

## Steps for tagging the inducing commits

- Step 1. Preliminary tagging. Each labeler locates the file and code module where the vulnerability occurred by examining the fixing commits and IMPACT information, and then scrutinizing the file's historical commits to identify the inducing commit of the vulnerability. To ensure accuracy, we assign two-thirds of the 429 vulnerabilities to each labeler, allowing each repository to be tagged twice. After the preliminary tagging, we used Krippendorff's  $\alpha$  coefficient to calculate the inter-coder agreement for each domain to ensure the validity of the results. After the calculation, the range of inter-coder agreement for domains is 0.85~0.91, which satisfies reliability requirements.
- Step 2. Resolving conflicts. After preliminary tagging, there are 13/145, 21/145, and 17/145 repositories that are tagged differently by the labelers. The labelers of these repositories discuss with each other to resolve conflicts. When they cannot agree among themselves, the third labeler will make the final decision on the tagging result.