Change-Link 2.0: A Digital Forensic Tool for Visualizing Changes to Shadow Volume Data

Timothy R. Leschke
Computer Science and Electrical Engineering
University of Maryland, Baltimore County
Baltimore, MD 21250
tleschk1@umbc.edu

Charles Nicholas
Computer Science and Electrical Engineering
University of Maryland, Baltimore County
Baltimore, MD 21250
nicholas@umbc.edu

ABSTRACT

We present Change Link 2.0, a coordinated and multiple view tool for digital forensics which supports an understanding of how shadow volume data have changed over time. An improvement over the original Change-Link tool [25], Change-Link 2.0 provides an *overview*, a *directory-tree view*, a *directory content view*, and a *metadata view* in a side-by-side, split-screen, linked-view interface that supports easy browsing and detection of files and directories that have changed over time. This data visualization approach supports faster comprehension of digital forensic data, quick detection of anomalous data, and a better understanding of "what happened?." Input to Change-Link 2.0 is an evidentiary hard drive containing multiple versions of files and directories which have been archived by the Microsoft Volume Shadow Copy Service [28]. Our contributions include data visualization techniques that support an overview of the entire dataset, as well as an understanding of how the directory-tree structure, individual directory content, and file and directory metadata have changed over time. Change-Link 2.0, and its predecessor, are the first data visualization tools that we are aware of which support the forensic examination of shadow volume data.

Categories and Subject Descriptors

H.1 [**Models and Principles**]: User/Machine Systems—*Human Information Processing*
; H.5 [**Information Interfaces and Presentation**]: User Interfaces—*Screen Design and User-Centered Design*

Keywords

Coordinated and multiple views, linked view, change over time, overview+detail, digital forensics, data visualization.

VizSec '13 October 14 2013, Atlanta, GA, USA

1. INTRODUCTION

Digital forensics involves the use of scientifically derived and proven methods to support the extraction and analysis of data from digital artifacts such as computers, cell phones, and GPS receivers, for the purpose of facilitating or furthering the reconstruction of events that are found to be criminal or disruptive to planned operations [34]. We design, implement, and demonstrate Change-Link 2.0, a second generation data visualization tool which helps the user identify files and directories that have changed over time. This data visualization approach supports faster comprehension of digital forensic data, quick detection of anomalous data, and a better understanding of "what happened?."

We define a changed file or directory as one that has been added, deleted, or moved within a directory structure, or had its contents or timestamps altered through user or process interaction. We define the primary baseline as the fresh installation of the operating system. Changes to files and directories that have occurred since the baseline are considered to be of higher importance since they are more likely to contain information of evidentiary value.

A secondary baseline is defined in terms of a specific moment in time from which all changes are compared. Consider a computer intrusion investigation in which case the most important data is that which was changed during the time period of the criminal activity. Understanding what has changed within the artifact supports a better understanding of how that device was used to support criminal activity.

Change-Link 2.0 is designed to work with data from the Microsoft Volume Shadow Copy Service (VSS), which is found in Windows Server (2003 and 2008), Windows Vista (2007) and Windows 7 (2009) [28]. This service makes automatic backup copies of user, application and system data to support recovery from accidental data loss and system instability. The repositories in which these backup copies are stored are called *shadow volumes*. There can be as many as 64 active shadow volumes resulting in as many as 64 versions of each file and directory [28].

We have developed an overview+detail data visualization tool with four side-by-side, split screen, linked-view interfaces that support easy browsing and comprehension of shadow volume data. Our contributions include:

- A new tool, Change Link 2.0, for exploring changes to directories and files within shadow volume data.
- The development of a visualization technique that shows how individual files and subdirectories within a select-

ed directory have changed over time.

- The demonstration of how a polar plot can be used to support pattern recognition among directory and file timestamp data.
- The demonstration of how a bar-chart can be used to show changes to directory or file size.
- A general approach for applying data visualization techniques to digital forensic data in order to gain insight.

The rest of the paper is organized as follows: Section 2 presents digital forensic tools that support shadow volume forensics, and four data visualization techniques that support our examination goals. Section 3 explains shadow volume data and presents the test set that is used to evaluate the tool. Section 4 explains the design and implementation of Change-Link 2.0. Sections 5, 6, and 7 present the results, user reactions, and conclusion.

2. RELATED WORK

Change Link 2.0 is a much improved and expanded version of Change-Link [25]. Both versions of this tool make use of the *segmented box and whisker* glyph. [24].[1] Change-Link 2.0, however, adds a directory content view to show how individual files and subdirectories within a directory have changed over time. A polar plot and bar-chart are added to show when a file or directory was accessed and how its size has changed over time.

We investigate digital forensic tools that support shadow volume examinations as well as data visualization techniques that support visualizing large datasets to support an *overview*, visualizing changes in tree structure to support a *directory-tree view*, visualizing time series data to support a *directory content view*, and visualizing radial/axis data to support a *metadata view*.

