Forensic Analysis of Geodata in Android Smartphones

Stefan Maus, Hans Höfken, Marko Schuba

FH Aachen, University of Applied Sciences,
52066 Aachen, Germany
maus@aixis.net, hoefken@fh-aachen.de, schuba@fh-aachen.de

Abstract. Smartphones provide a lot of interesting data for a forensic investigator. As many apps make use of location information to provide or enhance their services, the log files of smartphones can contain valuable geodata. Based on geodata it is possible to reconstruct the whereabouts of the phone and its user at certain points in time. The functionality to analyse geodata based on GPS coordinates is standard for many mobile forensics tools. Unfortunately, many smartphone apps log their geodata in human readable format only, which makes it harder for investigators to get and analyse this important information. This paper proposes a simple but comprehensive approach to the analysis of all geodata on a smartphone. The approach uses simple app-specific descriptions of the geodata information stored in smartphones and thus enables an automatic extraction and presentation of all relevant geodata. The approach is currently being implemented for Android smartphones.

Keywords: mobile, smartphone, forensics, geodata, Android

1 Introduction

The market for smartphones has been rapidly increasing during the last few years. In quarter four 2010 the worldwide shipments of cell phones totalled 401 million units [1]. About 25 percent of those units (101 million [2]) have been smartphones. What distinguishes those smartphones from feature phones and "dumb phones" is their higher computing ability and connectivity, which allows end users to install and run advanced applications. Characteristics of smartphones include:

- They are no longer optimized for telephony but for usability of a broad range of applications. This is also reflected in the user interface, which typically consists of a large screen with an alphanumeric keyboard or a touchscreen.
- They often run a complete operating system with an open API, allowing users to install applications developed by third parties.
- They offer a variety of connectivity methods, for example WLAN and Bluetooth (for local wireless access), GSM, UMTS, HSDPA etc. (for wide area mobile access) and USB (e.g. for synchronizing with a local computer).

- They contain additional sensors like accelerometers, magnetometers, light sensors, proximity sensors or GPS receivers.

The combination of this rich feature set with data flat rates, which are nowadays common among operators and internet service providers, has led to a vast number of smartphone applications (so-called apps). Apart from the applications that are known from ordinary PCs, new types of apps have emerged that make use of the sensor capabilities of the smartphones. Games, for instance, use the accelerometers (in combination with the touchscreen) as user interface for game control. Navigation apps, which formerly required special devices, can now easily be developed as the phone's position and direction is known, and maps are available online.

From a digital forensics' point of view, smartphones offer great opportunities but also cause a lot of problems. While the forensic analysis of ordinary cell phones typically results in a well known set of data (e.g. call history, text messages, contacts, photos) an analysis of a smartphone reveals a plethora of information, because each app stores application-related data. The type of data ranges from emails, over browser history, chat conversations to routes of the navigation app. In crime investigations this application-related data could contain very useful evidence. Unfortunately, the retrieval of information from a smartphone is not trivial. First of all, there exist a number of different operating systems for smartphones, the top four currently being Android (33% market share), Symbian OS (31%), iOS (16%) and Blackberry OS (14%) [2]. Phones based on those operating systems differ with regard to data storage, file system or interface, sometimes even for different releases of the same smartphone vendor. As a consequence, methods for the logical or physical analysis need to be reinvented for almost every new smartphone released. On top of this problem, the mass of apps in the market, which store their data in different, sometimes proprietary ways, make a forensic analysis even more difficult.

This paper focuses on the forensic analysis of a type of smartphone data that is particularly interesting for crime investigations: geographical (or spatial) data, also called geodata. A large number of apps in smartphones use geodata to provide additional functionality, for instance by providing location specific information or to store the location as context information for later usage. Geodata that is logged in the phone can be forensically extracted and used to provide a (graphical) location history of the phone (the user).

The analysis done in the context of this paper has been conducted on an Android smartphone (HTC Desire) running Android 2.2. Android has been selected as Android-based cell phones currently have the highest growth rates [2] and an application base of more than 200,000 apps [3], what makes Android phones and apps a likely candidate in crime investigations.

