Reconstructing and Visualizing Evidence of Artifact from Firefox SessionStorage

Matsumoto, Shinichi
Institute of Systems, Information Technologies and Nanotechnologies

Onitsuka, Yuya
Department of Informatics, Faculty of Information Science and Electrical Engineering, Kyushu University

Kawamoto, Junpei
Department of Informatics, Faculty of Information Science and Electrical Engineering, Kyushu University

Sakurai, Kouichi
Department of Informatics, Faculty of Information Science and ElectricalEngineering, Kyushu University

出版情報：Lecture Notes in Computer Science. 8909, pp.83-94, 2015. Springer International Publishing
バージョン：
権利関係：

Reconstructing and Visualizing Evidence of Artifact from Firefox SessionStorage *

Shinichi Matsumoto[1,2], Yuya Onitsuka[2],
Junpei Kawamoto[2], and Kouichi Sakurai[2]

[1] Institute of Systems, Information Technologies and Nanotechnologies,
2-1-22, Momochihama, Fukuoka, Japan,
smatsumoto@isit.or.jp,
[2] Department of Informatics, Faculty of Information Science and Electrical Engineering,
Kyushu University, Fukuoka, Japan
onitsuka@inf.kyushu-u.ac.jp
kawamoto@inf.kyushu-u.ac.jp
sakurai@csce.kyushu-u.ac.jp

Abstract. Importance of digital forensics is expected to increase in the future.Many of researches on digital forensics are targeted to persistent memory. These researches concerns about the extraction of evidence directly or via filesystem. On the other hand, there is a movement to employ the Web browser supports HTML5 as software platform. In this situation, it is considered that the forensics techniques for extracting evidences from HTML5 browser is important.
In this paper, we experimented to retrieve the artifacts left by Web-Storage feature for the Web browser for personal computer from the file system. In addition, we implemented a tool that constructs and visualizes the evidence from the artifacts.

Keywords: Computer Forensics, Mobile Forensics, Web browser, Privacy

1 Introduction

1.1 Background

PC and other digital devices that can connect to network into commodity goods. However, users who don't have enough information literacy has utilized these devices on a daily basis. Various activities that have been performed in the real world so far, are now come through the network that they connect by a terminal.

As a result, criminal acts have also moved to the network. Therefore, to investigate these acts, research of evidence on the network or on computer terminal has become essential. These activities of investigation of evidence focus not only on criminal acts, but also corporate governances and litigation of business [1]. With regard to these evidence investigation, the following three points are cited considering the characteristics of digital data.

* The first author's work is supported by JSPS KAKENHI Grant Number 26330169.

– Retrieve data as an evidence.
– Find and summarize the relationship between the data.
– Certify that the data has not been tampered.

In order to keep the effectiveness of evidence acquired in the survey process, it is necessary to pay attention to the above. It is possible that for this purpose, to take advantage of a number of tools, including those from open source that can be used free of charge .

1.2 Motivation

With the spread of HTML5, Web browser is becoming the platform to running applications.

These applications expand the functionality and improve the usability of the terminals. It can be regarded as comparable to native applications, in the viewpoint of the operating speed and usability.

These applications are described in languages relevant to Web technology and aims at high portability and development efficiency. Mobile platforms encompass execution environment which can run both styles of application.

Furthermore, some of software platforms are based on the Web browser supporting HTML5 as an application execution environment. Table 1 summarizes the software platforms from the viewpoint of the application execution environment. It includes platforms that have not been released.

Table 1. Relationship Between Mobile Terminal Platform and Web Technologies

Platform	Dedicated Apps. Support	Web Apps. Support
Android	✓(Dalvik bytecode)	✓
iPhone (iOS)	✓(Objective-C)	✓
WindowsPhone	✓(CLR)	✓
Tizen	✓(C++)	✓
Firefox OS	Not Supported	✓
Ubuntu OS	✓(Qt & Javascript)	✓
Sailfish OS	✓(C++ & Qt)	✓
Chrome OS	Not Supported	✓

1.3 Related Works

Forensic research on Web browsers is well developed especially for private browsing mode and portable browsers. Private mode of Web browsers is provided for privacy against the network and privacy against local machines. The former one prevents identification of the user over the network. On the other hand, the later one prevent leaving the evidence on the terminal (local machine). Portable Web

browser is the browser that is primarily designed to be installed on a removable disk (e.g. USB flash drive). These browsers can be used if the user does not want to leave the evidences of a browsing activity on personal computer terminal primarily.