2.1 Digital Forensic Tools

Many tools have been developed which support the forensic examination of shadow volume data including *EnCase VSS Examiner EnScript* [21], *VSC ToolKit* [13], *ShadowExplorer* [37], *ShadowScanner2* [9], *X-Ways Forensics* [10], *TimeTraveler* [2], and *ProDiscover* [41]. With the exception of TimeTraveler, which makes minimal use of a visual timeline [2], all of the other tools are strictly text-based. Only ShadowScanner2, X-Ways Forensic, ProDiscover, and TimeTraveler provide the ability to identify what has changed between shadow volumes. The other tools only provide access to the data. Identifying change is based on a manual comparison of hash values, which does not scale well beyond the comparison of three shadow volumes. None of the tools provide an overview of the dataset which provides the necessary context for browsing large datasets. None of the tools provide insight into the nature of the change such as changes in timestamps, which implies the accessing of data, and changes in file size, which suggests file modification. Lastly, none of the tools support pattern recognition, which supports data analysis and better comprehension of the dataset.

[1]Citation number eleven of the original Change-Link paper incorrectly cites the source of the segmented box and whisker glyph[25]. This paper corrects that citation.

2.2 Visualizing Change in Large Datasets

The Visual Information Seeking Mantra is "Overview first, zoom and filter, then details on demand" [38]. This mantra is the foundation upon which good browsing and searching data visualization tools are based [39]. An overview of the entire dataset supports easier navigation, and an enhanced comprehension of the data due to an understanding of context. Data visualization techniques that support large datasets must make efficient use of the screen real estate. We are unaware of any visualization techniques that use the display space more efficiently than the *TreeMap* [19].

TreeMaps with millions of nodes have been developed [40]. The ability to display large amounts of data is enhanced by high-resolution monitors [7]. Additional dimensions can be added through the use of color and texture (e.g., orientation, size, and contrast) [15] [45] [43]. Effective aggregation of data has been demonstrated with the *ManyEyes* TreeMap tool that has been used to summarize the "U.S. Prison Population by State" [17]. Although no other data visualization technique supports the display of more data points than TreeMaps, TreeMaps do not support an understanding of individual time periods, which is an essential requirement. Many other tools have been developed for visualizing large data sets, but they are domain specific and inappropriate for our goals. Thus, a new technique is needed.

2.3 Visualizing Change in Tree Structures

Several visualization tools have been developed to support the comparison of two or more trees. *TreeJuxtaposer* [32], *TaxoNote* [31], *Point-set View* [1], and *TreeWiz* [36] are used by biologists to compare taxonomic trees that describe, identify, and classify living things, and phylogenetic trees that organize living things such that one can understand how species are related. *TreeVersity* was developed to compare tree representations of the U.S. Federal Budget [12]. Organization hierarchies of businesses are compared with *TimeTree* [5]. *CandidTree* compares file system trees [23]. Although each of these tools is effective when applied in its own domain, until now, only Change-Link is suitable for showing the differences between multiple shadow volumes.

The three ways to convey change between two or more trees are *juxtaposition*, *superposition*, and *explicit difference* or *aggregation* [11]. Juxtaposition is the displaying of the trees in a side-by-side view in order to support the comparison of similarities and differences (i.e., the net change). Examples of juxtaposition include TreeJuxtaposer [32], TaxoNote [31], and TreeVersity [12].

Superposition is the technique whereby several states are expressed either at the same time or at the same location. Examples include CandidTree [23], TreeVersity [12], Change-Link [25], and almost any visualization that makes use of a consensus tree. A consensus tree, by its very nature, is an example of superposition because it expresses the states of several trees at the same time. Coloring is used to identify the similarities and differences among trees. CandidTree colors nodes blue if they have been added and red if they have been deleted [23]. TreeVersity outlines the nodes in black if they are deleted and white if they are added [12].

Explicit difference and aggregation express complimentary approaches to the same goal. Explicit difference expresses the difference in change between data points whereas aggregation expresses the sum of the change between data points. The *DiffTree* of TreeVersity is a good example of these prin-

ciples [12].

2.4 Visualizing Change in Time Series Data

The purpose of time series analysis is to "obtain insight into phenomena," "discover repetitive patterns," and "predict the future" [44]. Insight into phenomena through time series analysis has been shown regarding the renaming of a directory [25] and an understanding of daily sunshine intensity [46].

Discovering a repetitive pattern helps predict future events and supports an understanding of *intent*, which is very important in a criminal prosecution. For example, being able to identify a pattern that shows an individual routinely accessed child pornography shows an interest in illicit images and an intent to violate this particular statute.

Many tools have been developed to support time series analysis such as *ProcessLines* [26], *ThemeRiver* [14], a tool that we call *Cluster and Calendar* [44], *Cyber Forensic TimeLab* [33], *LifeFlow* [47], *LifeLines* [35], *XTG* (Time-line Display Generator for X-Windows) [20], and *TimeSlice* [49]. Of particular interest to us is CyberForensic TimeLab which supports an overview of the timestamp data from the files of a hard drive [33]. By understanding the type and quantity of files that were accessed during a particular time period, insight into the data is achieved.

LifeFlow has been applied to health care to support an understanding of "bounce backs," which are patients that are returned to a higher level of care after being decreased to a lower level of care within a hospital [47]. LifeLines has been applied to medial patient records to reveal phenomena such as allergic reactions, rashes, complaints, and changes in patient health that can be attributed to a previous event such as the administration of a drug or treatment [35]. XTG helps users understand network delays due to remote connection requests, maintenance, slow software execution, heavy usage, and time-outs because of network failures [20]. TimeSlice supports a historical perspective of famous people, including their gender, profession (Engineer, Philosopher, Politician, etc.), and continent of origin [49].

