A Study of the Data Remaining on Second-Hand Mobile Devices in the UK

Olga Angelopoulou
University of Warwick, olga.angelopoulou@warwick.ac.uk

Andy Jones
University of Hertfordshire, andy1jones.aj@gmail.com

Graeme Horsman
Cranfield University, g.horsman@tees.ac.uk

Seyedali Pourmoafi
University of Hertfordshire, s.pourmoafi@herts.ac.uk

Follow this and additional works at: https://commons.erau.edu/jdfsl
Part of the Computer Law Commons, and the Information Security Commons

Recommended Citation
DOI: https://doi.org/10.58940/1558-7223.1785
Available at: https://commons.erau.edu/jdfsl/vol17/iss2/5
A Study of the Data Remaining on Second-Hand Mobile Devices in the UK

Cover Page Footnote
The funding for the purchase of the mobile phones was generously provided by Comparitech.com, a company that carries out reviews of technologies.
A STUDY OF THE DATA REMAINING ON SECOND HAND MOBILE DEVICES IN THE UK

Olga Angelopoulou\textsuperscript{a,*}, Andy Jones\textsuperscript{b}, Graeme Horsman\textsuperscript{c}, Seyedali Pourmoafi\textsuperscript{b}

\textsuperscript{a}University of Warwick, Warwick Manufacturing Group, 6 Lord Bhattacharyya Way, Coventry CV4 7AL, UK; olga.angelopoulou@warwick.ac.uk
\textsuperscript{b}University of Hertfordshire, Department of Engineering and Computer Science, College Ln, Hatfield AL10 9AB, UK; \{a.jones26,s.pourmoafi\}@herts.ac.uk
\textsuperscript{c}Cranfield University, College Rd, Cranfield, Wharley End, Bedford MK43 0AL, UK; graeme.horsman@cranfield.ac.uk

ABSTRACT

This study aimed to identify the level and type of information that remained on portable devices purchased from the secondhand market in the UK over the last few years. The sample for this study consisted of 100 secondhand mobile phones and tablets. The study aimed to determine the proportion of devices that still contained data and the type of data they contained. The study attempted to determine the level of personal identifiable information associated with the previous owner where data were identified. The research showed that when sensitive and personal data was present on a mobile device, in most cases, there had been no attempt to remove it. However, fifty-two percent of the mobile devices had been reset to the factory settings or had all the data erased, demonstrating the previous owner’s attempt to remove permanently personal identifiable information. Twenty eight percent of the devices sold were not functional or recognized by the software used in the research. Twenty percent of the devices that contained data contained data that gave away the previous owner’s identity.

Keywords: data erasure, digital forensics, mobile devices, mobile phone, portable devices, residual data, smartphone, survey

1. INTRODUCTION

Around the middle of the 1980s, the first mobile phone was introduced to the public and became popular in the early 1990s. The continuing development in technology quickly advancing to feature phones. The term smartphone emerged in the late 1990s with the development of Personal Digital Assistants [1]. However, smartphones became widely used in the early 2010s. At around the same time, tablet devices were introduced and became popular. The constantly increasing use of smartphones and tablets over the last two decades has provided instant access to any type of service and information a user may require. It is claimed that in many cases, they have substituted the use of computers with a portable device. Nonetheless, these devices transformed communication and access.
to information for a vast number of mobile device users around the globe.

Statista [2] predicts that the number of smartphone users worldwide will reach more than 6.3 billion users by the end of 2021. However, the number of mobile phone users worldwide is expected to exceed the seven billion mark by the end of the same year [3]. A similar study from [4] predicted that the total number of tablet users worldwide would increase to 1.28 billion by the end of 2021. As a result, the number of portable devices being disposed of or made available to the secondhand market yearly is expected to grow constantly.