Donny [2] examined private mode of Web browsers mainly with memory forensic techniques. As the experimental result shown, in all private mode and portable browsers, evidences are left. Left evidences are residuals of Web browsing history, e-mail address, browsed pictures. In addition , for some browsers, browsed movie is acquired from main memory. Furthermore, Aditya [3], Mulazzani [4] and Aggarwal [5] also used memory forensics techniques as well, and their experiments show these evidences are left is private mode of browsers.

Amari reported on forensics techniques from the viewpoint of memory forensics [6]. On the other hand, in terms of anti-forensics or privacy protection, there is a study [7]. In the position of forensics that targets featured phones, Willassen [8] discussed about the acquisition of evidence left by featured phone. He has acquired the evidence from the flash memory in the feature phone. In addition, [9] and [10] is discussing on this theme.

1.4 Challenging Issues

We carried out the experiment to acquire the artifact left by WebStorage that is a part of HTML5 standard. This acquisition is performed via file system and not from Web user interface and/or APIs of Web framework. Reading the data handled by HTML via Web user interface and/or Web framework API means reading the data via Web browser framework. In this case, it is difficult to ensure the admissibility of evidence. In order to ensure the admissibility, it is necessary to retrieve the evidence via side channel. In this research, we try to retrieve it via file system.

Furthermore, size of retrieved artifact may be become quite large. Therefore, find fragments of evidence from the artifact, and correlate these fragments to construct new evidence may be humanly impossible. This task is hard to perform if there is no automated assistance by computer. However to automate this, it is necessary to elucidate the structure of the artifact. Therefore, it is necessary to analyze the encoding format and data structure of artifacts. Then, based on its results, we have to design/implement the tool for evidence structuring/ visualization.

1.5 Contributions and Result

This research is about the forensics experiment related to HTML5 that is still under standardization process. HTML5 runtime is expected to be the foundation of mobile devices, especially smartphones. These devices From the native nature of the devices, these devices have aggregate information related to the behavior of user. Therefore, retrieve the evidence of user behavior from the foundation layer of these devices is very effective in forensics.

In this paper, we experimented the acquisition of the evidence of artifact left by HTML5 sessionStorage from file system. This is intended to be acquire the artifact of Web browser from the lower layer of the system. Especially, it is important that acquisition is performed with not mediated by the Web browser framework.

Furthermore, we investigated the format of this artifact. In the result, it is found that the artifact is encoded as JSON. Furthermore, we investigated the structure of this artifact and found it records the user's Web browsing history. URL of browsed page and its referrer page is recorded according to the browsing order. In Addition, it was found that artifact recording the additional information of Web pages. At the same time with the investigation, it is found that the size of the artifact can be enormous. Therefore, analyze the artifact and retrieve some evidence from it by human will be difficult. Since forensics work need some manually task, it is necessary to work cooperative with automation tool. Therefore, we designed and implemented the tool that visualize and correlate the evidences from the artifact.

By using this tool, it makes that a forensics investigation of artifacts left by HTML5 Web browser more realistic.

1.6 Comparison with Existing Work

Donny [2] and Aditya [3], Mulazzani [4] have investigated the artifact left by Web browsers. They examined especially on private browsing mode of Web browsers and verified that privacy can be acquired from artifacts. Forensics analysis on artifact of Web browser is also discussed by Satvat [11], Murio [12] and Junghoon [13].