2.5 Radial/Axis Visualization Techniques

A radial (axis) visualization technique uses an elliptical fashion to layout the data (single circle, concentric circles, nested circles, spirals, etc.) [8]. Radial visualization techniques are a natural way to visualize an expanding dataset as is done with the Hyperbolic Browser [22]. Another strength is its ability to represent patterns in "serial periodic data" [8]. Seven design patterns for radial visualizations have been identified. They are divided into three groups: polar plots (including *tree* and *star* plots), space-filling plots (including *concentric*, *spiral*, and *Euler* plots), and ring plots (including *connected* and *disconnected* plots) [8].

Polar plots are characterized as having a center with some special meaning or semantic significance, such as a root node, origin of a coordinate system, or search term from a query [8]. Tree plots and star plots are both classes of polar plots that have line segments that radiate from the center node. Whereas tree plots allow for branching within the structure, star plots do not. Examples of tree plots include the Hyperbolic Browser [22], and MoireGraphs [18]. An example of a star plot includes Starstruck [16], which represents a document as the central node in the star, and the themes in the document are represented as lines that radiate from the center. The length of each ray represents how closely each theme is related within the document [16].

InterRing is a space-filling visualization that uses concentric circles to represent the hierarchical structure of a node tree [48]. Siblings are represented as partitions within the same ring; children are represented as partitions within the corresponding section of the inner ring; and parent nodes are represented as corresponding sections of the outer ring.

An unnamed space-filling visualization that uses an Archimedes spiral, and twelve evenly spaced radials to represent months, shows the consumption of the flowering Baphia Capparidifolia tree by chimpanzees in Tanzania [6]. One revolution of the spiral represents a year. Well-defined patterns show that March, April, May and October are the months of greatest consumption.

An Euler pattern is characterized by a set of circles placed either inside or next to a larger circle such that each smaller circle is understood to be the child node of its neighboring parent node. MoireTrees use an Euler pattern to represent a hierarchy [30].

TimeWheel is a connected ring whereby nodes located along the ring edge are connected by lines to points along the horizontal time axis located at the center of the wheel [42]. The user is allowed to rotate the wheel so that the visual clutter caused by the plotted lines is diminished.

An unnamed disconnected ring tool has been developed which supports queries of large tabular data such as polling data [8]. By allowing the user to interact with the polling data in real time, the user can draw their own conclusions and avoid the editorial predisposition of the polling agency.

3. SHADOW VOLUME DATA

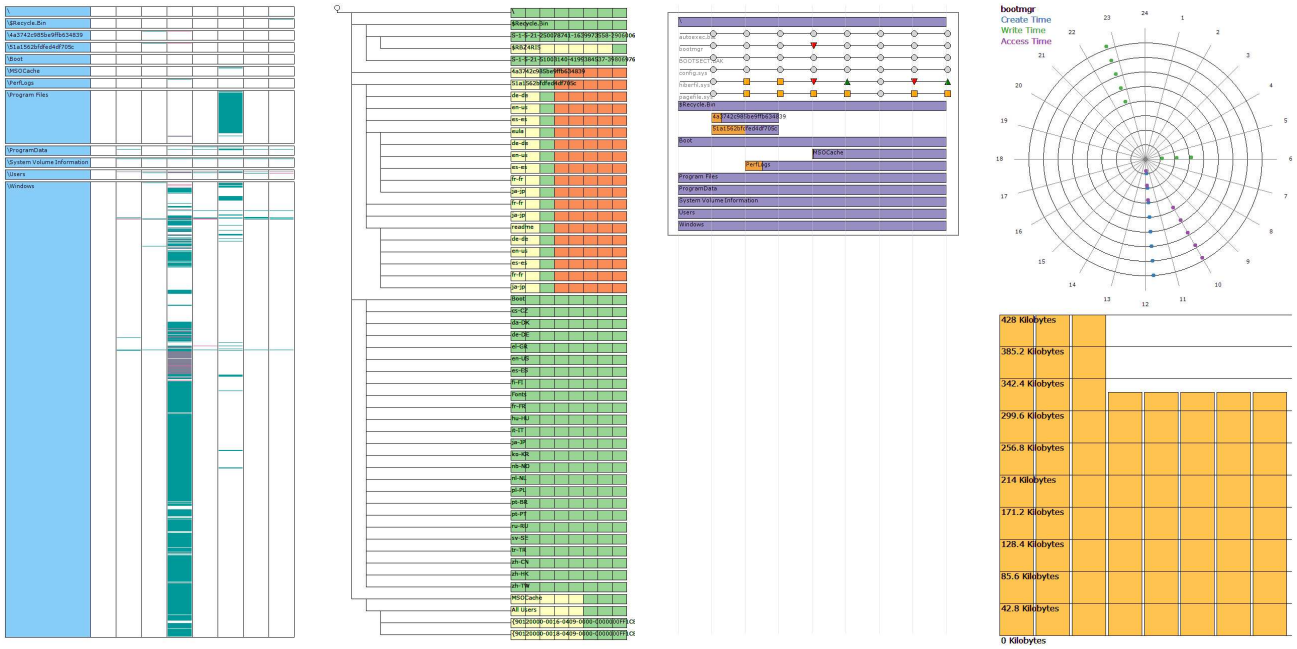
In testing Change-link 2.0, we used the same dataset as before which are backup copies of user, application and operating system data that are maintained by the Microsoft Volume Shadow Copy Service [28]. As many as 64 active versions of the data are maintained in repositories known as *shadow volumes*. Up to 512 shadow volumes can be maintained through an archiving process [28]. This data can be examined by law enforcement to recover evidence of previous criminal activity such as images of child exploitation.

