EXPOSING THE TOR FAILURES ON MOBILE DEVICES USING PARABEN’S E3:DS TECHNOLOGY

Ton Don, Georgia Southern University, td02284@georgiasouthern.edu
Hayden Wimmer, Georgia Southern University, hwimmer@georgiasouthern.edu
Lei Chen, Georgia Southern University, lchen@georgiasouthern.edu
Queen E. Booker, Minnesota State University, queen.booker@mnsu.edu

ABSTRACT

Privacy and anonymity tools on electronic devices are increasingly commonplace. These tools provide challenges for digital investigators especially when trying to track down and uncover illegal activity such as the use of the “Darknet”, an illegal online service trading market which is hidden from the normal internet users and is often difficult to detect with digital forensic tools, especially if the user uses privacy and anonymity “protection” on their devices. However, digital forensic tools continue to evolve with the ability to detect internet usage, even those supposedly hidden by such privacy tools. These tools present privacy concerns to users. This work seeks to examine the extent that the Tor Browser Bundle protects a user’s privacy. Using a mobile digital forensic software by Paraben, E3:DS, we extract the mobile device’s data, a Samsung Galaxy Note 5 using the Android 6.0.1 operating system. While past research has claimed that using Tor protects user privacy, analysis of the data revealed we were able to extract the websites visited via Tor and extracted search terms from common shopping and social media websites. Implications from this research are three-fold: first, the study shows that E3:DS can reveal internet activity even while running applications that are supposed to provide privacy and anonymity; second, user’s activities are not fully protected when using the Tor Browsing Bundle to surf the internet; and, finally, there are weaknesses in the Tor Bundle that Tor developers need to address in their software to support their privacy claims.

Keywords: Mobile Digital Forensics, E3:DS, Tor Browser Bundle, Internet Privacy.

INTRODUCTION

With electronic technology growing rapidly every day, it is difficult for digital forensic tools and methods to keep pace (Lillis, Becker, O'Sullivan, & Scanlon, 2016). Digital forensics (DF) is an interdisciplinary research area and practice between forensic science and computer science. DF uses methods which are scientifically derived and proven to acquire evidence through a well-defined process. Using this evidence, investigators can reconstruct a crime scene in order to find the culprit or help with future unauthorized activities or operations (Carrier, 2003; Reith, Carr, & Gunsch, 2002). In this study, we investigate a digital forensics tool, E3:DS created by Paraben. Our main interest is to examine artifacts from Orfox, a Tor-based web browser, on a Galaxy Note 5 Android smartphone. Using E3:DS by Paraben, we were able to access the internet history of the user of the Galaxy Note 5 device. After carefully examining the information found on the device using E3:DS, we present our findings and show the security and privacy weaknesses which currently exist in the mobile Tor Browser Bundle. With the findings, law enforcement can improve their current digital forensics process on mobile devices. As for Tor users and developers, this will alert them as the users will be more careful while using the software and the developers will try to figure out how to overcome the vulnerabilities.

LITERATURE REVIEW

Digital Forensics Research

With technology growing at a rapid rate, digital forensic tools and methods are struggling to keep up with the pace. This leads to the issue and question about how investigators can obtain evidence from mobile devices, especially since different vendors have different log configurations and encryption for their devices. In fields that research computers and cellphones, however, the rate of change is faster than the normal times required for peer-reviewed
publication. As a result, establishing effective strategies for discovering the facts associated with a set of digital data can be difficult. The rate of change in technology exacerbates this issue, as the lifespan of some analysis techniques or interpretative guidance can be short lived due to the rate of release of updated versions of software which may change the internal structure or fundamental workings of an application and its associated data. (Sommer, 2010)

Thus, frameworks which are artefact specific can be of limited use. Yet frameworks providing a practitioner with the support to implement robust testing on any form of digital data are arguably a valuable tool and remain in short supply. (Horsman, 2018)

Watson and Dehghantanha (2016) described the challenges created by the massive network of mobile devices. Three main issues they discussed were:

1. the onboard data storage is not accessible using traditional DF methods,
2. cumulative data sets might be in multiple locations, and
3. the existing tools might not be able to read retrieved data.

Because of these existing challenges, it is difficult for investigators to identify illegal activity evidence and present their findings to justify the search and capture of the criminals.