However, these studies do not address the evidence left by APIs added in the HTML5 standard. In contrast to these, in this research, we focus on retrieving the evidence left by HTML5 related APIs. From the viewpoint of memory forensics, basics of its technique is described in Kristine [6]. In addition, memory forensics on Windows machine is discussed bu Runn [14]. Compared to them, our study aims at extraction of evidence from the file system.

2 Overview of Digital Forensics

Subjects of digital forensics is spreading rapidly in response to the transforming in IT. Nowadays, devices handle digital data, such as networks, cloud system, information appliances, and mobile devices are included in this target.

2.1 Applications of Digital Forensics

Digital forensics investigation is included not just those related to the criminal case, also related to civil litigation. In addition, forensics investigation involving patents and disputes between companies, diplomatic international dispute has also gathering attention in recent years. Uses of digital forensics is widely spreading as follows [1].

Criminal Investigations The term "digital forensics" is told in this context primarily. Its main objective is to find and retrieve the electronic evidences left by the criminal act and conserve these evidences validity.

Civil Litigation This application of digital forensics is called "eDiscovery", and its market is spreading rapidly . eDiscovery is defined as "refers to any process in which electronic data in sought, located, secured, and searched with the intent of using it as evidence in a civil or criminal legal case".

Intelligence This application field is called "DOMEX (Document and Media Exploitation)" especially. Modern governments and terrorists are utilizing digital devices. DOMEX is "the collection and exploitation of captured equipment, documents, and media to generate actionable intelligence [15]".

2.2 Mobile Forensics

Mobile forensics, or mobile device forensics is the techniques to investigate on portable device to retrieve the evidence left on it. In generally, investigation is targeting on the cellular phone and smartphones, etc. Because the user is always carrying the device, it aggregates various private information of users. Examples of such private information are, phone book, e-mails, photos and music files. In addition, mobile device is consist of many sensors, such as microphone, camera, acceleration sensor, barometer and GPS antenna. These can also capture private information.

There are many literatures on mobile device forensics. Most of those are targeting feature phones still often [8] [9] [10]. In addition, Report by SANS [16] is targeting not only cellular phone but also MP3 players. However, literature on forensics that targets smartphones has been increasing [17] [18].

When considering digital forensics that targets smartphones, Web browser framework is not negligible. This frameworks is located at the core of the smartphone platform, and many services are implemented thereon.

Especially, Web framework provided in the smartphone platform has been utilized as the execution engine of the HTML5 standard currently under development. Analyzing the artifact left by Web browser supports HTML5 will occupy important place in the mobile forensics.

2.3 About Web Browser Anti-forensics

For the progress of the research on digital forensics, research on countermeasure against it is also progressing. The countermeasure techniques, called anti-forensics, has many definitions [1]. Harris [19] defines it as "consider anti-forensics to be any attempts to compromise the availability or usefulness of evidence to the forensics process".

Of course, anti-forensics that targets cellular and/or smartphone have been studied. Azadegan [20] designed and developed the tool to disrupt the connection between smartphone and forensics device. From the standpoint of anti-forensics on Web browser, it is important to make hard or impossible to analyze or retrieve the artifact of Web browser. Anti-forensics techniques for making evidence

unavailable is classified into four categories by Harris [19]. These categories are "Destroying","Hiding", "Eliminating source" and "Counterfeiting".

3 HTML5 as a Application Development Language

3.1 Abstract of HTML5

Currently, standardization process of HTML5 [21] as the latest version of HTML, is in progress by the W3C and the W3C. It is expected to be modified in many ways to the previous edition [22]. HTML5 has been developed to improve the appearance and usability for Web browser user, and to improve the expressiveness for Web page developers . In addition, HTML5 has been developed to improve the perfection as an application description language. For this reason, many API definitions has been added in HTML5.

3.2 WebStorage

The cookie [23] is a technique that is defined for the purpose of having to maintain information about any status set by the server primarily. In addition, Local Shared Object, is called "flash cookie" is the another method to record some information in client side [24] [25]. In some ways, flash cookie is more useful than cookie. e.g. it never expire. However, this property may cause problems in terms of privacy.