The rest of the paper is structured as follows. Section 2 gives a brief introduction to the determination and forensic usage of geodata in smartphones. Section 3 describes state-of-the-art approaches of geodata forensic analysis today. Section 4 presents an extended approach for the analysis of geodata on an Android smartphone. Section 5 explains the current implementation status of the approach in a small tool called AFT (Android Forensic Toolkit) and what work is planned for the future. Finally, section 6 provides a short summary of the presented work.

2 Geodata in Smartphones

2.1 Measuring Geographical Positions

Various methods exist to determine the geographical position of a cell phone. Most common are GPS, using cell-ids or signal strength measurements of mobile networks or using the locations of WLAN hotspots in the mobile's proximity (for an overview see e.g. [4]). From a digital forensics' perspective, the method used to obtain the position is of importance, as it determines the accuracy of the collected geodata. Table 1 depicts the accuracy that can be obtained by different systems to date.

Table 1. Accuracy of example positioning systems

Positioning Method	Accuracy	Usage
GPS	± 8 m	Outdoor
Assisted GPS (aGPS)	5-50 m	Indoor and outdoor
Cell-id	100-3000 m	Indoor and outdoor
GSM Cell Tower Triangulation	± 25 m	Indoor and outdoor
WLAN Positioning System	20-30 m	Indoor and outdoor

Most of today's smartphones are equipped with assisted GPS (aGPS) and WLAN, so the positions determined will be quite accurate in most cases.

When using aGPS, the geodata is delivered to the smartphone using a latitude/longitude-coordinate system that presents locations in degrees from 180° west through 180° east along the Equator and 90° north through 90° south along the prime meridian. One description format is [-]d.d, [-]d.d in decimal degrees with negative numbers for south and west. Fig. 1 shows an example of a GPS position that was cached on an Android smartphone. Note that apart from latitude and longitude the data set includes an accuracy attribute (in meter) and a timestamp attribute (UTC time in milliseconds since Jan 1, 1970 GMT).

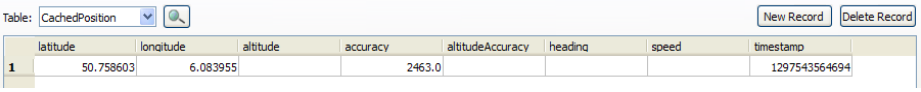

Fig. 1. Logged GPS position in an Android smartphone

2.2 Relevance of Geodata for Digital Forensics

Obviously, a comprehensive history of location information with timestamps can be useful in a crime investigation. But can it be expected to find a sufficient amount of geodata in an acquired smartphone? The answer is: yes, it is very likely. The rapid increase in number of smartphones during the last two years has led to the development of a vast number of location aware apps, many of them being very popular. For the targeted smartphone platform Android a recent study of almost 50,000 apps showed that 25% of the apps requested access to coarse location

information. 15% of the apps even requested fine grained access to location information [6]. A few examples of widely spread apps that use geodata:

- Smartphones come with a set of pre-installed apps, including map or navigation systems.
- Most smartphones also have a built-in camera. When geo-tagging is enabled, those cameras store the geographical position as part of the image's metadata.
- Some third party location based apps have become extremely popular. Foursquare, an app to explore cities and sharing locations, had 6.5 million users by February 2011 [7]. Facebook, the world's leading social network, announced their location based service "Places" in August 2010. Within two months about 30 million people tried it out [8]. The famous microblogging service Twitter also added location information as an option to their Tweets during 2010.

As many location based apps log their activities there should be a sufficient amount of geodata stored in the smartphone, justifying the effort to do a forensic analysis on this data.

3 Existing Approaches to Smartphone Geodata Analysis

Forensic tools analyse data of a seized smartphone in a number of steps. The first step is to extract as much data as possible from the smartphone and its storage media. This is done in various ways that can differ a lot between different smartphones. A general distinction is the logical or physical acquisition of data. The result is a logical or physical memory image that can be analysed further.

In the next step, data seen as relevant (e.g. photos, call history etc.) is extracted from the memory image. In order to be able to extract geodata from a smartphone, forensic tools first need to find the relevant data in the memory image. Depending on storage medium and type of geodata, the approach can be different

As any smartphone provides cell phone functionality, the location information related to the mobile network usage can be extracted in a standard way from the phone's SIM card. This method is supported by a number of forensic tools (see e.g. [9] for an overview).

