10. Appendix: User Study Text and Task Descriptions

This section contains the presented text and questions used in the different stages of the user study.

10.1. Training Stage

The Training Stage did not have an explicit script as participants were encouraged to ask questions during the system tutorial. Instead, after introducing the idea of "resolving conflicts insights in collaborative visual analytics" for large and deep datasets, the following lists itemize the necessary features and interactions that had to be covered to describe each system.

CoVA and Baseline shared features and interactions:

Login Interface:

- Log in and log out the system
- Create and share a project

Project Interface:

- Page Header:
  - "CoVA" Link: Go back to Login Interface
  - Project name shown in the middle
  - "Review" or "Explore" Button: Switch between Explore Interface and Review Interface
- Explore Interface:
  - Data visualization Board (left pane)
    - Visualization output Board (upper pane)
      - User study preset graph choices
      - Visualization chart rendering canvas
    - Visualization input Board (bottom pane)
      - Data description
      - Chart spec input text box
    - Middle function bar
      - "Update" Button: Update visualization after changing spec
      - "Insight" Button: Form new nodes based on the visualized data dimensions
      - "Provenance" Button: Add the visualization as a provenance to the Insight Graph
  - Insight creating and sharing Board (right pane)
    - Git menu (upper pane)
      - "Commit" Button: Save current insight graph to local
      - "Push" Button: Push the local stored insight graph to remote repository
      - "Pull" Button: Pull the remote repository stored insight graph and merge with the local stored insight graph
    - Insight graph board (bottom pane)
      - Create insight graph to store findings

- Functions include: Modify node, Remove node, Add link, Modify link, Remove link, Hover on link, Drag node/link, Click to change node/link label, Hover on node/link to view details, Double click to fix details, Click chart to see bigger view, Restore chart, Lasso function

- Review Interface
  - Git tree board (left pane)
    - Show the history of commits (the building process of current insight graph)
    - Click on each commit to see insight graph in that commit stage
  - Trail insight graph board (right pane)
    - Show the insight graph of the chosen commit

- All the windows are draggable and resizable

CoVA only features and interactions:

Resolve Conflicts and Merge Insights Interface

- Conflict List Board (left pane)
  - Show all the conflicts between their (collaborators) and our (users) Insight Graphs
  - Each conflict provides
    - Cause of the conflict
    - Position of the conflict
    - "Keep Ours" Button: Resolve the conflict based on "Ours" (Users) graph
    - "Keep Theirs" Button: Resolve the conflict based on "Theirs" (Collaborators) graph
    - "Keep Both": Resolve the conflict by combining the content of the conflict in two graphs

- Resolution Board (middle and right pane)
  - "Theirs" Graph: The current Insight Graph stored in the remote repository, created by collaborators
  - "Ours" Graph: User's current Insight Graph
  - "Merged" Graph: The final merged Insight Graph generated based on the chosen resolutions for each conflict
  - Users can edit the three insight graphs with the same features and interactions provided in "Insight graph board"
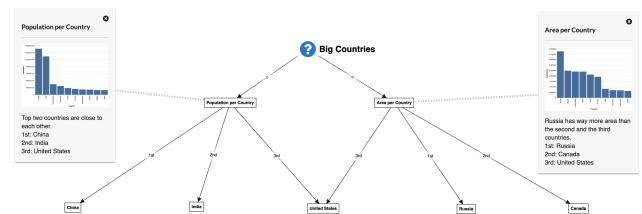

Figure 13: *A screenshot of the initial Insight Graph provided by the fake collaborators in the training stage.*

Training Task

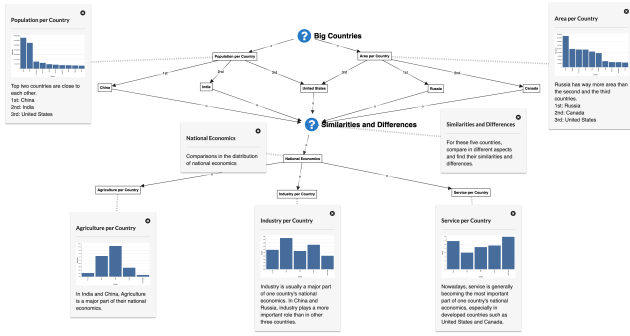The introductions for the training task were printed on a sheet of

Figure 14: *A screenshot of the Insight Graph with new findings provided by the fake collaborators in the training stage.*

paper and handed to the participant. The participant was guided to finish the training task, and at the same time, gained familiarity with all the features and interactions of the provided system. The text for the document is as follows:

Training Task

You and your teammates are trying to investigate "big" countries all over the world. Before you join this project, your teammates have already done some data exploration and found five big countries. Please review their findings and continue to find similarities and differences in these countries in one of the following aspects:

1. Geographical data, or
2. Demographical data

Check the preset visualizations in the upper left corner of this system, and then create reasonable nodes and links in the insight graph to save your findings.