We created a dataset that simulates the characteristics of actual digital forensic data. We are working with law enforcement to test Change-Link 2.0 with real case evidence.

The test dataset comprises eight shadow volumes that are established at the following times, respectively:

1. after installing Vista (but before activation)
2. after activating Vista
3. during the installation of Service Pack 1
4. after installing Service Pack 1
5. after installing Microsoft Office
6. after activating Microsoft Office
7. after creating user directories
8. after deleting user directories

The third shadow volume was created automatically by the operating system during the installation of Service Pack 1. The other seven shadow volumes were created manually by establishing a restore point, which triggers the archiving process of the Volume Shadow Copy Service. Each of these eight shadow volumes contains a snapshot of the files and directories within the directory-tree structure at the time the shadow volume was created.



(a) Overview (b) Treeview (c) Directory Content (d) Metadata

Figure 1: The Change-Link 2.0 user interface is comprised of four linked views. The *Overview* window supports quick browsing by highlighting locations where data has changed within each time period. The *TreeView* window shows how the directory-tree structure has changed over time. The *Directory Content* window shows how all of the files and directories within a selected directory have changed over time. The *Metadata* window uses a polar plot and bar graph to reveal patterns of change for a selected file or directory.

4. DESIGN AND IMPLEMENTATION

Change-Link 2.0 is implemented with approximately 8,500 lines of C# code. This language was chosen because of its availability in our research lab, and because of its support of the .NET framework. The .NET framework was chosen because of its support of Windows Forms, SQLite database references, and easy extraction of directory metadata from the dataset being explored.

Change-Link 2.0 is executed by means of the PsExec tool, which elevates its directory and file access rights to "System" level [29]. This allows the tool to parse all of the data on a hard drive, including those directories and files that are locked to programs with only "User" level access rights.

A drop-down menu allows a user to direct Change-Link 2.0 to process the evidence hard drive, which is connected to the examination station via a write-block device in order to prevent the original evidence from being altered. Change-Link 2.0 uses the "vssadmin list shadows" command to identify the shadow volumes on the evidence hard drive, and the "mklink" command to create dynamic links to each shadow volume. Change-Link 2.0 performs a concurrent breadth-first traversal of the directory trees found in each shadow volume. Data such as name, size, write time, access time, creation time, and path for each file and directory are extracted and saved in a SQLite database. Calculations that support the rendering of the Change-Link 2.0 visualizations are made and saved in a second SQLite database. The visualizations that are created by Change-Link 2.0 are drawn from the data that are found in these SQLite databases.

The Change-Link 2.0 user interface is partitioned into four equal windows which display four different semantically zoomed representations of the data set: an *overview*, a *directory-tree view*, a *directory content view*, and a *metadata view* (Figure 1). This paper emphasizes our recent contributions which are the directory content view (Figure 1(c)) and the metadata view (Figure 1(d)). The overview and directory-tree view are presented for completeness.

4.1 Overview Window

The Change-Link 2.0 overview window is comprised of rectangle glyphs, one for each of the directories that exist at the root of the evidence hard drive (Figure 1(a)). These glyphs are labeled with the name of the directory which they represent. The height of each glyph represents the number of files and directories that are found within that directory's sub-structure. Each glyph is partitioned vertically according to the number of time periods.

Within each partition are horizontal color-coded lines that represent the locations within that directory's sub-structure where the "change" has taken place. The change that is identified within the current implementation consists of only the creation or deletion of a file. These lines are colored green and red respectively. Future implementations are expected to highlight other events such as directory creation and deletion, file growth and reduction, as well as highlight specific file types such as Word, PDF, JPEG and others.

The Overview window of Change-Link 2.0 is an improvement over the previous work in that this version includes a glyph labeled "\" which contains all of the files (if any) that are found at the root of the evidence hard drive. This ver-

sion also uses an easier to read font (Verdana) and a more pleasing color pallet.

4.2 Directory-tree View

The directory-tree view window of Change-Link 2.0 (Figure 1(b)) shows a semantically zoomed-in view of the portion of the directory structure that has been selected for browsing by clicking on one of the green or red "change lines" found in the overview (Figure 1(a)). Color-coded segments of the segmented box and whisker glyphs correspond to time periods in which the corresponding directory *does not yet exist* (yellow), *does exist* (green), and *once existed but has been deleted* (orange). By aligning the segments in columns, the user is able to see how the directory-tree exists at each period of time, and through comparison, see how the directory structure changes over time.

In response to the original Change-Link usability study [25], vertical lines are added to connect the whiskers of the segmented box and whisker glyphs of Change-Link 2.0 in order to make the parent-child relationships among directories more explicit. ColorBrewer2 is used to select segment colors that are more appealing and do not conflict with colors of the neighboring windows [3].

