

Evaluation Difficulty

- Information is not stored in Git, so it is not possible to compare with existing archive
- There is no "Gold Standard" (i.e., no current repository with this information)
- Algorithm is based on heuristics, so we cannot prove that they work in all cases

Evaluation

- Generated test repositories: precision and accuracy
- Case study: accuracy assessment on OSS
- Manual verification: precision on sample

$$\text{precision} = \frac{\# \text{ correctly identified commit origins}}{\# \text{ identified commit origins}}$$

$$\text{accuracy} = \frac{\# \text{ commit origins identified}}{\# \text{ commits in repository}}$$