A Term Distribution Visualization Approach to Digital Forensic String Search

Moses Schwartz*† and L. M. Liebrock*

*New Mexico Institute of Mining and Technology

†Sandia National Laboratories

Sandia National Laboratories

NEW MEXICO TECH
SCIENCE · ENGINEERING · RESEARCH · UNIVERSITY

Digital Forensic String Search [1,2,3,5]

- Search physical media for readable strings to identify forensic artifacts
- Used to recover data from
  - Allocated space (files)
  - Unallocated space (deleted files)
  - Slack space (unused space at block boundaries)
- State of the Art method: Grep

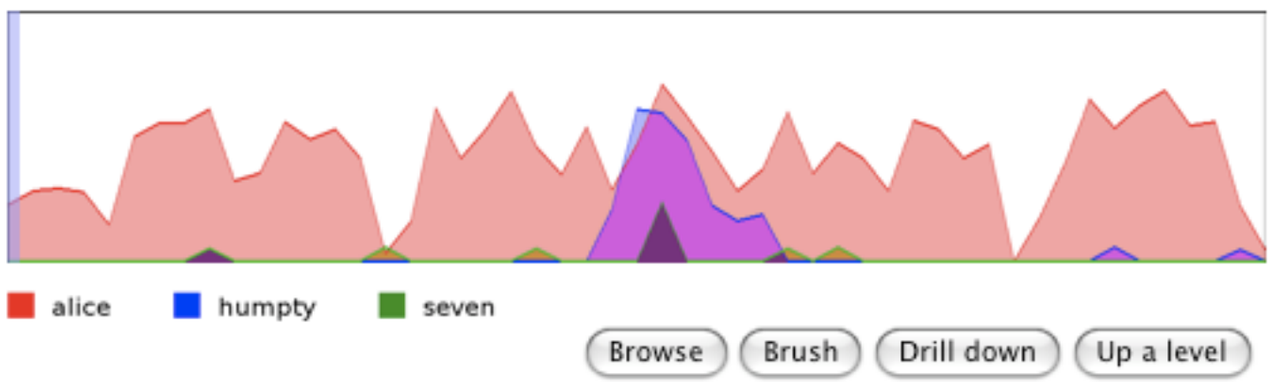
The Problem [1,2]

- Large datasets
- Unstructured data
- Very high hit rate
- Very high false positive rate
- All hits must be manually reviewed by a human analyst

Solution Approaches [1,2]

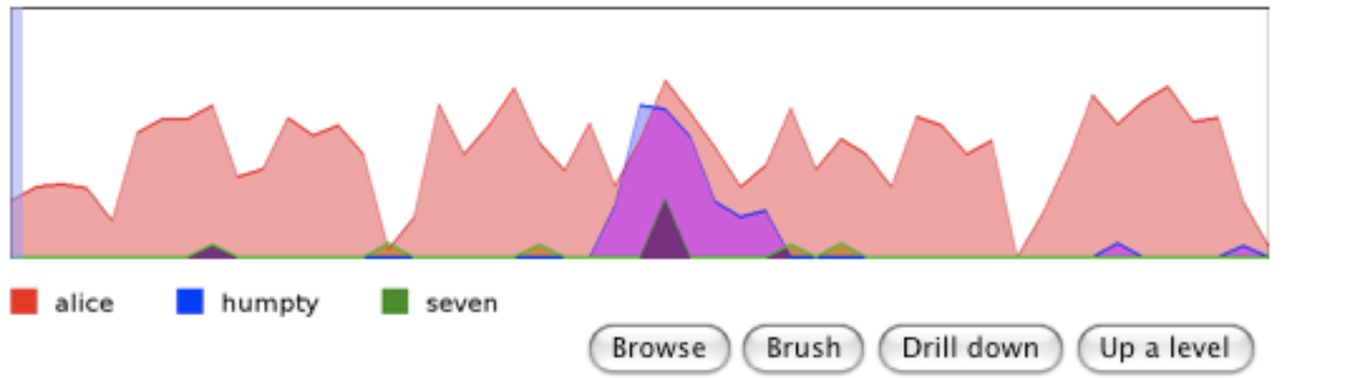
- Advanced information retrieval techniques
  - Clustering algorithms
- Information Visualization
  - Leverage the eye's bandwidth
  - Decrease cognitive load
  - Simplify the human analyst's task