After finishing this task, commit to save your current insight graph to the local repository and then push it to the remote repository. If needed, pull the remote insight graph to local and then merge it with your own graph.

Participants were originally provided with the initial Insight Graph created by the fake collaborators shown in Figure 13, which shows the main task of finding "big" countries and five big countries found through ranking countries by population and area. After participants started working on the task, the system automatically pushed the Insight Graph with new findings to the remote repository to simulate the fake collaborators' progress in data exploration. So when participants finished this task, committed their findings as a new Insight Graph to the local repository, and wanted to push it to the remote repository, they received a warning from the system asking for pulling before pushing. After they pulled the remote Insight Graph (Figure 14) to local, the results were combined through the conventional graph combining algorithm [CYM*10], and the participants in the CoVA group were required to merge the pulled results with their own graphs by solving potential conflicts using CoVA's

visual interface. Finally, participants had the options to further modify the combined Insight Graph and were required to commit and push the final Insight Graph. Each participant was provided with the same initial Insight Graph Figure 13 and the same Insight Graph with new findings Figure 14 in this training task.
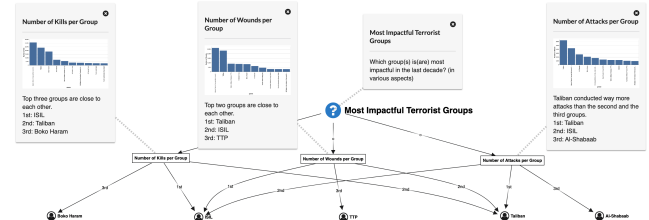

Figure 15: *A screenshot of the initial Insight Graph provided by the fake collaborators in the task stage.*
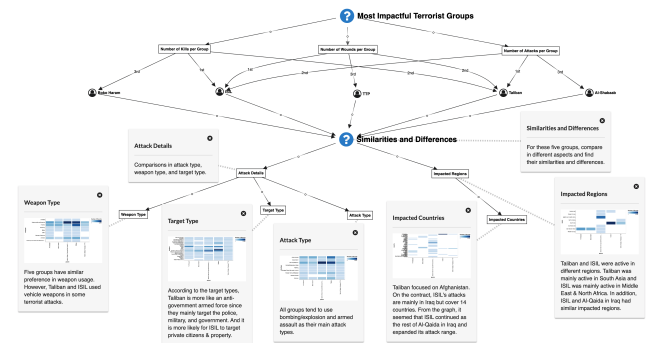

Figure 16: *A screenshot of the Insight Graph new findings provided by the fake collaborators in the task stage.*

10.2. Task Stage

The Task Stage instructed users to perform a single task. The introductions for the stage were printed on a sheet of paper and handed to the participant, who read the instructions and subsequently began the appropriate task. The text for the document is as follows:

Task Stage

In this stage, you're going to investigate a dataset called the **Terrorism Dataset**. The terrorism dataset shows information about terrorist attacks around the world since 1970.

In the 1st part of this study, we will have you complete 1 task. We will give you a sheet of paper that describes the task. After you read it, you are free to start. Please remind me when you start, so I can start my stopwatch. When you're done, tell me "I'm done" and I will stop my stopwatch. You can use this sheet of paper as a scratchpad if you want.

If you are stuck, confused, or need help, feel free to ask me!

Task 1

You and your teammates are trying to investigate the most impactful terrorist groups in the last decade. Before you join this project, your teammates have already done some data exploration and found the five most impactful terrorist groups. Please review their findings and continue to find similarities and differences in these groups in one of following aspects:

1. Kills and Wounds
2. Activities
3. Success rate

Check the preset visualizations in the upper left corner of this system, and then create reasonable nodes and links in the insight graph to save your findings.