When it comes to the geodata of apps on the smartphone: their storage is mostly proprietary. For each combination of smartphone, platform and app, the geodata storage might look slightly different regarding the geodata format and storage location. The usual approach of forensic tools is therefore to either look for specific geodata formats (usually GPS coordinates) or to analyse data of specific applications (for which the geodata storage location and format is known). The review of 13 forensic tools conducted in [10] showed that many tools indeed support those methods: in this review tools were rated (among other things) according to their capability to find GPS coordinates of different apps or to extract the location history of a specific app (Google Maps in that case).

Unfortunately this approach leaves a wide area of potential geodata evidence untapped: geodata of many apps that is not stored in the usual GPS format. Why is

this relevant? Instead of storing numbers of longitude and latitude, which are unreadable for users, many apps convert GPS coordinates to textual address information before they log them. Fig. 2 and 3 show two examples of apps that store geodata only in text format.

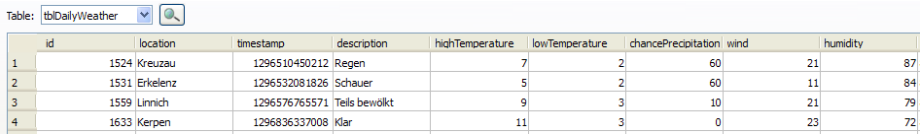

Table: tblDailyWeather

	id	location	timestamp	description	highTemperature	lowTemperature	chancePrecipitation	wind	humidity
1	1524	Kreuzau	1296510450212	Regen	7	2	60	21	87
2	1531	Erkelenz	1296532081826	Schauer	5	2	60	11	84
3	1559	Linnich	1296576765571	Teils bewölkt	9	3	10	21	79
4	1633	Kerpen	1296836337008	Klar	11	3	0	23	72

Fig. 2. Logged city locations of a weather app
(app presents weather forecast at current user location)

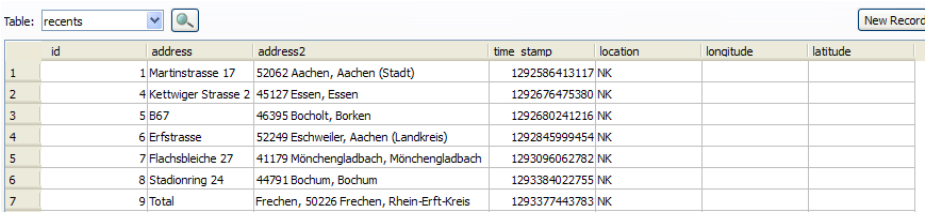

Table: recents

New Record

	id	address	address2	time_stamp	location	longitude	latitude
1	1	Martinstrasse 17	52062 Aachen, Aachen (Stadt)	1292586413117	NK		
2	4	Kettwiger Strasse 2	45127 Essen, Essen	1292676475380	NK		
3	5	B67	46395 Bocholt, Borken	1292680241216	NK		
4	6	Erfstrasse	52249 Eschweiler, Aachen (Landkreis)	1292845999454	NK		
5	7	Flachsbleiche 27	41179 Mönchengladbach, Mönchengladbach	1293096062782	NK		
6	8	Stadionring 24	44791 Bochum, Bochum	1293384022755	NK		
7	9	Total	Frechen, 50226 Frechen, Rhein-Erft-Kreis	1293377443783	NK		

Fig. 3. Logged streets and city locations of a navigation app
(note that GPS coordinates are not logged)

While these logs make the geodata useful in the context of the app, it makes it more difficult for forensic tools to find and present them. New solutions are needed that also make use of this geodata in forensic investigations. One possible approach is described in the following section.

4 A Comprehensive Approach for the Forensic Analysis of Geodata

In order to do a comprehensive analysis of smartphone geodata, a modular approach is proposed (see Fig. 4). Even though this method currently focuses on smartphones based on the Android platform, it can easily be extended to other platforms.