Sequential Histogram [4,6]

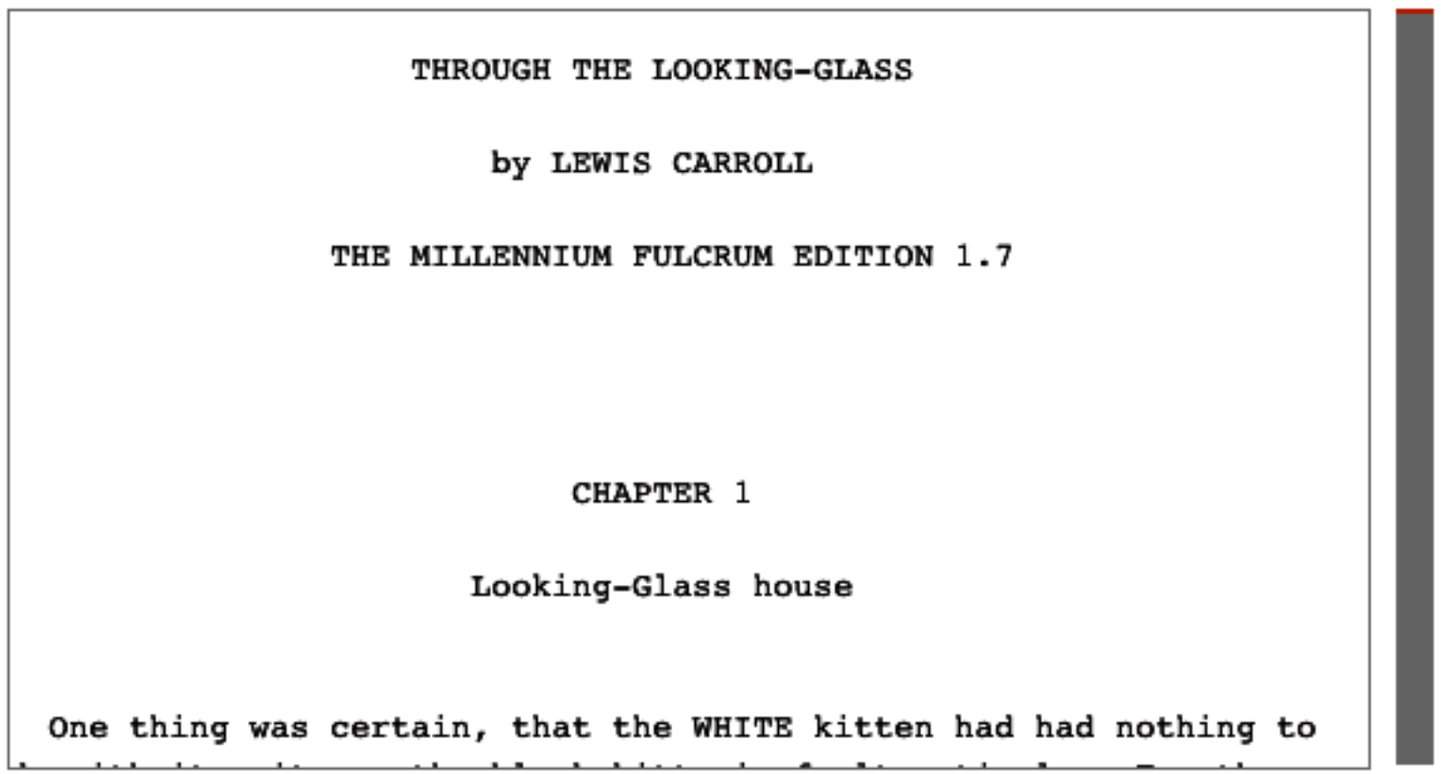


- Visualize the distribution of search terms throughout a dataset as a sequential histogram

Corresponding Text [6]