Every data storage device requires measures to be taken before it is made available to the market or passed on to a new user if privacy is maintained. All mobile devices allow the user to revert the device to the factory settings. This option allows the user to remove any Personal Identifiable Information (PII) from the device. While the actual definition of PII varies in different countries, a general description of Personally identifiable information (PII) is data that could identify a specific individual. Information that can distinguish an individual’s identity from another or deanonymize anonymous data is also considered PII. The issue of remnant data on second hand storage media is an ongoing topic of interest and has been researched for over a decade. Prior research [5-18] has shown throughout the period that from hard disks to USB thumb drives, micro SD memory cards, and mobile devices and in all types of storage medium end-users dispose of; there is a significant likelihood that they will contain personal identifiable information. The secondhand mobile device market is undoubtedly growing worldwide as the availability of increasingly feature rich mobile devices on offer grows. The focus of this study is to investigate the situation on a national level and explore the occurrence of data remaining on secondhand devices in the UK. The growth in popularity of mobile devices suggested a need to investigate the amount of PII that could have been made available to the secondhand market through mobile devices.

This paper is organised as follows. Section 2 presents the existing literature review; Section 3 explains the scope of the study and Section 4 outlines the research procedure used for the study. Section 5 presents the study’s results; Section 6 summarises some of the most notable case studies identified during the analysis and discusses the possible implications, while section 7 contains the conclusions, discusses future directions, and provides recommendations.

2. A LOOK AT EXISTING LITERATURE AND COVERAGE

With an increase in the diversity of mobile device types available for purchase comes a natural ‘turnover’ of device ownership. It is arguable that this began to be witnessed around 2007-08 [15] with the emergence of the first truly game changing mobile devices when the first iPhone and Nokia’s competing device the ‘N95’ were introduced. Fast developing device capability arguably led to a drive to possess such devices, but the cost can be an issue, leading to the emergence of device trade-in companies. Those seeking to part company with their devices may rely on an inbuilt ‘factory reset’ function, though it was suggested in mainstream media reports that as early as 2014, such functionality lacked effectiveness [19], [20], [21]. The first reports of the threat of personal data remnants being left on second hand mobile devices were raised as early as 2008 [15]. Following media
acknowledgment of the risks of data recovery from secondhand devices, a number of formal academic studies have acknowledged concern [28], albeit not always within the mobile device context.

One of the earliest studies confirming the need for appropriate sanitization of digital media before discarding a device is that of Garfinkel and Shelat [24] in 2003. Later, Szewczyk and Sansurooah’s 2012 study involving the purchase and attempted recovery from 78 memory cards purchased through Australian auction sites revealed that only 29% of memory cards had no recoverable data. In comparison “Fifty-one percent (40) of memory cards purchased had all data intact with no evidence to suggest that the seller attempted to remove the data” [25]. In 2013, Szewczyk et al’s [26] follow-up study indicates that only 27% (38) of the sampled memory cards had no recoverable data. In 2016, Jones et al’s [27] study of secondhand hard drives purchased in the UAE indicated similar concerns, with 26 (65%) of disks containing recoverable data. Jones et al’s previous work has shown consistent data recovery from secondhand device purchases [15], [16], [28], [29], [30], [31], with work from Medlin and Cazier confirming such concerns [32]. Finally, concerns not only lie with more mainstream forms of digital storage media, where Podhradsky et al’s [33] 2011 study indicated gaming consoles may also pose a risk.

Concerns over the continued failure to effectively remove information from a device before discarding or selling it remain, although there appear to be fewer recent formal studies evaluating this issue. Arguably there remains a need for such sampled approaches to continue to assess whether the threat of the non-removal of data remains, regardless of whether this is through poor sanitization practices or via a failure by any given device to effectively remove content when a wipe function is triggered.

3. SCOPE OF RESEARCH

The scope of this research was the acquisition, analysis, and reporting on the data remaining on 100 mobile devices (mobile phones, smartphones and tablet computers) purchased on the secondhand market in the UK. A prerequisite for selecting a device to be purchased was for the device to be clearly advertised as ‘used’ or ‘pre-owned’. Within the scope of the research was data that could be recovered from the devices using freely available software tools or through a manual examination process. No attempt was made to jailbreak or root the devices. The use of “chip off” technology and processes is outside the scope of this research. If data was extracted from a device, either using the software tools or through manual examination, it was considered that the result was a successful extraction. No attempt was made to determine the potential proportion of available data that was recovered.