4.3 Directory Content View

The directory content window of Change-Link 2.0 (Figure 1(c) and Figure 2) shows a semantically zoomed-in view of the directory that has been selected for browsing by clicking on a segmented box and whisker glyph found in the preceding tree view window (Figure 1(b)). The semantically zoomed-in representation consists of a *directory content box* outlined by a thin black line that conveys the container-like quality of the directory being viewed. The width of the box spans the entire visualization window, which uses faint vertical lines to partition the space into a sequence of time periods that are defined by the shadow volumes of the underlying dataset. Although the horizontal lines are evenly spaced for easier viewing, the duration of time that is represented by each partition can vary from seconds to days, depending on the duration of the underlying shadow volumes.

The directory content view is comprised of directory glyphs and file glyphs. Directory glyphs are purple rectangles that span the time period for which the directory is known to exist. The name of the directory is written along the left edge of the inside of the glyph. The gap between the preceding shadow volume line and the left edge of the directory glyph is colored yellow. Yellow is used to color the gap between when the directory exists and when the preceding shadow volume was created. A "large" gap suggests that the directory may have been made by a manual process, whereas a "small" or missing gap suggests the directory was created by an automated process - because only automated processes are fast enough to create directories within a few seconds of a shadow volume. This coloring technique helps establish user intent by distinguishing between user behavior and automated process behavior.

Along the top internal edge of the directory content box is a directory glyph that represents the directory that is being visualized within this window. The directory glyphs that appear at the bottom of the directory content box represent sub-directories of the directory being browsed. All of the ancestor directories, up to the root of the directory tree, are represented by directory glyphs that appear outside and above the directory content box (not shown).

File glyphs are thin, horizontal, black lines with shapes plotted where the file line intersects with the faint, vertical, shadow volume lines. Gray circles represent that the associated file has not changed since the previous time period. Yellow squares represent that some or all of the timestamps for the file have changed since the previous time period. Green and red triangles that point up or down respectively represent that the file changed in size in the corresponding direction. It is understood that any change in size is also accompanied by a change in the file's *accessed* and/or *written* timestamps.

The name of the associated file is written just below the file glyph line, beginning at the left edge of the glyph.

4.4 Metadata View

The last of the four visualization windows provides the file and directory metadata view (Figure 1(d)) This is a combination of a polar plot and a bar-chart which visually represent the changes to the selected file or directory timestamp and size information. Files and directories are selected by clicking on the file glyph within the previous directory content view window (Figure 1(c)) . The name of the selected file appears at the top of the metadata view window.

The polar plot represents a twenty-four hour day as one revolution around the circle. Each concentric circle represents a different shadow volume. Although shadow volumes span different time periods, only the most recent change to a file or directory is recorded in each shadow volume. The most recent shadow volume is represented by the outermost ring. The *creation time*, *write time*, and *access time* for the file or directory is plotted on each shadow volume ring as a blue, green, or purple dot respectively. This supports the identification of patterns of use.

In order to avoid occlusion caused by plotting two or more points at the exact same location, a technique called jittering is used [27]. This technique is applied by plotting creation times directly on the concentric ring, write times are plotted just inside the concentric ring, and access times are plotted just outside the concentric ring. This allows for each of the three data series to be plotted on the same polar plot without causing occlusion. This saves valuable screen real estate by using one polar plot rather than three.

Below the polar plot is a bar-chart that shows the size of the file or directory as recorded within each shadow volume. The size records are displayed from left to right, from the oldest to the most recent. Ten horizontal lines, in 10% increments of the largest size value, are superimposed over the bar-chart to denote the scale of the graph. The units of the graph (B, KB, GB, etc.) are automatically adjusted to ensure bars of maximum height are drawn for easy viewing.

5. RESULTS

Results for the Overview and TreeView visualization techniques have been reported previously [25]. This paper provides results for the Directory Content and Metadata visualization techniques.

Figure 1(c) and the zoomed-in view provided by Figure 2 provide visualizations of the root directory. The directory content box contains eighteen glyphs, which we number from top to bottom to support this discussion. The first is a purple directory glyph that represents the root directory ("\").

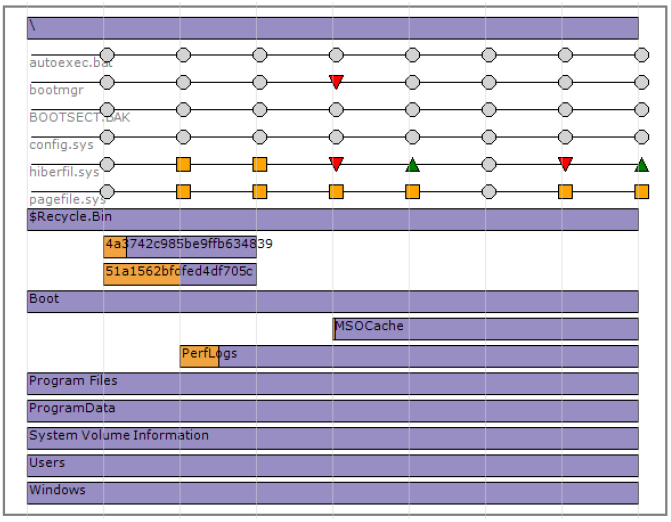


Figure 2: The Directory Content glyph showing the root directory ("\") with 6 files and 11 sub-directories that span 8 time periods.