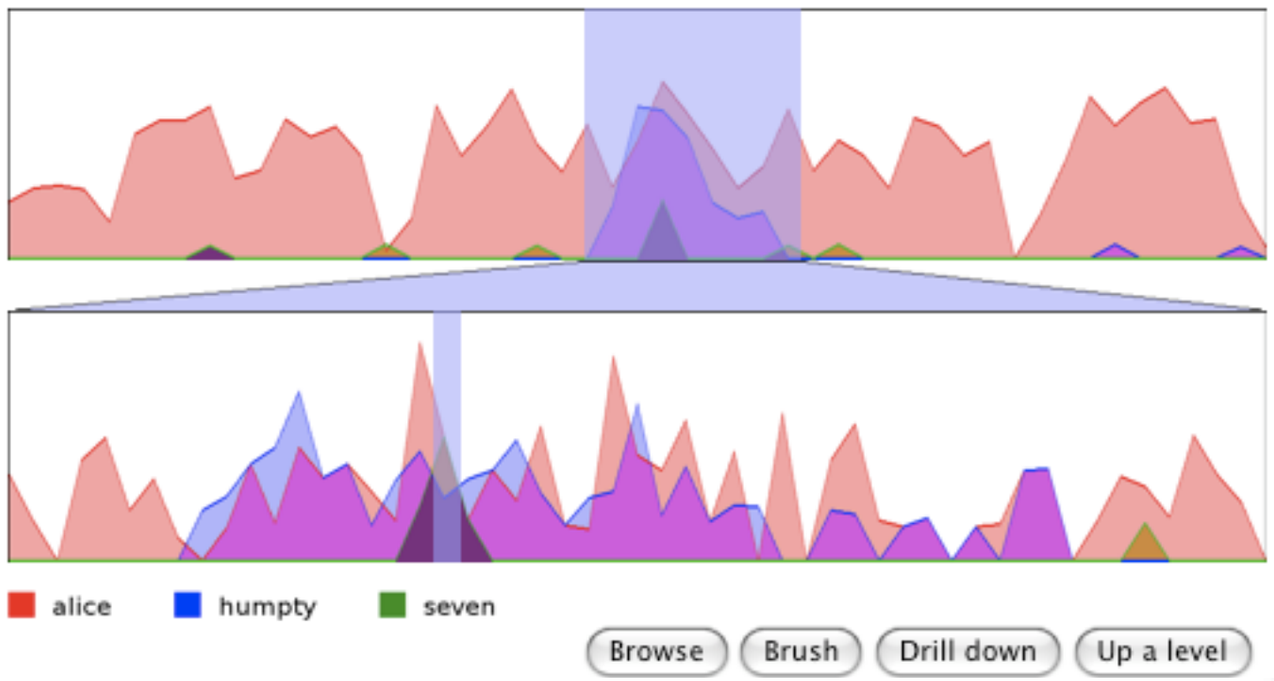


- Display text corresponding to a brushed section of the histogram



Focus+Context [6]

- Provide a Focus +Context mechanism to enable dataset exploration and information retrieval



'I though you meant "How old ARE you?"' ALICE explained.

'If I'd meant that, I'd have said it,' said HUMPTY Dumpty.

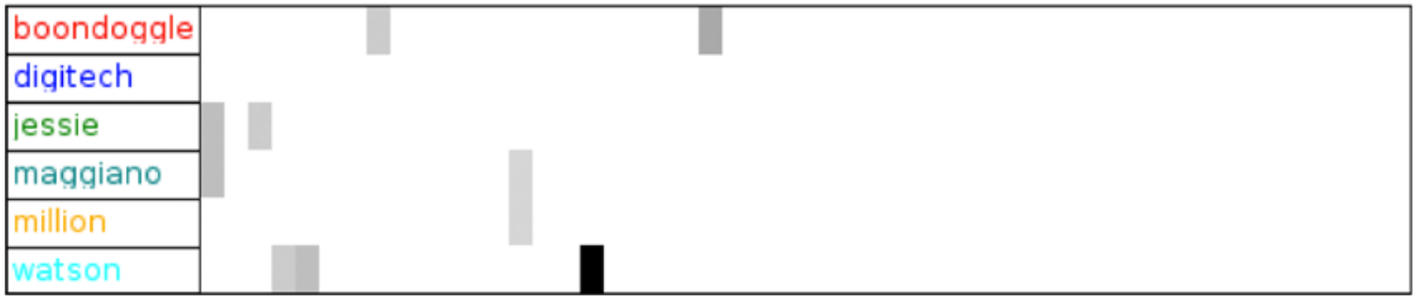
ALICE didn't want to begin another argument, so she said nothing.