Dhar and Pingle (2016) discussed the framework for digital forensic investigation (DFI). DFI has four steps, (1) investigation preparation, (2) evidence acquisition, (3) analysis of evidence, and (4) result dissemination. Dhar and Pingle (2016) also mentioned a three-step diagram for DFI for Internet of Things (IoT). Since mobile devices are a part of IoT, this article gives some insight into the things that needed to be examined. Flowchart 1, adapted from Dhar and Pingle (2016), illustrates the framework:


Horsman (2018) also discussed a framework for digital forensics investigation. The proposed Horsman framework is called Framework for Reliable Experiential Design (FRED). The framework has six core stages, namely “plan”, “implement”, “evaluate”, “repeat”, “analyze” and “confirm”. The plan stage involves establishing the goals and hypotheses. The implement stage requires the user to carry out a series of actions to simulate user behavior in line with the planned methodology. These actions constitute the “data set” which is used during testing. Following a successful test implementation, the outcome must be evaluated. In order to do this, the effect of the implemented tests on a system or series of artefacts must be identified and collected. This process is the evaluate stage. Repeatability is the key to establishing robust knowledge as the standard to be achieved through the utilization of FRED is that of fact, where repeatability is a useful measurement of reliability. Reliability cannot be established from one test alone, and result must be capable of being replicated. The testing process that must be capable of repetition. The analysis stage involves the interpretation of results generated from testing carried out and collected during the “evaluate” and “repeat” stages of FRED. The final stage of FRED is the ability to affirm as a matter of fact, the outcome and interpretation of testing which has taken place and to document the process. A full explanation of the framework can be found in Horsman (2018).

The Darknet

According to Dayalamurthy (2013), the Darknet is the place where criminals and unethical users have access to illegal services and hidden links. These links can provide access to things such as child pornography, drug dealing, crime-as-a-service, and hidden hosting devices. Because of Tor’s seemingly ability to protect user identity and privacy, the Darknet was quickly affiliated with Tor. Jardine (2015) reviewed several articles written about Tor, of which the majority focused on how Tor assisted the Darknet in becoming a place for ill intentions such as drugs trading, crime as a service, child pornography, etc.

The term Darknet gained traction after Microsoft researchers Peter Biddle, Paul England, Marcus Peinado and Bryan Willman gave a paper at the 2002 ACM Workshop on Digital Rights Management. Although there is no clear evidence that the Darknet is necessarily a “bad place,” the Darknet infrastructure specifically was designed for
anonymity. Mansfield-Devine (2009) stated that the Darknet could be a tool for many legal but non-publicly available uses, if used properly. Nevertheless, it is best known as a place on the internet where people engage in illegal activity without expectation of being monitored or “caught” by law enforcement. As such, cyber-criminals have formed groups based on their languages and then later link back to a larger group of criminals.

**Tor as an Application**

The Onion Router (Tor) is a network of volunteer-operated servers which purports to give people the ability to improve their anonymity and privacy over the Internet. There are a series of tunnels, to which Tor’s users can connect while they are on the Internet. This purportedly allows them to access blocked or Darknet sites or content without being tracked. Tor also promotes a hidden service, which allows websites to be published anonymously. The more users use Tor, the more secure the network is (TorProject.org, 2017).

Tor networks create a random pathway through several relays on the network. Furthermore, a twisty and hard-to-follow route is used so that users cannot be tailed. Users’ footprints are also periodically erased. Each relay on the network only knows where to receive and to send data. Because of this, no relay knows the complete travel path of the information. All the traffic between the relays is encrypted except the last relay to the destination, refer to Figure 1. However, each time the client hops from one relay to another, the client requests a new encryption key. This further improves the security and protects the users. Another feature of the Tor network is that it requests a new path through the relay every ten minutes or so, supposedly making a user’s path impossible to follow. These features are supposed to secure Tor’s users’ anonymity and protect their privacy (TorProject.org, 2017).