4. RESEARCH PROCEDURE

The study aimed to determine the proportion of devices that still contained data and the type of data they contained. It followed the processes and procedures that were used in similar studies [5-13], [16-17] that examined residual data in hard disks, SD cards, and mobile devices obtained from the secondhand market. Therefore, the research methodology used to investigate mobile devices remained the same as in the earlier research. The approach was to recover data that anyone who purchased the device could gain access to without the use of specialized tools.
The analysis focused on identifying the extent to which data from mobile devices had been securely erased, and the devices had been reset to the factory settings. When data were identified, the analysis focused on determining the amount of PII that could be recovered from the device, and that could be used to identify an individual or an organisation. For this study PII follows the definition as provided in Article 4 of the UK GDPR (General Data Protection Regulation):

"Personal data means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person" [34].

Between April and December 2018, 100 secondhand mobile phones were purchased from second hand sources, including eBay, conventional auctions, and secondhand shops in the UK. All the mobile phones were purchased singly and over a protracted period to prevent any distortion in the market from a large number of purchases in a short period of time, and effort was made to ensure that the sources of the devices were spread geographically around the UK. The sampling of the devices was the same as in previous studies, and it has been considered a representable sample size to provide reliable and valid results over the years [6-13], [26]. The sample was randomly selected from a broad range of devices available on the secondhand market and included smartphones, tablets, and other mobile phones. They were selected to encompass a broad range of devices based on availability. It is accepted that this may not reflect the current balance of market share for the main operating systems. The mobile devices were supplied ‘blind’ to the researchers, with no indication of the source and the only identifier being a sequential serial number.

A formal digital forensics methodology was followed for the duration of the study comprising collection, examination, analysis, and reporting of the findings. Figure 1 outlines the research procedure that was followed for the study.

Te collection initiated with the mobile phone purchase, followed by an attempt to forensically acquire a logical image of each device. If a logical image could not be acquired, then a manual examination of the device was attempted. This would entail turning on the device, attempting to recover any data present, and photographing the screen as any data was revealed. The manual examination was primarily used for devices not compatible with the mobile forensic tools selected to be used (see Figure 2). The scope of the study remains unchanged, with a manual examination to identify whether PII was present on a given device. The manual examination was attempted for two out of the twenty-eight devices that could not be logically acquired.

The mobile devices were labeled and stored securely in a cabinet after the acquisition and for the duration of the study. Two mobile forensic tools were used for imaging and analysis of the mobile devices. The tools used were XRY Forensics v.7.6.2 and Mobiledit Forensic Express 5.2.0.12555. These tools were chosen because they were available to the researchers and were primarily designed for the imaging and analysis of mobile devices. In all cases, a logical acquisition was undertaken as the study was based on determining what user data could be retrieved. The acquisition was initially attempted with XRY Forensics for all mobile devices for consistency.
Verifying findings and device compatibility are the main reasons for employing both tools. No attempt was made to compare the effectiveness of these tools. A preview of each device’s data was initiated during imaging to determine whether any meaningful data were present in the device. For the preservation of evidence and to maintain the chain of custody, the analysis was performed on the image and not the original device, similar to the procedures used in previous studies and in line with industry best practice. Still, it is important to note at this stage that most of the data recovered from the devices could have been achieved even without the use of specialist software since the data was readily available on the devices. Even though similar studies have examined secondhand devices in the past, only two other studies of mobile devices were identified in the literature, and these were from more than a decade ago [15],[16]. Therefore, the sample for this study included all types of mobile devices up to twelve years old, from conventional mobile phones up to the latest technology smartphones and tablets. The oldest device on the sample was a Motorola StarTAC mr501, a mobile phone released in 1996. The purpose of including a wide range of devices was to compare the amount of data that could remain in older, more conventional devices in relation to the more recent and technologically advanced smartphones. A diverse sample of devices also identified different operating systems installed on the devices. Figure 3 demonstrates the distribution of mobile devices per vendor.
he market share and popularity of the devices [3] were also considered in the selection of the sample. The availability of the devices in the secondhand market during the sample selection played a significant role in the final purchase of the devices. However, in some cases, the researchers dealt with non-compatibility issues. This was largely restricted to a total of 12% (12 devices) of the purchased phones that were not smartphones and were mostly outdated. There was no pattern identified in the devices that were not compatible. The issue was mainly in relation to the tools that were selected for the study. In those cases, and when applicable and appropriate, a manual attempt was undertaken to identify whether data was present on the devices. As previously mentioned, this was feasible for only two out of the twelve devices.