'SEVEN years and six months!' HUMPTY Dumpty repeated thoughtfully. 'An uncomfortable sort of age. Now if you'd asked MY advice, I'd have said "Leave off at SEVEN"--but it's too late now.'

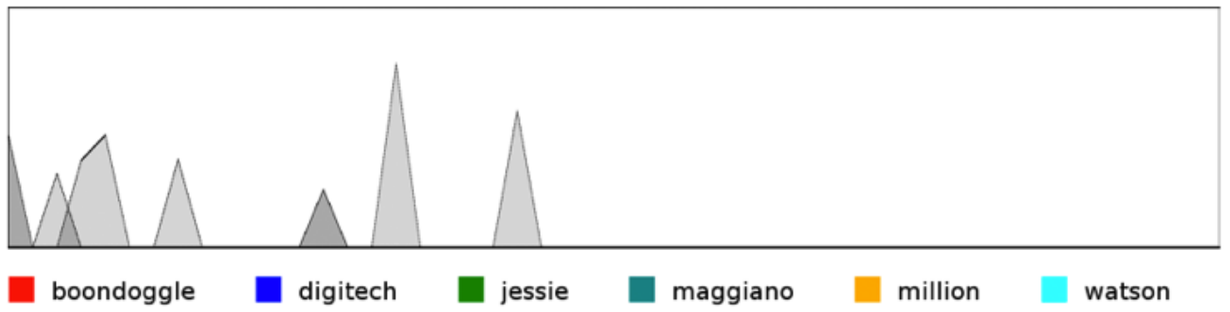
'I never ask advice about growing,' ALICE said indignantly.

'Too proud?' the other inquired.