![Figure 1. How Tor Works (TorProject.org, 2017)](image)

Because of Tor’s seemingly ability to protect user identity and privacy, the Darknet was quickly affiliated with Tor. Jardine (2015) reviewed several articles written about Tor, of which the majority focused on how Tor assisted the Darknet in becoming a place for ill intentions such as drug trading, crime as a service, child pornography, etc. The author also argued that government regulations were the reason people use Tor or the Darknet. The more governments sought policies and laws aimed at reducing potential illegal activity on the internet or using the internet for political unrest organizing, Jardine argued, the higher the usage rate for Tor. This occurred because the perceived anonymity of using Tor or the Darknet allowed users to feel they could freely express their opinions. Jardine (2016) continued that political regressions of one’s country were the reason why Tor’s usage was rising. To express their political opinions without being monitored by their own government, the Tor network became the people’s tool of choice. Jardine showed how political regression affected the number of users on the Tor network. The relationship between political regression rate and Tor usage formed a U shape. If a country allowed its citizens to exercise their liberal and political rights, they were less likely to have people using the Tor network. According to Jardine (2016), in 2013, there were 620 million internet users recorded. With a political regression scale of 14, this would cause an increase of 1,317,996 Tor bridges and 37,785,528 Tor relay users per year reinforcing that Tor could be useful for countries which were politically suppressed. Jardin (2015) believed the best option to promote democratic ideas while reducing illegal activity was to police the network so that the Darknet could bring more positive effects.
In 2013, Sandvik (2013) reported about 100,000 downloads for the Tor Browser Bundle every month. Sandvik performed a forensic analysis of the Tor Browser Bundle on three operating systems: OS X 10.8, Windows 7, and Debian 6.0 Squeeze Linux. The processes of setting up all operating systems and Tor Bundle were the same. Many traces of the Tor Browser Bundle were found on each operating system. The researcher also suggested using The Amnesic Incognito Live System (TAILS) so that no traces would be left on the system. Tor should erase user’s footprint from the system. Unfortunately, because of the default settings of the operating systems, the bundle could not remove the left-over tracks.

To learn more about the Tor network, Al Barghouthy and Marrington (2014) performed a study over a rooted Samsung S2, Android v.4.1.1, using Orweb, a Tor browser built for Android devices. The main purpose of this study was to find digital evidence left over after using Orweb. The image of the device was created using rootkit method and recovery mode method. After the images were examined, there were traces of information of the places users visited such as URLs, email, IDs, encrypted messages, the RSA key, ports used, rejected ports, and date/time stamp. The researchers concluded that it was not necessary to root the device because acquiring the image using the Rootkit method was better since the device did not have to be rebooted, which was better for the artifacts on Random Access Memory (RAM).

Prior Research on Tor

Al-Khaleel, Bani-Salameh, and Al-Saleh (2014) examined the Tor claims by collecting artifacts from the Tor Browser Bundle by extracting them from memory. To do so, they had to understand how Tor worked. After that, they designed an investigation model and experimental setup. There were six experiments. After the data from all the experiments were collected and carefully examined, the researchers concluded all the valuable information, which could be extracted from the Tor Browser Bundle, was only available when Tor and its’ browser were open. The moment the browser or Tor was closed, all the artifacts were erased from memory thus confirming Tor developers claims of anonymity.

Chrane and Kumar (2015) also conducted a study on Tor to see if Tor could become the source for anonymity on the internet as a mainstream browser. Unfortunately, it was not simple to make Tor work. For a novice Tor user, if the software was not properly configured, there would be no anonymity. It is not user-friendly. On top of that, with the rapid growth rate, there were no resources available, such as nodes, for Tor to be consistent. With a limited number of nodes, the Tor network would slow down or fail. There was a fix for this issue which was Universal Rate Limit. However, the Tor software was not compatible. Even though there were limitations for Tor, the authors still believed that Tor would be the future of the Internet.

Saleh, Qadir and Ilyas (2018) surveyed various studies conducted on the Tor network. They quantified the studies into three groups: (1) deanonymization, (2) path selection, (3) analysis and performance improvements. Their study shows that majority of the research works were in the deanonymization group, followed by performance analysis and architectural improvements. In the deanonymization group, most of the research was devoted to breaching attacks followed by traffic analysis. In the path selection group, most of the research focused on the development of new algorithms. Their analysis over simulations and experiments shows that 60% of studies used experiments and 86% of those experiments were carried out on private testbed networks. Among simulations, 75% of the studies developed their own simulator to analyze Tor network. Analysis of simulation parameters shows that there is no distinct pattern of parameters. However, majority of the studies used bandwidth and latency. In deanonymization group, research was observed in six different categories: (1) Hidden services which limit their scope to hidden servers identification, (2) Fingerprinting which are based on pinpointing Tor network, (3) Attacks which are focused over breaching Tor network, (4) Traffic analysis which analyze Tor traffic to pinpoint the weaknesses, (5) Studies studying improvements in Tor to avoid deanonymization, and (6) Anonymity without Tor which suggest alternate methods to provide anonymity by pinpointing weaknesses in Tor.