The process that was followed for the examinations and analysis was the following:

1. Examine whether the device:
   a. contains data
   b. has been reset to the factory settings or data has been wiped
   c. is password protected
   d. is not recognized or is not functional

2. Keep a record of the device properties for those devices that belong to 1a, 1b, 1c, including model and operating system. Group 1d was excluded from further examination as it added no value to the study.

3. Use data extraction techniques with the support of the selected mobile forensics tools to recover data that contain PII, or any information that could distinguish an individual. All readable areas of each device are included in the analysis following a formal digital forensics methodology; contacts, messages, conversations, image files, geolocation information, email applications, social media retrievable data, etc.
4. Extract relevant data, filter out PII and anonymise the findings.

5. Analyse and measure the amount of retrievable PII and examine further those cases that a full identity could be recovered.


Further information was maintained for the devices that belong in group 1d – not recognized or not functional – depending on their unique characteristics. This included the reason why the device was not recognized, such as the device was password locked or a SIM card was requested. Similarly, when a device was not functional, such as constantly re-booting, the device was not responsive or did not start, the device was not charging, or arrived damaged. However, it is beyond the scope of this paper to expand on these findings.

5. RESULTS

The data extraction results from each individual device were analysed to better understand the findings and their significance in terms of PII. This section presents the results of the study in percentages out of the 100 devices that were analysed in the sample.

Overall, for 28% of the purchased devices, an image could not be acquired; for 72% of the devices, logical images were acquired. Figure 2 demonstrates the balance between the devices that images were acquired for and not acquired for during the study.

As an initial step, the operating system of the acquired devices was identified. Table 1 presents the different operating systems that were installed on the devices. Even though the versions of the operating systems were also identified, this aspect was not examined further as it was beyond the scope of this study. The most popular operating system in the sample was the Android with the equivalent of 40% of the devices, while the Apple iOS followed with 28% of the devices. The Blackberry OS was 10% of the sample, while Microsoft Windows and other operating systems equated to approximately 11% of the sample.

The analysis of the mobile devices revealed that 52% of all the purchased devices for the study were either reset to their factory settings or had all the user data effectively removed. A significant finding demonstrates the users’ familiarity and confidence with erasing their PII from their devices before selling them to the secondhand market. Only 19% of all the devices that were examined contained any data; however, in most of these cases, PII was fully retrievable and could fully identify the previous owner of the device.

In more detail, for 28 out of the 100 mobile devices purchased, (28 percent) an image could not be acquired for a variety of reasons; eleven of them were faulty, and seventeen were not recognised by the tools available. Either the connectors were not available, or the tools were returning errors. Two out of these seventeen devices that were not recognised were those used in the manual analysis.

Still, there were 72 mobile devices for which images were successfully acquired (72 percent); fifty-three of these devices were reset to the factory settings or had had all user data erased by the previous owner or the
Table 1. Identified operating systems on the devices for which an image was acquired

<table>
<thead>
<tr>
<th>Number of devices</th>
<th>Operating system</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>Google Android</td>
</tr>
<tr>
<td>20</td>
<td>Apple iOS</td>
</tr>
<tr>
<td>7</td>
<td>Blackberry OS</td>
</tr>
<tr>
<td>8</td>
<td>Windows Mobile</td>
</tr>
<tr>
<td>8</td>
<td>Other</td>
</tr>
<tr>
<td></td>
<td>Total Number of Imaged Devices</td>
</tr>
</tbody>
</table>

Two devices contained nothing that could be retrieved. One of these two devices was password protected but contained a sticker at the back of the device with the previous owner’s full name and home address in London. The remaining 17 devices (24 percent) contained data. The examination of the devices is focused on the residual data that can potentially contain PII and not the device’s manufacturer. Hence the findings are not linked to a device make and model.

Two devices contained an SD memory card. The SD memory cards were not treated as separate entities as they were part of the device’s storage area. However, it is worth noting that in these two cases, most of the information was retrieved from the SD memory card, and there had been no attempt from the previous owner to erase it. These cases are further discussed in the “Case Studies’ section.