WebStorage [26] is one of the API newly defined in HTML5 standard. It is another method to store data in the client side. There are characteristics of WebStorage below.

Storage Capacity
: Storage capacity of Cookie is 4KB, but capacity of WebStorage is up to at least 5MB. Enlarged storage capacity can increase the amount of data that can handle in client, and thus increasing the flexibility for application developer.

Expiration Time
: Cookie has a limited lifetime and when the cookie expires, it is deleted. In contrast with it, WebStorage can retain the data until deleted explicitly.

Data Transmission
: Cookie is sent over the network when the client interact with server every time. On the other hand, WebStorage must not sent over network. It lessen the burden of network bandwidth and is preferred from the viewpoint of security.

Store Format
: WebStorage is maintained by key-value pair. This is similar to NoSQL style database.

WebStorage is classified as localStorage and sessionStorage, these are defined for different purpose.

3.3 localStorage and sessionStorage

localStorage is one of the kind of WebStorage and the mechanism to store some data in Web browser side. As described in the previous section, localStorage is isolated based on the concept of Web origin [27].

localStorage can be shared between another tabs and/or windows if even have the same Web origin. Furthermore, contents of localStorage is kept after browser has been closed and retained until deleted explicitly.

On the other hand, sessionStorage is the another kind of WebStorage. However, unlike local storage, session storage is not shared between different windows and tabs even if have the same origin. Furthermore, sessionStorage is kept until session. Therefore, when the session is finished, the session storage space will be removed and inaccessible.

3.4 HTML5 and Mobile Devices

As discussed in 3.1, HTML5 has been developed to improve the perfection as an application description language. Especially, it is expected that it will be used as the foundation of mobile platform. Some mobile platforms equip only Web browser framework as the foundation of application execution platform. These platforms have in view to take advantage of the portability and development efficiency of HTML. The potential of HTML5 as a mobile platform is discussed by Juntunen [28].

4 Investigation Result and Proposal Method

Our goal is the realization of a mobile forensics that targets smartphone, but in this paper, we carried out experiments and tool development with Web browser for personal computer. We examined with Firefox 26.0 for Windows and browse dozens of pages. After that, extract the artifact file of sessionStorage before it has been deleted.

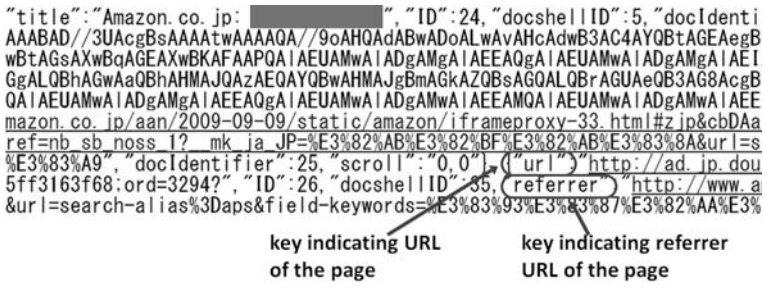

Fig. 1. Artifact Example

4.1 Artifact Format of Web Browser

Example of artifact by Firefox is shown in Figure 1. As is apparent in this figure, artifacts left by Firefox browser is text data and it is encoded in JSON format. In this JSON format artifact, "key indicating URL of the page" indicates the URL of the Web page browsed. In Addition, "key indicating referrer URL of the page" indicates the referrer of Web page. The referrer, to present the Web page visit before making a transition to the Web page.

As a result we analyzed that the structure of artifact as in tree structure. Tree structure of evidence is constructed as Figure 2. In this format, [windows] and [_closedWindows] nodes describes the Web browser windows, respectively. These nodes have [tabs], [_closedTabs] nodes as subsidiary and These nodes may have multiple [entries] nodes. [entries] node describes web page browsed respectively and these have subsidiary nodes [url], [title], [ID], [referrer], etc. [url] node describes the URL of the Web page. [title] node describes the title of the Web page. [referrer] node describes the URL to link the original Web page. In addition, [children] contain the information about the pop-up page kicked by parent window. If there are multiple [entries] nodes as subsidiary of a [tabs] node, it means that this session includes multiple tabs.