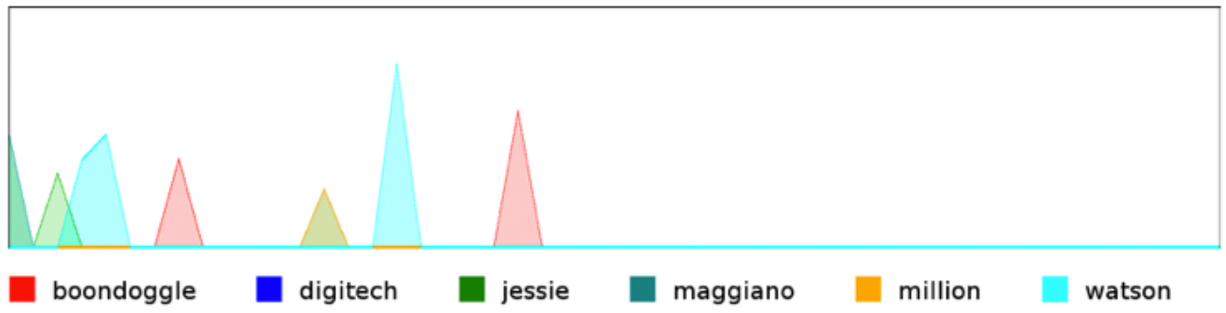
(a) TileBar



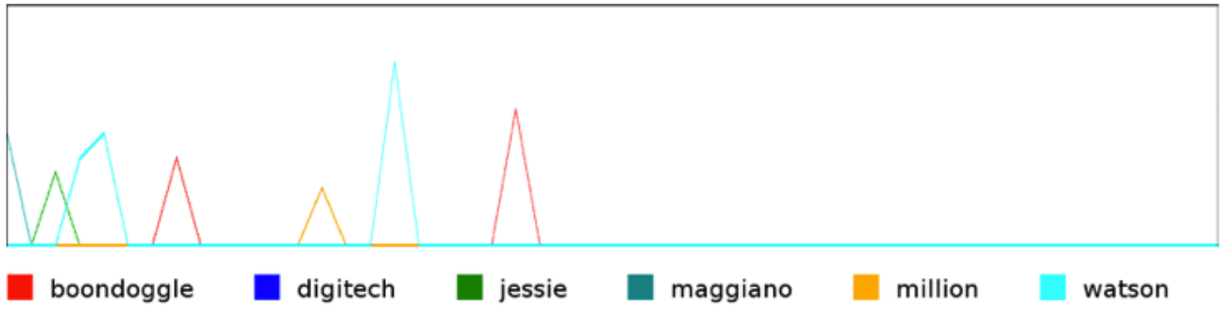
(b) Greyscale Histogram



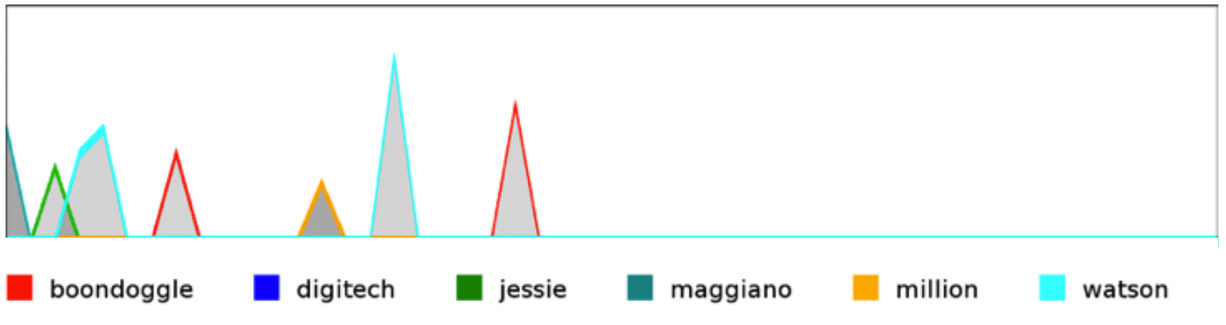
(c) Color Histogram



(d) Line Histogram



(e) Line/Fill Histogram



undefined

Browse Brush Drill down Up a level

THROUGH THE LOOKING-GLASS

by LEWIS CARROLL

THE MILLENNIUM FULCRUM EDITION 1.7

CHAPTER 1

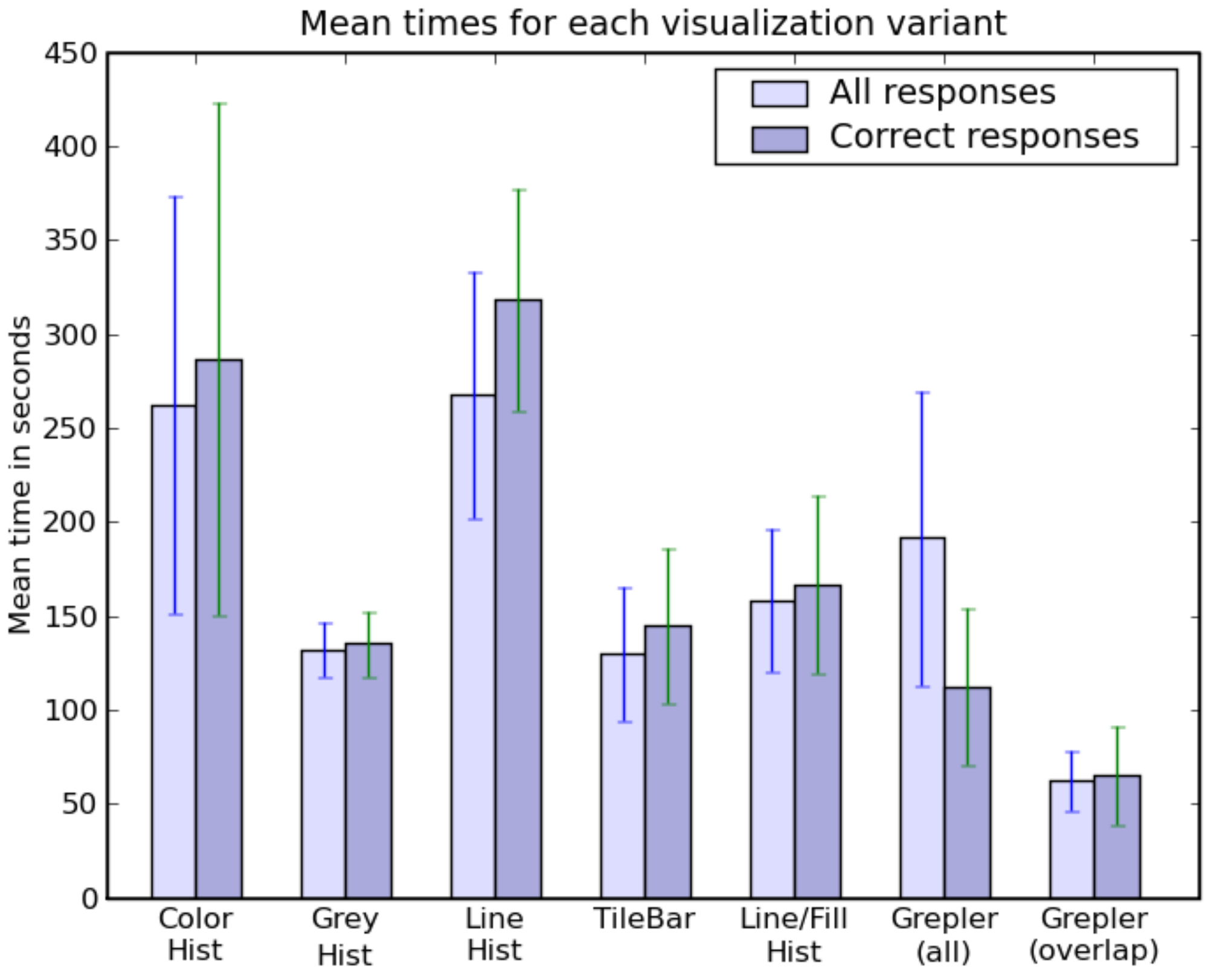
Looking-Glass house

One thing was certain, that the WHITE kitten had had nothing to

User Study

- Show subjects seven documents in the interface, sequentially
- For each subject, randomly assign one visualization to a document
- Ask a quiz question about the document
- Measure time until quiz is answered

Initial Results (5 subjects)



Conclusions

- This visualization model appears effective for information retrieval tasks, including the specific task of digital forensic string search
- Additional user studies are warranted to obtain statistically significant results

References

[1] N. Beebe and G. Dietrich. A New Process Model for Text String Searching. Norwell: Springer, 2007.