As a result, PII was retrievable and could identify the previous owner for 17% of the total sample of devices. In some cases, there was no attempt to delete any PII before selling the phone. Five of the seventeen devices appeared to have simply been switched off before selling them, with no encryption method activated and no security settings to prevent anyone from accessing the personal information. Figure 4 illustrates the distribution of the attempted identification of data during the imaging process.

As a result, PII was retrievable and could identify the previous owner for 17% of the total sample of devices. In some cases, there was no attempt to delete any PII before selling the phone. Five of the seventeen devices appeared to have simply been switched off before selling them, with no encryption method activated and no security settings to prevent anyone from accessing the personal information.

Of the devices for which an image could not be acquired, the manual investigation revealed that two of the phones contained meaningful data. One of the two phones had rich content, with the previous owner’s email account containing 6464 emails in the inbox, while the last email was received in December 2017. The emails contained PayPal payments, including addresses and contact details of the buyers, indicating the phone’s previous owner was an eBay seller. It also contained 344 images in the gallery, 133 contacts, and text messages, including a text message in the Spanish language. The second phone that was examined without an image being acquired contained personal photos and a welcome message to the previous owner.
6. CASE STUDIES

In this section, we present the notable findings from the devices that were analysed. Each bullet point represents a single case from a single device.

As a general observation, the data retrieved from smartphone devices included Wi-Fi networks, MAC addresses, and a standard list of applications as part of the recovered files.

The findings of the analysis are:

- One phone contained an SD card, and all files were retrieved from there. The previous owner’s name was present, along with an email address and a home address in York. A ‘living social’ voucher that was purchased in December 2013 is valid until December 2022, which makes it valid at the time of writing. In addition to these, a P11D Expenses and Benefits form for 2012-13 was retrieved that contained: Employer’s name, PAYE (Pay As You Earn) reference, payroll number National Insurance Number and date of birth. The form was not fully completed. There was also a Hootsuite profile and some personal pictures linked to the owner’s profile.

- Another device contained a contacts list with thirty entries, including the previous owner’s number, a recent calls list, one hundred and fourteen text messages that included sexting, and seven multimedia messages.

- Another phone contained an SD card; again, most of the data was retrieved from there. The previous owner appears to be a university student at the time of using this mobile phone. Their phone number and email address were retrieved, including their bank account details, 532 personal pictures, and 16 videos. Also, a list of calls was retrieved from the phone in the period between October 2015 and January 2016. Several different GPS location information in the area around Birchington and Manchester in the UK were present on the phone. Also, some bookmarks and browsing history were retrieved as well as music files.
• On another device, the whole list of contacts was linked to Facebook, including the profiles of all 126 contacts. The device was still connected to the previous owner’s social media platform, and since the profiles of the contacts were retrievable, full names, dates of birth, work information, location, email addresses, and home addresses could be retrieved for some of them. The same information about the previous owner was also recovered. Also, present were chats from two WhatsApp groups, 13983 photos, and some family videos. In addition, a Booking.com reservation for Glasgow airport was retrieved, 2 Skype accounts, and an Instagram account.

• The previous owner’s mobile and home phone numbers were identified on another device. The area code appears to be in the Portsmouth and Southampton area. There were text messages from T-Mobile and promotions on the phone. The list of calls was from the period between January and April 2004. There were 46 contacts on the list and six text messages. Also, 740 pictures were retrieved, three videos and 31 recordings. Most pictures seemed blurred and included shots from television programmes, a motorbike, and family photos.

• On another device, the previous owner’s name, occupation, and home address were retrieved from the device and their email account. Sixty-five email messages were retrieved that mostly contained travel bookings and email confirmations, utilities information, and PayPal payments between April and June 2016. The place of residence appears to be Hunsworth in Yorkshire, UK. Their Apple ID and password were present and retrieved from the device, as well as their eBay user account name and password. Also, 408 pictures and three videos were retrieved, including visits to London, the British museum, and the island of Corfu, numerous selfies and a car with the registration number visible. The geolocation information revealed 375 different locations over the period the phone was used, from 2015 until 2018. The web browsing history was also retrieved from the phone.