E3:DS

E3:DS by Paraben is software for mobile forensics supports more than 26,000 devices. E3:DS allows users to access: image devices logically and physically, parse app data, carve data, bypass passwords, and unique content analysis. E3:DS supports different operating systems: iOS (up to version 10.3.x), Androids (up to version 7.x), Blackberry (up to version10.x), Windows (up to version 10), etc. (Corporation, 2017).
RESEARCH METHODOLOGY

This study follows the framework suggested by Horsman (2018) which outlines an experiential design for digital forensics to ensure the dependable interpretation of digital data.

Plan Stage

For this study, our hypothesis was that the software in the Tor bundle could effectively provide anonymity for internet users. Our study environment was a Samsung Galaxy Note 5, model number SM-N920G with 32 GB of storage running Android Marshmallow version 6.0.1. Version 6.0.1 was the newest version what was available on an Android device at the time the study was conducted, with an exception for Google’s devices which had Android version 7 installed.

Orbot is open-source software which includes Tor, Polipo, and LibEvent. With Orbot, users can have a local HTTP proxy (port 8118) and a SOCKS proxy (9050) which allow them to access the Tor network. On a rooted Android device, Orbot can also redirect the traffic through the Tor network. The version which we use for this study is 15.2.0-RC-8-multi. We installed Orbot through the Android Play Store. Orfox is also an open-source software and part of the Tor Project. It is built using the same source code as the Tor browser. Orfox is modified with more features which enhance privacy and make the software compatible with the Android operating system. We installed Orfox Fennec-45.5.1esr/TorBrowser-6.5-1/Orfox-1.2.1, which was uploaded on December 1, 2016.

We used (Lee, 2016) to download all the necessary software and perform the rooting process. The software applications used were ODIN v3.12.2, Magisk v12.0, Samsung antir root removal v2.4, phh’s SuperUser v1.0.3.3, and TWRP for Galaxy Note 5 International SM-N290G. The steps for the rooting process can be found on the site.

For our Samsung Galaxy Note 5, we navigated to the device’s settings and chose Developer options. While in Developer options, these options were turned on: developer, stay awake, OEM unlock, USB debugging, and verify apps via USB.

Implement Stage

To generate data, we browsed through two shopping sites: eBay and Amazon using Orfox when Orbot was active. To ensure that the device connected to the Tor network and the software functions properly, we visited check.torproject.org for a connection test.

First, we started E3:DS via our USB dongle. Once the application started, there were two options which were suggested by the software. One of the options was “Acquire Device”, this was what we would choose if we already had a case. The other option was “Add Evidence”, we chose this option. If “Add Evidence” option was chosen, a new case would be created automatically. Then, we navigated to “Mobile Data” and chose “Mobile Data Acquisition”. “Acquisition Wizard” appeared and asked for device type: “Portable Device”, “Android”, and “Samsung GSM”. Since we were using a Samsung Galaxy Note 5 with an Android operating system, we chose the Android option. Then, the type of acquisition was asked. We chose “Full Logical Acquisition”. As for the next screen, all the options were left as default and we chose “Start Acquisition” and waited until the process was finished.

Below, Figure 2, was the screen shot of all the folders which were collected by E3:DS after the acquisition process. Under the device, we can see there is a file system folder, contacts, authentication data, call history, installed applications, SMS, MMS, media store, calendar, default browsing history, and settings. As for this study, we paid more attention to the file system folder.
Evaluate Stage
After the acquisition process was finished, we explored our interested folders and files. After a few hours of examining all the collected files and directories, we found our target. Located inside “SM-N920G/FileSystem/data/User/0/info.guardianproject.orfox/files/mozilla/51xiouku.default/”, there was a file named “browser.db-wal”. Using E3:DS, we could display the content of this file in text format. There were traces of what the user did during the Orfox browsing session.