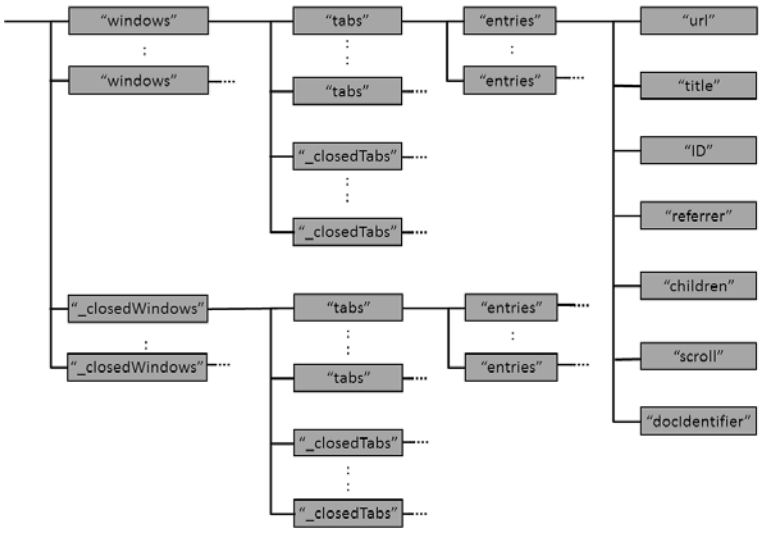

Fig. 2. Structure of Artifact

It is specified that sessionStorage is deleted when the browsing has been finished. However, as the result of our examination, we confirmed that the artifact of sessionStorage is deleted when the browser is launched, not browser is finished.

It is thought to be in order to allow the recovery of the last tab, as a function of the browser.

4.2 Location of Artifact

Location of the artifact of Web browser is depend on the Web browser implementation. As a result of survey, Table 2 summarize the location of WebStorage of major Web browsers.

Table 2. Stored Location of WebStorage

Browser	Version	Stored Path
Internet Explorer	8 or later	N/A
Mozilla Firefox	3.6 or later	<usershome>\AppData\Roaming\Mozilla\Firefox\Profiles\<profileFolder>
Google Chrome	8 or later	<usershome>\AppData\Local\Google\Chrome\UserData\Default
Opera	11 or later	<usershome>\AppData\Roaming\Opera␣Software\Opera␣Stable
Safari	5 or later	<usershome>\AppData\Local\Apple␣Computer\Safari

4.3 Our Proposal Method

In this section, we describe the design and implementation of tool that retrieve the evidences from the artifact left by Web browser (Firefox browser) and visualize it.

We assume to investigate the evidence from the artifact of sessionStorage left by criminal act. Artifact of sessionStorage is accumulating the history that user browsed. But it has a large amount of information (page display size and window size, etc.). These are not related to the investigation. Therefore, we must extract the information that need to be investigated. It is expected to take huge time as human task, and we propose the tool that process the structuration and visualization of evidence semi-automatically.

In this method, we examined on the artifact of Firefox's sessionStorage. About Firefox, if more than one window is open, until all of those windows are closed is considered as period of the same session. Therefore, the browsing history within its period is accumulated in the sessionStorage. By using this property, it is possible to retrieve the evidence not only on the window that is closed at the end, also on the window or tab that has been closed during the same session may be obtained from artifacts file of sessionStorage. Furthermore, Firefox keep the artifact of the last session as backup. From these properties, it is possible to obtain the evidence about session

From these properties, we can retrieve the evidence about sessions on the closed browser and evidence about previous session.

4.4 Design and Implementation

Processing procedure of this tool is described as bellow.