• The activity on another device placed it in 2014. A full contacts and recent calls list were present, as well as four email accounts that were used on this phone. Both personal and some promotional emails were retrieved. Also, a text message from Lloyds bank was present, some finance details and a bank account number. In addition, the device contained calendar events, one multimedia message, five pictures, and a video.

• Another device dated back in 2012 had the last owner’s name and phone number written on the box it was delivered in. The device itself contained six contacts and three text messages.

• A tablet device contained the previous owner’s email account that was connected, logged in, and active.

• Another device was sold with the SIM card present that revealed the phone number. The previous owner’s full name and a list of recent calls were retrieved. In addition, the Blackberry PIN for Blackberry messenger was present, together with calendar events, one note, and a Tesco mobile welcome text message.

• Another device contained home and work geolocation coordinates, together
with some family pictures, a voice note, and a video.

- A device that appears as a schoolchild’s phone from Ringwood, Hampshire was identified. Some contacts and two notes were present on the phone, including a landline number that allowed us to identify the area from the area code.

- The mobile number was retrieved from a pay-as-you-go Tesco Mobile phone. The contacts present only corresponded to Tesco mobile services.

- A device with the Orange network contained a history of 1 call received. It seems as if the device had very light use after a factory reset and before it was sold.

- A picture gallery from 2010-2013 was retrieved from another phone. The previous owner’s mobile number was present, as well as draft text messages and mobile numbers for their contacts. The recent calls list contained calls that the area codes indicated the areas of Alton, Cheltenham, and Carmarthen in the UK.

7. CONCLUSIONS, RECOMMENDATIONS AND FUTURE DIRECTIONS

This study found that 52% of the devices were reset to the factory settings or had had the data effectively removed. This result indicates that the users considered erasing their PII before selling their devices to the second-hand market and offered a slightly different perception to the existing secondhand device studies. The data was retrievable from most of the devices when different types of electronic devices were examined in other similar studies. However, modern smartphones and tablets offer several advantages related to communication and accessibility to their users. The low level of knowledge and effort that it takes a non-technology or computer literate user to reset their device to the factory settings is indicative of the results from the study.

On the other hand, most devices containing PII could identify the previous owners. This only applied to 17% of the total sample of mobile phones that was examined. At the same time, the study results indicate that a forensic analyst can achieve 17% successful retrieval of PII, the same as a malicious buyer who intends to commit fraud. Even though the majority of the devices were reset, the threat of PII leakage is evident since the users had made no effort to contain or eliminate their personal data.

This study used specialist forensics software to acquire and analyze the devices to preserve potential evidence. In many cases, though, PII was available in plain sight. The devices that contained data that was retrievable with specialist software also had an indication of available content when powered on. Nevertheless, there were purchased devices that included phone numbers, names on the phone’s box or on the device itself, and in one case, a home address.

This research will be repeated and grown periodically in the future. It is also being considered to expand the study to countries with large secondhand smartphone markets, such as India and China. This could provide a more robust sample of the PII that can be found in the secondhand mobile devices market coming from countries with large populations and reputable secondhand markets. Various challenges could emerge when expanding the study in that direction, espe-
cially during the collection and analysis phase. The increased cost of purchasing the devices would need to be considered for sampling and shipping costs. However, the main barrier could be the language. The team will need to expand and include researchers who could understand the selected markets’ languages.

A weakness identified in the methodology followed for this study was the wide range of portable devices purchased, as some of the devices were outdated, and it was impossible to acquire an image and analyse them with the tools and software available.

The problems arising from the disposal of secondhand portable devices are likely to increase due to the availability and constant technological advancements in these devices. A future version of this study will focus on more recent types of devices to compare the results with the current study.


[17] Garfinkel, Farrell, Roussev and Dinolt, Bringing Science to Digital Forensics with Standardized Forensic Corpora, DFRWS 2009, Montreal, Canada


[26] Szewczyk, P., Robins, N. and Sansurooah, K., (2013), Sellers continue to give away confidential information on second hand memory cards sold in Australia.

Market In the UAE. Journal of Digital Forensics, Security and Law.