4.1 Data Acquisition

The first step of the approach is to acquire all data from an Android smartphone. Compared to an ordinary PC, where all applications are executed under the rights of the currently logged in user, all Android apps are executed as separate processes with individual permissions each. By default, processes are not allowed to access data of other apps. Gaining access requires an explicit permission.

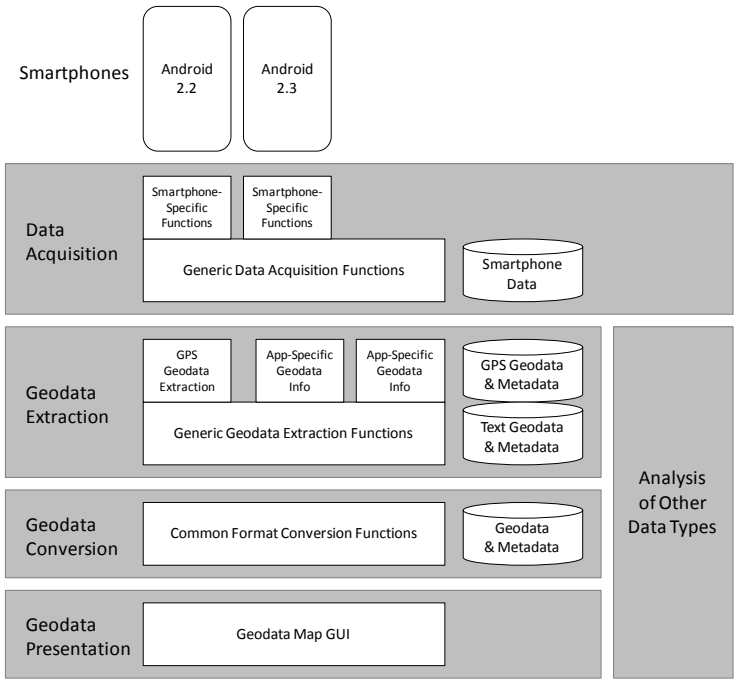

Fig. 4. Modular Architecture for Geodata Forensic Analysis

As a consequence a forensic extraction process needs suitable access rights to get the data from all apps of a smartphone. Such a general access right can for instance be obtained by rooting the smartphone. The detailed rooting process is not part of this paper: there are many documents and instructions available that describe the details for different smartphone types (e.g. [11], [12] or [13] for Android). For the results described here, an HTC Desire with Android 2.2 (Froyo) was used, which was rooted as described in [12].

The copied data is stored in a database (Smartphone Data, see Fig.4) and automatically sorted according to their generating apps. In the current implementation, the database actually consists of a set of app-specific SQLite databases. From these databases data can be extracted for different examination purposes.

4.2 Geodata Extraction

GPS geodata can be extracted by scanning the databases for attribute names (e.g. “longitude”, “latitude”) or by a structured analysis of known data attributes of installed apps or stored files (e.g. geo-tagged photos). Once the data has been extracted, it is stored in a database for GPS geodata.

Textual geodata is more tricky to identify. A generic approach is to scan the databases for keywords like ”location” or “address”. While this might deliver many hits, the manual examination of the data will take quite some effort.

A more efficient approach to extract textual geodata is to use short descriptions of the data attributes for different apps. Looking back at the two example apps shown in Fig. 2 and 3, the attribute structure is very simple. A description for extracting the textual geodata only needs a few parameters: the app name, its database, the attribute name(s) of the textual geodata and the names of further metadata of interest. Below, a brief XML description for the two example apps of Fig. 2 and 3 is given.

```
<textgeodata>
  <app>
    <name>LocalWeather</name>
    <dbname>weather.db</dbname>
    <dbtable>tblDailyWeather</dbtable>
    <city>location</city>
    <time>timestamp</time>
  </app>
  <app>
    <name>Navigon</name>
    <dbname>navigon.db</dbname>
    <dbtable>recents</dbtable>
    <street>address</street>
    <city>address2</city>
    <time>time_stamp</time>
  </app>
</textgeodata>
```

Based on a set of such short descriptions, which are either available or can be easily produced for a new app, textual geodata can be quickly extracted from an app’s database and stored in a Text Geodata Database.

4.3 Geodata Conversion

The Geodata Conversion functions read geodata and metadata from the GPS and Text Geodata Databases and convert them into a standard format for further processing.