1. Copy the artifact file of sessionStorage to another path and launch the tool to load its file.
2. Parse artifact file to JSON objects and extract "entries" that denotes Web page browsed.
3. Classify each "entries" node to root entry(has no "referrer" node immediately below) and not root entry(has "referrer" node immediately below).
4. Inspect [url] node that immediately below of node root entry, and search the "referrer" that has [url] node same to former node.
5. If search succeed, trace to search for the "entries" node with URL as a key.
6. If search failed, finish to trace, and print the traced "entries" and value of its subsidially node [url], [title], [ID], [docshellID], [referrer].
7. In addition, represent the distance from root as the number of "*".
8. Move to the next root node, and start the search.
9. When process to all root has been finished, exit.

In this process, we treat just Web page browsed directly. It means we do not treat popped-up page opened by other page.

4.5 Evaluation

We examined on the tool implemented. We used to browse with the Web browser Firefox 26.0. When move between pages, we record the URL of the page move source and the URL of the page move destination. After reading dozens of pages, retrieve the artifact of sessionStoarge.

Result of examination indicates, we found that mismatch between the value of [referrer] and [URL]. The reason for this is that wen keyword searching, when loading the page, referrer of the page viewed has been changed. It is observed only when using a specific search site.

5 Conclusion

In this research, we examined the artifact of Web browsers. For four Web browsers (Google Chrome, Mozilla Firefox, Apple Safari and Opera browser), we located the path of artifact that is left by sessionStorage function.

Furthermore, we examined the format of the artifact, and we revealed that it is encoded as JSON format. Based on these results, we have designed and implemented a tool that structures the evidence of artifact and visualize its result. Which makes it possible to extract the evidence necessary to investigate digital forensics from data fragment left by sessionStorage and present the findings to the investigator.

From the viewpoint of anti-forensics, it must be prevented to locate the artifact file and to be analyzed. Especially later is more important. To prevent

the contents of the artifact file to be analyzed and achieve the anti-forensics, it is necessary to revise the implementation of Web browser. According to the classification by Harris [19], adopt "Destroying" or "Eliminating source" for anti-forensics methods of Web browser running is difficult. Adoption of these methods would be to impair the normal function of the Web browser. Therefore other methods, "hiding" and/or "Counterfeiting" would be effective as anti-forensics techniques. In particular, it is expected "Hiding" of artifacts fileusing any encryption teqhnology to be effective.

5.1 Future Work

In this research, we examined on Firefox browser and another Web browsers, MS Internet Explorer, Google Chrome, etc. is not examined sufficiently. We must examine other Web browsers and carry out implementation of the tool using the same verification and evaluation. In addition, the artifact by localStorage must be examined and correlation between the evidence left by WebStorage from file system and other evidence.

As a part of this effort, we are studying about the memory forensics in WebStorage if Web browsers on Windows [29]. In addition, to apply to mobile forensics on the results of this study, it is necessary to survey the browser of the mobile OS and on iPhone and Android.

References

1. Sammons, J.: The Basics of Digital Forensics. Elsevier (2012)
2. Ohana, D.J., Shashidhar, N.: Do private and portable web browsers leave incriminating evidence? a forensic analysis of residual artifacts from private and portable web browsing sessions. Security and Privacy Workshop (SPW) (May 2013) 135–142
3. Aditya Mahendrakar, James Irving, S.P. In: Forensic analysis of private browsing artifacts. IEEE (2011) 197–202
4. Mulazzani, M.: New challenges in digital forensics: online storage and anonymous communication. PhD thesis, Vienna University of Technology (2014)
5. Aggarwal, G., Bursztein, E., Jackson, C., Boneh, D.: An analysis of private browsing modes in modern browsers. In: USENIX Security Symposium. (2010) 79–94
6. Amari, K.: Techniques and tools for recovering and analyzing data from volatile memory. `http://www.sans.org/reading-room/whitepapers/forensics/techniques-tools-recovering-analyzing-data-volatile-memory-33049` (2009)
7. Waksman, A., Sethumadhavan, S.: Silencing hardware backdoors. SP '11 Proceedings of the 2011 IEEE Symposium on Security and Privacy (2011) 49–63
8. Willassen, S.: Forensic analysis of mobile phone internal memory. In: Advances in Digital Forensics. Springer (2005) 191–204
9. Jansen, W., Ayers, R.: Guidelines on cell phone forensics. NIST Special Publication **800** (2007) 101
10. Ahmed, R., Dharaskar, R.V.: Mobile forensics: an overview, tools, future trends and challenges from law enforcement perspective. In: 6th International Conference on E-Governance, ICEG, Emerging Technologies in E-Government, M-Government. (2008) 312–23