The first site we discussed in this section was Amazon. Figure 5.1a presented a link found in our text file. By analyzing the leftover evidence from the Orfox browser, we could see that the user was trying to sign in to the site. On top of that, there was information about the character encoding which is UTF-8. Figure 5.1a also showed us that Amazon was pulling the user’s order history and account status policy. Not only that, the user selected a few products which were available in the store. The selected products were Samsung Galaxy S8, S8 Plus, and a case for Samsung Galaxy S8 Plus, refer to Figures 3, 4, 5, and 6.

Figure 3. Information on Amazon log in page

Figure 4. User’s search history on Amazon
Figure 5: More user’s search history on Amazon

Figure 6: Item user was interested in on Amazon

The next shopping site which we used was eBay. As shown in Figure 7, there were traces left over from the user while at the sign-in page. Not only that, there was information about the kind of protocol which was used to secure users’ information. Furthermore, the user was visiting the Cell Phones and Smartphones section of eBay. The two devices which the user took an interest in were Samsung Galaxy S8 and S8 Plus, figure 8. The final piece of evidence was Spigen, a producer of Samsung Galaxy S8/S8 Plus [Air Skin] Ultra Slim Lightweight Case Cover, figure 9.
We intentionally chose the same products while surfing two shopping sites so that there would be a better comparison between them. After presenting all the collected evidence from the Galaxy Note 5, we saw that the information regarding the browsing sessions was identical. The product listing names were shown, refer to figure 5.1b, 5.1c, 5.1d and 5.2b, 5.2c. The only difference was the information about how the sites were set up, the format or the protocol that one site preferred over another, refer to figure 5.1a and 5.2a.

Repeat Stage
To replicate our findings from the shopping site, we decided to check two social media sites: Facebook and Twitter. As shown in Figure 10, the user accessed Facebook and performed a search query with the key word “georgia southern”. Next, the user visited the active friend list on this Facebook account. Another search was performed, and the search query for this time was “danny don”, refer to figure 11. The last piece of evidence which we found on the text file was presented on figure 12. It was another search query with the keyword “ton don”.

Another example for social media site was Twitter. By looking at figure 13, we saw the user’s Twitter ID, handle, and information about the website, the mobile Twitter site in this case. After signing into Twitter, the user’s search queries were found, figure 14, 15, 16. The last thing we want to present for the Twitter web browsing session was in figure 17. GASouthernNews was the last thing the user visited while browsing Twitter.
Comparing the evidence found on the device between Facebook and Twitter, there were a lot of similarities. All the information about user’s search queries were left behind after the web surfing session, figure 11, 12 compare to figure 14, 15, and 16. The major difference between Twitter and Facebook was the user’s ID and handle for Twitter were recorded, figure 13.

**Analyze and Confirm Stages**
Recall our hypothesis, that the Tor bundle could effectively provide anonymity for internet users. For the hypothesis to be accepted, we should not have been able to find evidence of user experiences while the Tor bundle was active. We evaluated four different websites, of which we were able to find trace evidence of all four on the Samsung hard drive. Thus, we can conclude that the Tor bundle does not effectively provide anonymity for users of the Samsung Galaxy Note 5, model number SM-N920G with 32 GB of storage running Android Marshmallow version 6.0.1. A limitation of this study is that our analysis was only for one device running one operating system. Thus our results are not generalizable but does indicate that the proposed reliability of Tor is not accurate.

**CONCLUSION**
In the past, multiple studies were conducted upon the Tor Browser Bundle. Since there were limited digital forensics tools and methods for the bundle (Dayalamurthy, 2013), most of the studies analyzed the memory dump from the mobile device (Al-Khaleel et al., 2014; Al Barghouthy & Marrington, 2014; Dayalamurthy, 2013). Because there has been a huge improvement in forensics tools and methods, we can better analyze the mobile device, Galaxy Note 5, from a new perspective. The Tor Browser Bundle is meant to leave no traces after the browsing session to protect user’s privacy, but we prove otherwise. We found not only the places where the user visited, but we also discovered performed activities by the user. The information from this study can be useful to digital forensics investigators, the Tor Browser Bundle users, and developers. From an investigator’s perspective, using this study, they can accurately examine mobile devices which have Tor software installed. For users, they can see that the Tor Browser Bundle doesn’t fully cover their tracks. As for the developers, this study will give them more insight into their software weaknesses. For future studies, the same process will be used on different mobile devices and operating systems. More scenarios will be created and further tested to discover more about the software.
REFERENCES


TorProject.org. (2017). Tor: Overview.