The conversion of geodata can be performed in both directions. Based on a textual address, the GPS coordinates can be easily determined using existing tools or services. An example for such a geodata conversion service is Google’s Geocoding API, which delivers latitude and longitude coordinates for a text based address [11]. To find out the coordinates of the author’s university (for example), you can simple use the URL

```
http://maps.googleapis.com/maps/api/geocode/xml?address
=Eupener+Str+70,+Aachen,+Germany&sensor=false
```

which returns an XML structure including the following GPS coordinates

```
<geometry>
  <location>
    <lat>50.7594900</lat>
    <lng>6.0826900</lng>
  </location>
</geometry>
```

In a similar manner, GPS coordinates can be converted into a human readable address. Taking the coordinates from above and using the URL

```
http://maps.googleapis.com/maps/api/geocode/xml?latlng=
50.7594900,6.0826900&sensor=false
```

the API returns an XML structure containing a textual address:

```
<result>
  <type>street_address</type>
    <formatted_address>Eupener Straße 70, 52066
    Aachen, Deutschland</formatted_address>
</result>
```

Together with the Geodata related metadata (e.g. timestamps, tagged element, app) both GPS coordinates and textual address are stored in a database, which is the basis for the graphical presentation in the next step.

4.4 Geodata Presentation

Now that a comprehensive set of geodata from different apps is available, it can be presented to the user. For a forensic investigation a graphical presentation of geodata in maps can be very useful. For the approach presented in this paper the following information has been selected to be presented:

- Map of the area including pins for investigator-selected locations.
- Context information in the form of metadata that has been extracted with the geodata (e.g. timestamps).
- Potential routes between the locations.

In order to generate the map, a standard map service (like Google Maps) with suitable API can be used [11]. Fig. 4 depicts an example of such a graphical representation, taking the first geodata entries of the navigation example of Fig. 3.

Fig. 5. Graphical presentation of textual geodata

From this presentation format an investigator can quickly identify, at which different locations the phone (user) has been at certain points in time. The routes also give indications on the movements of the user, which can be correlated to other evidence of the particular case.

5 Implementation Status

The implementation of the approach is ongoing as part of a project “Android Forensic Toolkit – AFT” at Aachen University of Applied Sciences. The work started with the data acquisition module, which has been implemented in C# and is currently working for the Android 2.2 version on a HTC Desire (see Fig. 6 for a screenshot of the latest AFT version).