[2] N. L. Beebe and J. G. Clark. Digital forensic text string searching: Improving information retrieval effectiveness by thematically clustering search results. In Digital Investigation, volume 4 supplement 1, September 2007.

[3] D. Forte. The importance of text searches in digital forensics. Network Security, pages 13–15, April 2004.

[4] M. A. Hearst. Tilebars: visualization of term distribution information in full text information access. In CHI '95: Proceedings of the SIGCHI conference on Human factors in computing systems, pages 59–66, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co.

[5] K. Mandia, C. Prosise, and M. Pepe. Incident Response & Computer Forensics. McGraw-Hill/ Osborne, California, 2003.

[6] M. Schwartz, C. Hash, and L. Liebrock. Term distribution visualizations with a focus+context model. Technical report, New Mexico Institute of Mining and Technology, 2008. Available at http://cs.nmt.edu/~liebrock/papers/SchwartzHashLiebrock.pdf.
Revised version submitted to ACM Symposium on Applied Computing

A Term Distribution Visualization Approach to Digital Forensic String Search

Moses Schwartz*† and L. M. Liebrock*

*New Mexico Institute of Mining and Technology

†Sandia National Laboratories

Sandia National Laboratories

NEW MEXICO TECH
SCIENCE · ENGINEERING · RESEARCH · UNIVERSITY