After finishing this task, commit to save your current insight graph to local and then push it to the remote repository. If needed, pull the remote insight graph to local and then merge it with your own graph.

In this stage, participants followed a process similar to the one in the training task. An initial Insight Graph (Figure 15) created by the fake collaborator was provided to each participant. After participants started their data exploration, the system automatically pushed the Insight Graph with new findings (Figure 16) into the remote repository. Then participants were required to go through the following process:

- Commit and attempt to push their findings to the repository.
- Pull the fake collaborator's latest findings (Figure 16) after notified by the system.
- Combine two Insight Graphs by resolving possible conflicts using either the conventional graph combining algorithm (the baseline group) or the visual interface for conflict identification and resolution (the CoVA group).
- Commit and push the combined graph.

10.3. Freeform Stage

After completing the Task Stage, each participant was handed a single sheet of paper with the following instructions printed at top:

Freeform Analysis Stage

This next stage is called the freeform analysis stage.

From the first stage, you should already get a general idea about the usage of insight graph in data exploration. In the basis of the current insight graph, please summarize all the given information and draw some new insights.

For example:
○ ISIL: a group of brutal terrorists
○ Taliban: an anti-government armed force

While you do this, you're going to use a technique called **think-aloud protocol**. Verbally speak whatever comes into your mind as you're using the system. This might include what you are looking at, what you're thinking, the action you're doing, what you're feeling, any new facts you are learning, etc.

You have 10 minutes for this. You can also use your paper and pencil if you want. I will tell you when the time is up. Afterwards, I might ask you some follow up questions about what you found during your investigation. (No pressure! There are no wrong answers and you cannot do anything wrong here!)

10.4. Review Stage

A Google Form questionnaire was used to ask the following set of questions to each participant after the Freeform Stage. Each "agreement" question (i.e., *I am comfortable with...*) was answered using a Likert Scale, with a range of 1 (strongly disagree) to 7 (strongly agree).

- Demographics Information.

  - Interface Style (options = [Baseline, Real System])
  - Gender (options = [18-24, 25-34, 35-44, 45-60, 61 and over])
  - Major / Field of Study
  - Age
  - I am comfortable with reading charts / visualizations that display data?
  - How familiar were you with the Terrorism dataset (prior to this study)? (options = [Quite familiar, Just know a little bit, Not at all])
  - Which of the following were you familiar with (prior to this study)? (options = [Node-Link Diagram or Graph Visualization, Concept Map or Knowledge Graph, Both, Neither])
  - How familiar were you with version control tools (e.g. Subversion and Git)? (options = [Quite familiar, Just know a little bit, Not at all])

- General Impressions.
  Overall, the system was:

  – easy to learn
  – easy to use

- Reflect on the functionality provided by the system.
  I felt the system was effective at:

  – Recording my insights derived through data exploration
  – Organizing insights and their relations in a clear and intelligible way
  – Saving insights for others to refer
  – Helping me understand the insights saved by my (synthetic) teammate
  – Allowing me to use my (synthetic) teammate's findings for continuing the data exploration

  Feedbacks:

  – Which functionality did you find the most useful, and why?
  – Which functionality did you find the least useful, and why?

- How useful were the interface features to you?
  I found it useful in my data exploration to:

  – Use the context menu to create nodes in the insight graph
  – Grab data attributes from the visualization (using the 'insight button') in order to create nodes in the insight graph
  – Add and review the provenances (visualizations and annotations) stored in nodes/links
  – Change the style of nodes/links (size, color)
  – Commit, push, and pull the insight graph
  – Review the history in the git tree

  (Real system only) I found it is helpful for me to understand the synthetic teammate's findings and merge our insight graphs by:

  – Visualizing the conflicts between my and the synthetic teammate's insight graphs when merging graphs
  – Choosing from different choices to solve conflicts when merging graphs

  Feedbacks:

  – Which interface feature did you find the most useful, and why?
  – Which interface feature did you find the least useful, and why?

References

[CYM*10] CHUNG H., YANG S., MASSJOUNI N., ANDREWS C., KANNA R., NORTH C.: VizCept: Supporting synchronous collaboration for constructing visualizations in intelligence analysis. In *2010 IEEE Symposium on Visual Analytics Science and Technology* (Oct. 2010), pp. 107–114. 2