Figure 3: The hiberfil.sys file glyph showing the file was accessed during the 2nd and 3rd time periods (yellow squares), decreased in size during the 4th and 7th time periods (red triangles), and increased in size during the 5th and 8th time periods (green triangles).

Glyphs 2-7 are file glyphs, and glyphs 8-18 are sub-directory glyphs. The horizontal length of these glyphs corresponds to the time periods for which each file or sub-directory exists.

5.1 File Glyph

Figure 3 provides a zoomed-in view of the 6th glyph, which represents the "hiberfil.sys" file. The thin black line represents the time period for this file's existence. The 1st shape (from left to right) is always a gray circle, which represents no change has taken place yet. The yellow squares denote the hiberfil.sys file was accessed (most likely read), but not changed in size during the 2nd and 3rd time periods. The red triangles that point down represent the hiberfil.sys file decreased in size during the 4th and 7th time periods. The green triangles that point up represent that the hiberfil.sys file increased in size during the 5th and 8th time periods. The gray circle located at the end of the 6th time period denotes that no change took place during that time period. Viewed collectively, these colored shapes provide an understanding of when and how the hiberfil.sys file changed over time. By comparing this file glyph to other file glyphs, correlations between when files have changed become possible. This provides insight into the data, and helps direct the examination towards data of interest.

5.2 Directory Glyph

Figure 4 provides a zoomed-in view of the directory named "4a3742c985be9ffb634839." There are faint vertical lines that divide the background into eight time periods. The glyph is colored purple for the time period for which the directory exists. The directory exists for part of the 2nd time period and all of the 3rd time period. This visualization also implies the directory was deleted sometime during the 4th time period. We are redesigning the directory glyph to make this deletion time more explicit. Yellow identifies the gap between the directory and the preceding shadow volume (as explained in Section 4.3).

Figure 4: The 4a3742c985be9ffb634839 directory glyph showing the directory existed for part of the 2nd time period, and all of the 3rd time period.

5.3 Metadata

The 3rd glyph from the top in Figure 2 represents the "bootmgr" file. The red triangle pointing down shows that this file decreased in size during the 4th time period. By clicking on this glyph, the bar-chart of the Metadata view of Figure 1(d) shows the "bootmgr" file decreased from 428 kilobytes to about 320 kilobytes. The blue dots of the polar plot (Figure 5) show the "bootmgr" file was created at about 11:45 AM. The purple dots show the initial access time-stamp for the "bootmgr" file is about 11:45 AM, which changes to about 10:00 AM during the 4th time period. The green dots show the initial write-time for the "bootmgr" file is about 5:50 AM, which changes to about 10:45 PM during the 4th time period. Visualizing the change in file size with the bar-chart allows the user to better understand when, and by what degree, the selected file (or directory) changed in size. This polar plot supports pattern detection, which allows the user to better understand when a file (or directory) was accessed and/or changed.

6. USER REACTIONS

We conducted an informal usability study with five digital forensic examiners with different levels of experience. Each participant answered 13 questions correctly to demonstrate they understand the visualizations being presented by Change-Link 2.0. Open-ended questions were asked to elicit feedback. A more formal usability study is planned for future work.

The comments that were received include the following: Change-Link 2.0 provides a "narrative that can be told at a glance" and "helps me narrow my search." By comparing data from different shadow volumes, a "user can discover data that they may have missed with other tools." Change-Link helps the user be more productive by allowing one to view data in a "quicker manner than other tools provided." Change-Link also supports better comprehension because "patterns and relationships with other files and directories become more apparent."

Change-Link 2.0 allows one to perceive change at the *file*, *directory*, and *metadata* levels, which is an appreciated ability that is not supported in the previous version.

Each participant was asked to evaluate the usability of Change-Link 2.0 based on the standard 10 question System Usability Scale [4]. The average score is 81. Given that the lowest scores were received from users with the least familiarity with data visualization and this tool, we think we can improve this score through training.

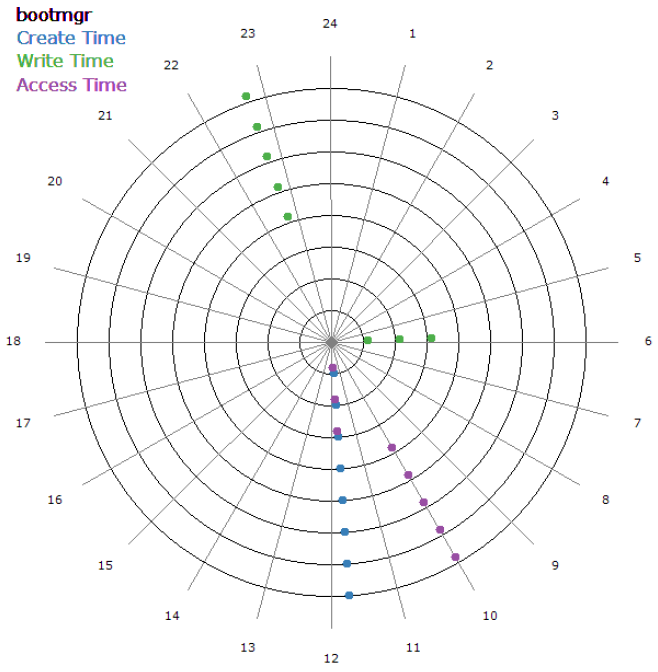


Figure 5: The Polar Plot showing the "bootmgr" file creation timestamps (blue dots), write timestamps (green dots), and access timestamps (purple dots).

