Merge-tree: Visualizing the Integration of Commits into Linux

Evan Wilde, Daniel German.

Presented by: Nick Bradley
March 16, 2017
Git Version Control

• Is a distributed version control system
• Supports non-linear workflows
• Uses directed acyclic graph (DAG)

Fig. 3

### Commits

- #1
- #3
- #6
- #8

### Branches

- Branch 1
- Branch 2
- Branch 3
- Master

① = commit
Domain Task

• **Show when and how a commit was merged into master**

• Challenge: commits cannot be changed
  - can’t link to later commits
  - can’t track merge dates

• **Solution: Linvis**
  - Shows topological view of merges
  - Supports aggregation and filtering
  - Supports two use-cases
    1. top-to-bottom: aggregate
    2. bottom-to-top: see flow into master

Source: http://li.turingmachine.org
Merge-tree

• Transforms DAG into trees
  • each rooted on master
  • such that all commits are assigned to exactly one tree

• Algorithm
  • Invert DAG
  • For every commit
    • compute distance to each subsequent commit
    • only keep link to closest (in time) in merge-tree
    • stop at master commit

• Relies on specific Git workflow
Convert DAG to Trees

**DAG Model**
- Newer commits link to older
- All links present

**Merge-tree Model**
- Older commits link to newer
- Removed links
  - Only keep links on shortest path
Linvis Live Demo

http://li.turingmachine.org
Linvis analysis

• What: data
  • DAG
• What: derived
  • Merge-tree
• Why: tasks
  • Search for commits
  • Summarize changes
• How: reduce
  • Filtering
• How: Manipulate
  • Navigate with pan/zoom
  • Select

• How: encode
  • indented outline (*list tree*)
  • tree map using nested circles and radial containment (*bubble tree*)
  • vertical node-link (*Reingold-Tilford tree*)
Limitations + Next Steps

• No evaluation of Linvis
  • quantitative user-testing: improvements to user workflow
  • qualitative user-evaluation: do users think tool is helpful

• Merge-tree cannot be constructed for most repositories

• Cannot search by filename

• Aggregate commit patches
• Aggregate authorship information
Compare existing - GitHub

Fig. 2

Couldn’t load network graph.
Too many forks to display.

Fig. 15
Compare existing - Gitk

Fig. 16

Fig. 17
Critique

• Strengths
  • Main contribution of merge-tree
  • Demonstrated on most complex DAG
  • Natural interactions (e.g. pan, zoom)
  • Different encodings of tree structure, all intuitive

• Weaknesses
  • Merge-tree algorithm not robust
  • Navigation between views cumbersome
  • Vis felt like afterthought
Suggestion

• Juxtapose:
  • Git log
  • Files
  • tree vis
Questions
Linvis Search View

- Allows filtering by
  - Merge date range
  - Commit author
  - Keyword
  - Commit ID

Fig. 7
Linvis Message Tab

- Displays Git log
- Included for completeness

**Fig. 9**
Linvis File Tab

- Files changed in leaf commits
- Added/Removed columns show number of changed lines
- Aggregates number of changes to single file across all commits

![Merge branch 'akpm' (patches from Andrew Morton)](image)

Fig. 10
Linvis Module Tab

- Linux specific
- Groups files
- Count column shows number of changed files for all leaf commits

*Fig. 11*
Linvis List Tree

• Text-based representations of the merge-tree
• Nested lists show the hierarchy
• Designed to model tree-views of file browsers which are familiar to developers
• Easy to search and navigate

Fig. 12
Linvis Bubble Tree

- Organizes the commits hierarchically by having the parent commit contain the child commits
- Similar to tree maps but clearly shows leaf commits
- Good for providing clear visualization of wide, hierarchical data
Linvis Reingold-Tilford Tree

- Intuitive representation of merge-tree
- Not effective at displaying large trees

Fig. 13