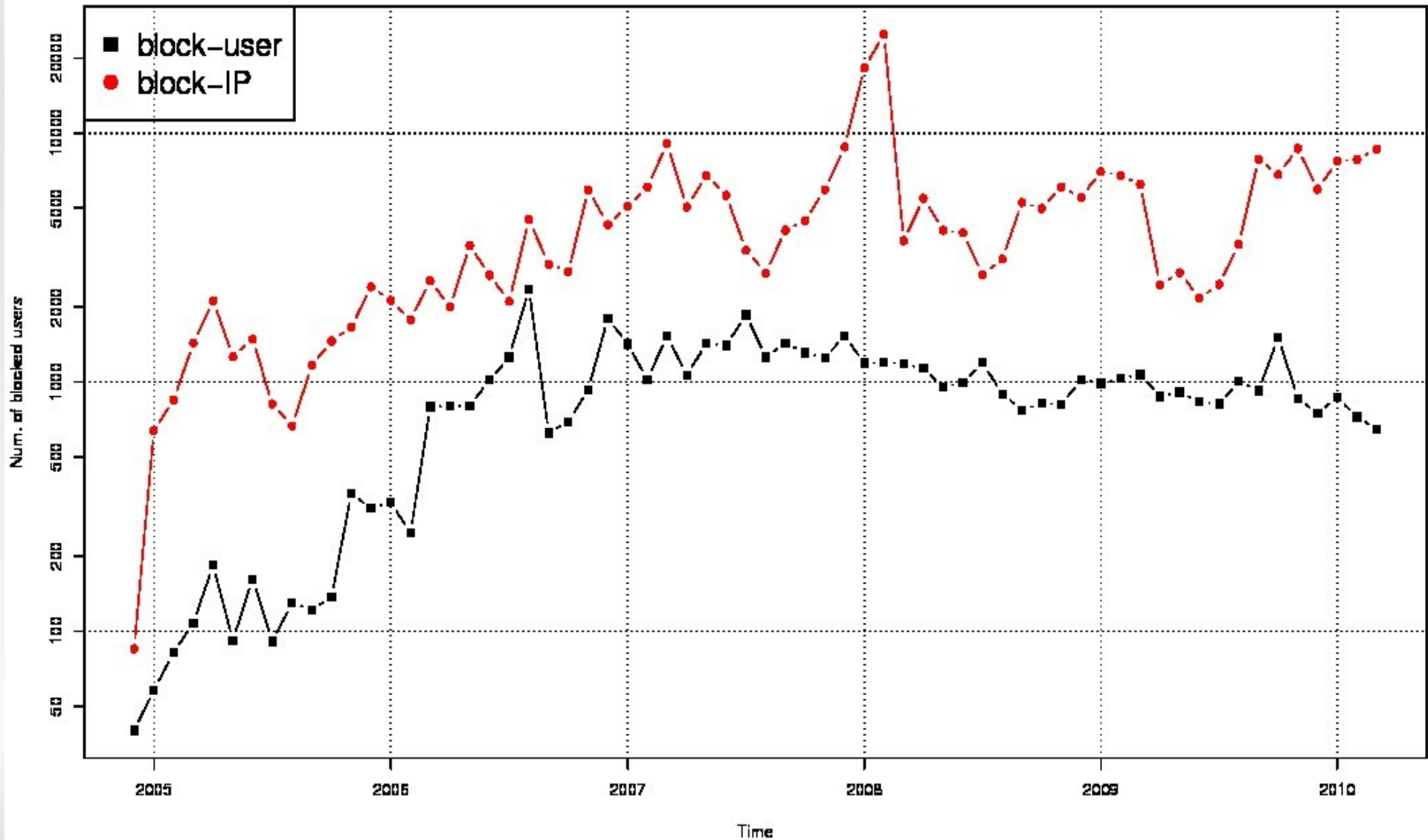
# Evolution % editors who revert

Development of %reverted pages over monthly editors



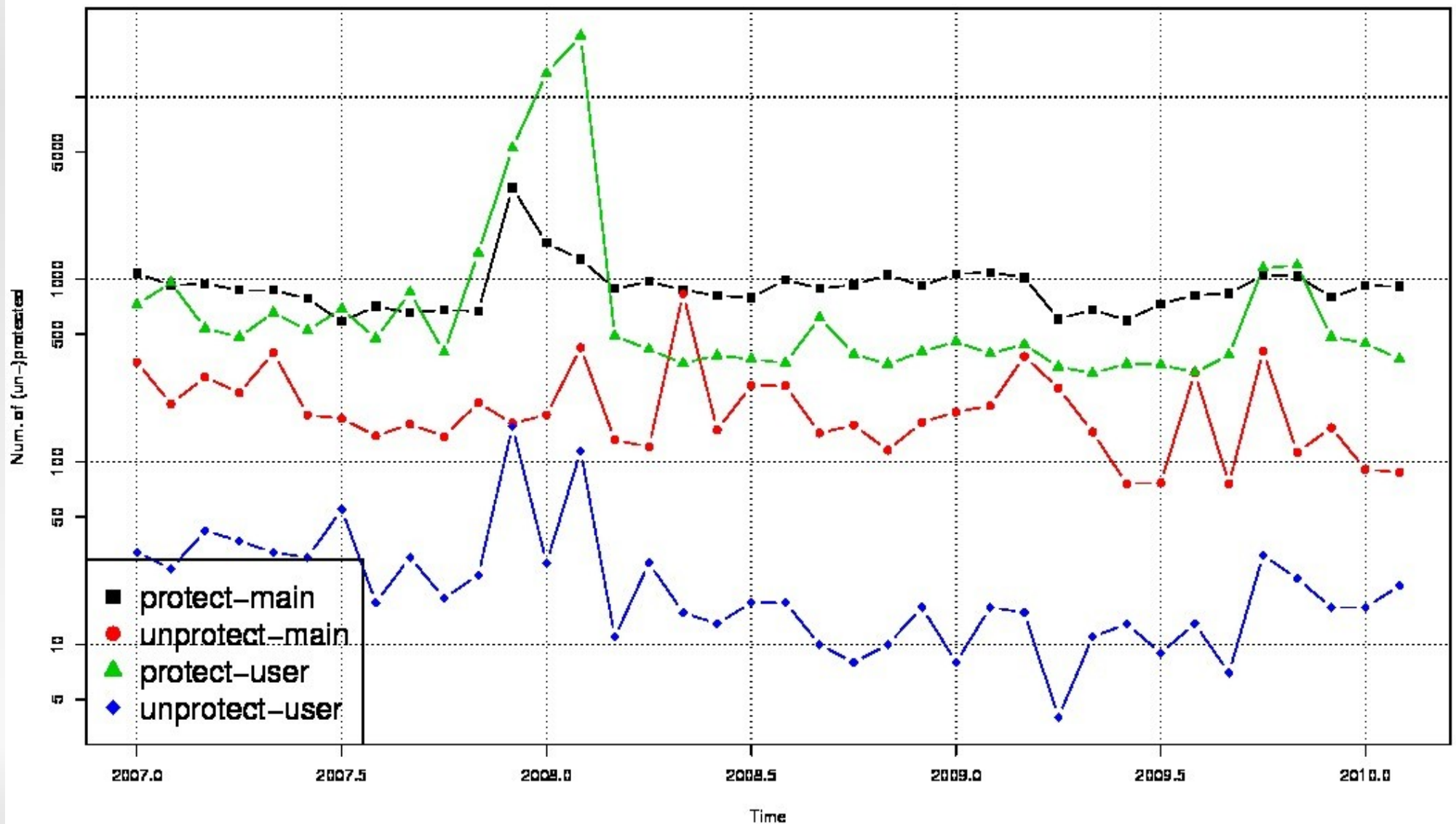
# Evolution blocked users

Development of num. blocked users



# Evolution of protection actions

Development of num. protections



# Conclusions

- In general, flagged revisions did not affected the anonymous editing.
  - Most revisions got approved very rapidly
- More activity on vandalism reverts.
  - Even faster than approval actions.
- Reduced impact of vandalism.
  - Growing number of reverts.
  - On an increasing number of pages.
- Mandatory comments had much more direct influence.



# Open questions / feedback

- Q: What did happen at the beginning of 2008 for such a high number of user pages protected?
  - A: mass-blocking of open proxies.
    - Creating the user page of blocked IPs with a template and protecting them.
- We need patterns for detecting reverts:
  - Russian Wikipedia
  - Polish Wikipedia
- Comments and feedback are very welcome.