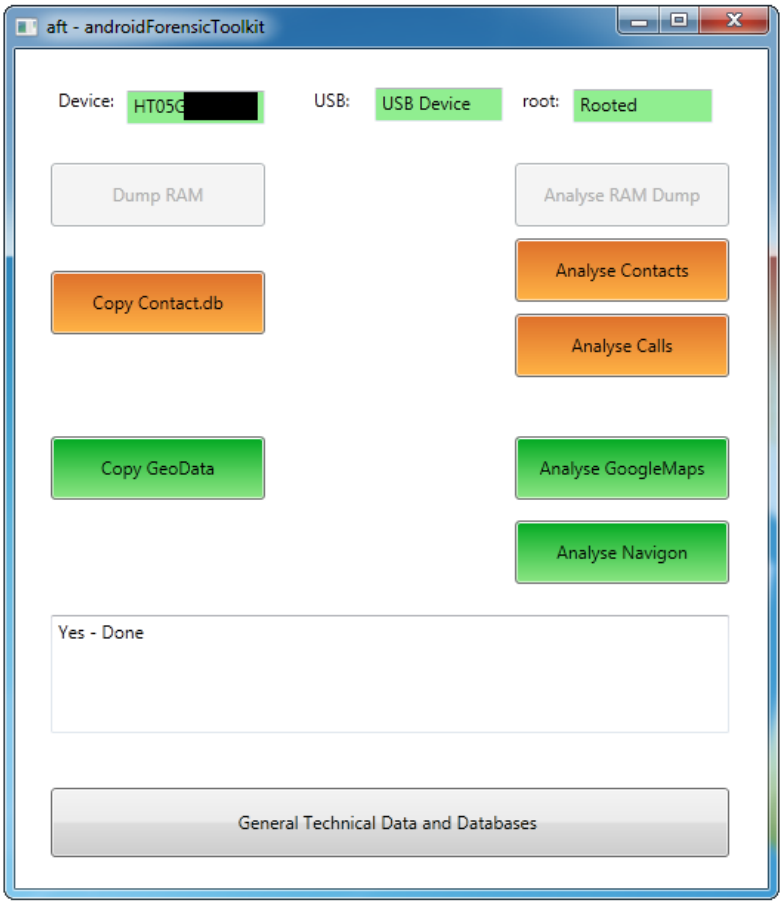

Fig. 6. Android Forensic Toolkit (Status Feb 2011)

The extraction, conversion and presentation modules are work in progress. The geodata presented in the context of this paper has been manually extracted from the acquired data and semi-automatically processed further.

Besides the automation of the geodata related functions, it is planned to implement support for other data types and other Android versions in later stages of the project.

6 Conclusions

Geodata logged on smartphones can provide forensic investigators with interesting information on the whereabouts of a person in certain time periods. The more geodata

available, the more accurate can the location history be reconstructed. In order to get as much geodata out of smartphones as possible, not only GPS but also textual geodata needs to be considered. With the presented approach that is currently implemented in a tool called Android Forensic Toolkit, such a comprehensive analysis is possible for the many location based apps that are available for smartphones today.

References

1. Zeman, E.: Top 5 Handset Makers Of 2010 Ranked, InformationWeek, http://www.informationweek.com/news/mobility/smart_phones/showArticle.jhtml?articleID=229200009 (Jan 18, 2011, retrieved on Feb 15, 2011)
2. Canalys research release 2011/013, http://www.canalys.com/pr/2011/r2011013.html (Jan 31, 2011, retrieved on Feb 15, 2011)
3. Brian, J.: Google Inc. (NASDAQ:GOOG) Android Has 200,000 Apps, http://www.stockbriefings.com/google-inc-nasdaqgoog-android-has-200000-apps-2/3179516 (Dec 31, 2010, retrieved on Feb 15, 2011)
4. Liu, H., Darabi, H., Banerjee, P., Liu, J.: Survey of Wireless Indoor Positioning Techniques and Systems, IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews, Vol. 37, No. 6 (Nov 2007)
5. Ionescu, D.: Geolocation 101: How It Works, the Apps, and Your Privacy, PCWorld, http://www.pcworld.com/article/192803/geolocation_101_how_it_works_the_apps_and_your_privacy.html (Mar 30, 2010, retrieved on Feb 15, 2011)
6. Vennon, T., Stroop, D.: Threat Analysis of the Android Market, http://threatcenter.smobilesystems.com/wp-content/plugins/download-monitor/download.php?id=8 (Jun 21, 2010, retrieved on Feb 15, 2011)
7. Foursquare – About, http://foursquare.com/about, (Retrieved on Feb 16, 2011)
8. Carlson, N.,: Foursquare Doomed? Facebook Places Has 7X More Users, http://www.businessinsider.com/facebook-places-may-have-30-million-users-but-none-of-them-use-it-very-much-2010-10, (Oct 29, 2010, retrieved on Feb 16, 2011)
9. Ayers, R., Jansen, W., Moenner, L., Delaitre, A.: Cell phone forensic tools: an overview and analysis update, NIST Technical Report 7387 (Mar 2007)
10. Hoog, A., Strzempka, K., iPhone Forensic White Paper, http://viaforensics.com/education/white-papers/iphone-forensics/, (Nov 2010, retrieved on Feb 16, 2011)
11. Wise, J.: Unrevoked, http://unrevoked.com/ (Jan. 20, 2011, retrieved on Feb 17, 2011)
12. Ryan, P., Unrevoked3, HTC Desire Root, http://unrevoked.com/recovery/ (Jan. 20, 2011, retrieved on Feb 17, 2011)
13. Bear, C. Forum XDA Developers http://forum.xda-developers.com/showthread.php?t=803682 (Oct 9, 2010, retrieved on Feb 17, 2011)
14. The Google Geocoding API, http://code.google.com/intl/en/apis/maps/documentation/geocoding/, (Retrieved on Feb 17, 2011)