7. CONCLUSION

Change-Link 2.0 provides an understanding of change within shadow volume data. This supports faster comprehension of digital forensic data, quick detection of anomalous data, and a better understanding of "what happened?."

Acknowledgments

We are grateful to David Andescavage, Bryce Blair, Eoghan Casey, Emil Kahihikolo, and Axil Valentin for participating in our usability study. We also thank Jian Chen for her helpful comments and ongoing mentoring.

8. REFERENCES

[1] N. Amenta and J. Klingner. Case study: Visualizing sets of evolutionary trees. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis'02)*, pages 71–74. IEEE Computer Society, 2002.

[2] BearsOnTheLoose. TimeTraveler. `http://www.bearsontheloose.com/`, 2012.

[3] C. Brewer and M. Harrower. ColorBrewer 2.0. `http://colorbrewer2.org/`, 2013.

[4] J. Brooke. SUS - A quick and dirty usability scale. `http://hell.meiert.org/core/pdf/sus.pdf`, 1986.

[5] S. K. Card, B. Suh, B. A. Pendleton, J. Heer, and J. W. Bodnar. TimeTree: Exploring time changing hierarchies. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*, pages 3–10. IEEE Computer Society, 2006.

[6] J. V. Carlis and J. A. Konstan. Interactive visualization of serial periodic data. In *Proceedings of the 11th annual ACM symposium on User interface software and technology*, pages 29–38. ACM, 1998.

[7] CNET. Apple thunderbolt display. `http://reviews.cnet.com/lcd-monitors/apple-thunderbolt-display/4505-3174_7-34850107.htm`, 2013.

[8] G. M. Draper and R. F. Riesenfeld. Who votes for what? A visual query language for opinion data. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1197–1204, 2008.

[9] EKLSoftware. ShadowScanner manual (v2.0). `http://www.shadowscanner.com`, 2012.

[10] S. Fleishmann. X-ways forensics/winhex. `http://www.x-ways.net/winhex/manual.pdf`, 2012.

[11] M. Gleicher, D. Albers, R. Walker, I. Jusufi, C. D. Hansen, and J. C. Roberts. Visual comparison for information visualization. *Information Visualization*, 10(4):289–309, 2011.

[12] J. A. G. Gomez, A. Buck-Coleman, M. L. Pack, C. Plaisant, and B. Shneiderman. TreeVersity: Interactive visualizations for comparing hierarchical datasets. In *Proceedings of the Transportation Research Board 92th annual meeting*, pages 1–22. The National Academies, 1995.

[13] J. Hale. VSC toolset: A GUI tool for shadow copies. `http://dfsteam.blogspot.com/2012/03/vsc-toolset-gui-toolset-gui-tool-for-shadow-copies.html`, 2012.

[14] S. Havre, E. Hetzler, P. Whitney, and L. Nowell. ThemeRiver: Visualizing thematic changes in large document collections. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):9–20, 2002.

[15] C. Healey and J. Enns. Large datasets at a glance: Combining textures and colors in scientific visualization. *IEEE Transactions on Visualization and Computer Graphics*, 5(2):145–167, 1999.

[16] B. Hetzler, P. Whitney, L. Martucci, and J. Thomas. Multi-faceted insight through interoperable visual information analysis paradigms.

[17] IBM. U.S. prison population by state. `http://www-958.ibm.com/software/analytics/manyeyes/visualizations/us-prison-population-by-state-5`, 2013.

[18] T. J. Jankun-Kelly and K.-L. Ma. MoireGraphs: Radial focus+context visualization and interaction for graphs with visual nodes. In *IEEE Symposium on Information Visualization (INFOVIS 2003)*, pages 59–66. IEEE, 2003.

[19] B. Johnson and B. Shneiderman. Tree-maps: A space-filling approach to the visualization of hierarchical information structures. In *Proceedings of the IEEE Visualization '91 Conference.*

[20] G. M. Karam. Visualization using timelines. In *Proceedings of the 1994 ACM SIGSOFT international symposium on Software testing and analysis*, pages 125–137. ACM, 1994.

[21] S. Key. Examining volume shadow copies - the easy way! `http://encase-forensic-blog.guidancesoftware.com/2012/06/examining-volume-shadow-copies-easy-way.html`, 2012.

[22] J. Lamping, R. Rao, and P. Pirolli. A focus+context technique based on hyperbolic geometry for visualizing large hierarchies. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computer Systems*, pages 401–408. ACM Press, 1995.

[23] B. Lee, G. G. Robertson, M. Czerwinski, and C. S. Parr. CandidTree: Visualizing structural uncertainty in similar hierarchies. `http://hcil2.cs.umd.edu/trs/2008-21/2008-21.pdf`, 2013.