11. Satvat, K., Forshaw, M., Hao, F., Toreini, E.: On the privacy of private browsing - a forensic approach. In: Data Privacy Management and Autonomous Spontaneous Security, Springer Berlin Heidelberg (2014) 380–389
12. Tito, M.: Forensic analysis of the firefox 3 internet history and recovery of deleted sqlite records. Digital Investigation: The International Journal of Digital Forensics & Incident Response archive **5** (March 2009) 93–103
13. Oh, J., Lee, S., Lee, S.: Advanced evidence collection and analysis of web browser activity. Digital Investigation: The International Journal of Digital Forensics & Incident Response archive **8** (August 2011) S62–S70
14. Ruff, N.: Windows memory forensics. Journal in Computer Virology **4**(2) (2008) 83–100
15. U.S. Army.: 2009 army posture statement. `http://www.army.mil/aps/09/information_papers/document_media_exploitation.html` (2009)
16. Martin, A.: Mobile device forensics. `http://www.sans.org/reading-room/whitepapers/forensics/mobile-device-forensics-32888` (2008)
17. Hoog, A., Strzempka, K.: iPhone and iOS Forensics: Investigation, Analysis and Mobile Security for Apple iPhone, iPad and iOS Devices. Syngress (2011)
18. Hoog, A.: Android Forensics: Investigation, Analysis and Mobile Security for Google Android. Syngress (2007)
19. Harris, R.: Arriving at an anti-forensics consensus: Examining how to define and control the anti-forensics problem. digital investigation **3** (2006) 44–49
20. Azadegan, S., Yu, W., Liu, H., Sistani, M., Acharya, S.: Novel anti-forensics approaches for smart phones. In: System Science (HICSS), 2012 45th Hawaii International Conference on, IEEE (2012) 5424–5431
21. Berjon, R., Faulkner, S., Leithead, T., Navara, E.D., O'Connor, E., Pfeiffer, S., Hickson, I.: HTML5 a vocabulary and associated APIs for HTML and XHTML W3C candidate recommendation 6 august 2013. `http://www.w3.org/TR/2013/CR-html5-20130806/` (August 2013)
22. Pieters, S.: Differences from HTML4. `http://www.w3.org/TR/2013/WD-html5-diff-20130528/` (2013)
23. Barth, A.: HTTP State Management Mechanism (2011) RFC6265.
24. Soltani, A., Canty, S., Mayo, Q., Thomas, L., Hoofnagle, C.J.: Flash Cookies and Privacy. `http://ssrn.com/abstract=1446862` (2009)
25. Ayenson, M.D., Wambach, D.J., Soltani, A., Good, N., Hoofnagle, C.J.: Flash Cookies and Privacy II: Now with HTML5 and Etag Respawning. `http://ssrn.com/abstract=1898390` (2011)
26. Hickson, I.: Web Storage. `http://www.w3.org/TR/2013/REC-webstorage-20130730/` (July 2013)
27. Barth, A.: The web origin concept. `http://tools.ietf.org/html/rfc4627` (2011)
28. Juntunen, A., Jalonen, E., Luukkainen, S.: Html 5 in mobile devices–drivers and restraints. In: System Sciences (HICSS), 2013 46th Hawaii International Conference on, IEEE (2013) 1053–1062
29. Matsumoto, S., Sakurai, K.: Acquisition of evidence of webstorage in html5 web browsers from memory image. AsiaJCIS 2014 (September 2014)