[24] T. Leschke, P. Rheingans, and A. T. Sherman. Using a fisheye view to visualize change-over-time in support of digital forensic examinations. In *Proceedings of Government Forum for Incident Response and Security Teams (GFIRST)*, pages 1–12. GFIRST, 2011.

[25] T. R. Leschke and A. T. Sherman. Change-Link: A digital forensic tool for visualizing changes to directory trees. In *Proceedings of the Ninth International Symposium on Visualization and Cyber Security*, pages 48–55. ACM, 2012.

[26] X. Lou, F. Tian, W. Liu, D. Teng, G. Dai, and H. Wang. Visualizing time-series data in processlines: design and evaluation of a process enterprise application. In *Proceedings of the 2010 ACM Symposium on Applied Computing*, pages 1165–1172. ACM, 2010.

[27] J. Mackinlay. Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics*, 5(2):110–141, 1986.

[28] Microsoft. Volume shadow copy service. `http://technet.microsoft.com/en-us/library/ee923636(28v=ws.10)29.aspx`, 2011.

[29] Microsoft. PsExec. `http://technet.microsoft.com/en-us/sysinternals/bb897553.aspx`, 2013.

[30] M. J. Mohammadi-Aragh and T. J. Jankun-Kelly. MoireTrees: Visualization and interaction for multi-hierarchical data. In *Proceedings of EuroGraphics, IEEE VGTC Symposium*, pages 59–66. IEEE, 2005.

[31] D. R. Morse, N. Ytow, D. M. Roberts, and A. Sato. Comparison of multiple taxonomic hierarchies using taxonote. In *IEEE Symposium of Information Visualization*, pages 19–21. IEEE Computer Society, 2003.

[32] T. Munzner, F. Guimbretiete, S. Tasiran, L. Zhang, and Y. Zhou. TreeJuxtaposer: Scalable tree comparison using focus+context with guaranteed visibility. In *ACM Transactions on Graphics*, pages 453–462. ACM Press, 2003.

[33] J. Olsson and M. Boldt. Computer forensic timeline visualization tool. *Digital Investigation*, 6:78–87, 2001.

[34] G. Palmer. A road map for digital forensic research. `http://www.dfrws.org/2001/dfrws-rm-final.pdf`, 2013.

[35] C. Plaisant, R. Mushlin, A. Snyder, J. li, D. Heller, and B. Shneiderman. LifeLines: Using visualization to enhance navigation and analysis of patient records. In *Proceedings of the American Medical Informatic Association Annual Fall Symposium*, pages 76–80. AMIA, 1998.

[36] U. Rost and E. Bornberg-Bauer. TreeWiz: Interactive exploration of huge trees. *Bioinformatics*, 18(1):109–114, 2002.

[37] ShadowExplorer. Manual. `http://www.shadowexplorer.com/documentation/manual/html`, 2012.

[38] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualization. In *Proceedings of IEEE visual languages*, pages 336–343. IEEE, 1996.

[39] C. Springfels. Interactive information visualization of a million items. `http://www.cs.umd.edu/hcil/VisuMillion/`, 2013.

[40] C. Springfels. Visualization. `http://www.cs.umd.edu/hcil/research/visualization.shtml`, 2013.

[41] TechPathways. Collection and differential analysis of volume shadow copy with prodiscover (updated for windows 8). `http://www.techpathways.com/webhelp/Collection_and_Differential_Analysis_of_Volume_Shadow_Copy.htm`, 2012.

[42] C. Tominski, J. Abello, and H. Schumann. Axes-based visualizations with radial layouts. In *Proceedings of the 2004 ACM symposium on Applied computing (SAC '04)*, pages 1242–1247. ACM, 2004.

[43] Y. Tu and H.-W. Shen. Visualizing changes of hierarchical data using treemaps. In *Proceedings of the IEEE Transactions on Visualization and Computer Graphics*, pages 1286–1293. IEEE Computer Society, 2007.

[44] J. J. van Wijk and E. R. Selow. Cluster and calendar based visualization of time series data. In *Proceedings of the 1999 IEEE Symposium on Information Visualization (INFOVIS '99)*, pages 4–9. IEEE Computer Society, 1999.

[45] C. Ware and W. Knight. Using visual texture for information display. *ACM Transactions on Graphics*, 14(1):3–20, 1995.

[46] M. Weber, M. Alexa, and W. Muller. Visualizing time-series on spirals. In *Proceedings of the IEEE Symposium on Information Visualization (INFOVIS 2001).*, pages 7–13. IEEE Computer Society, 2001.

[47] K. Wongsuphasawat, J. A. G. Gomez, C. Plaisant, T. D. Wong, B. Shneiderman, and M. Taieb-Maimon. LifeFlow: Visualizing an overview of event sequences. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*, pages 1747–1756. ACM, 2011.

[48] J. Yang, M. O. Ward, and E. A. Rundensteiner. InterRing: an interactive tool for visually navigating and manipulating hierarchical structures. In *IEEE Symposium on Digital Object Identifier*, pages 77–84. IEEE, 2002.

[49] J. Zhao, S. M. Drucker, D. Fisher, and D. Brinkman. TimeSlice: Interactive faceted browsing of timeline data. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pages 433–436. ACM, 2012.