3-22-2019

Google and Facebook Data Retention and Location Tracking through Forensic Cloud Analysis

Elizabeth Williams
Middle Georgia State University, johnathan.yerby@mga.edu

Johnathan Yerby
Middle Georgia State University, johnathan.yerby@mga.edu

GOOGLE AND FACEBOOK DATA RETENTION AND LOCATION TRACKING THROUGH FORENSIC CLOUD ANALYSIS

Beth Williams
Middle Georgia State University
Elizabeth.Williams7@mga.edu

Johnathan Yerby
Middle Georgia State University
Johnathan.Yerby@mga.edu

ABSTRACT

Mobile devices have hardware and software components that record large amounts of data. Some of the data is apparent to the device owner, some is discarded quickly, and some is hidden from the person using the device. For this study, the researchers used an Android smartphone as a typical user, carrying the device throughout the day, using Facebook and Google applications. Then the smartphone was analyzed using mobile forensic techniques and software. The investigation revealed security and privacy concerns. The researchers were able to retrieve social interactions, pictures, documents, and other personal attributes stored on the device. The most interesting find was location tracking information. This Android phone logged and stored location data when the researcher had location services enabled, but it also continued to collect and store location information after turning location services off. Within Google Maps, the sub-feature called Google Timeline, tracked location, date, and time as long as the phone was powered on. These findings will increase awareness for mobile devices users and may lead to more consumer-centric privacy settings in mobile operating systems.

Keywords

Privacy, security, social applications, Android, tracking, forensics, location

INTRODUCTION

In December 2018, Sundar Pichai, Google's chief executive was called to testify before the U.S. House Judiciary Committee. There were three themes that the Judiciary Committee was concerned with, focusing on distrust of large tech firms, specifically Google, privacy practices, and concerns about tracking user location (Wakabayashi & Kang, 2018). U.S. Senator Ron Wyden, who recently proposed legislation to limit the collection and sale of location data, said "Location information can reveal some of the most intimate details of a person's life — whether you've visited a psychiatrist, whether you went to an A.A. meeting, who you might date" (Valentino-DeVries, Singer, Keller, & Krolick, 2018). Privacy, data collection, and data utilization has garnered more news headlines in the past five years with major stories including Facebook's Cambridge Analytica scandal, passing the General Data Protection Regulation in Europe, and the Equifax breach exposing at least 143 million Americans information (Chatillon, 2018).

For most people in developed countries, smartphones have become vital communications devices to participate in a modern life. The smartphones that are within arm's reach of users for most hours of every day, contain powerful hardware and software that can track the activities going on within and near the device. The Androids and iPhones in people's pocket generate, transmit, use, and record, user, system, and application data locally, and in cloud accounts. Location is an important concept that could lead to being involved in a criminal case, cause embarrassment, or a myriad of other issues (Schilt, Hong, and Gruteser, 2003). People by nature prefer not to be monitored, but to have all the benefits of modern applications (Yerby, 2013). The end user of the device is not always able to see or control what the smartphone is doing. Device users have reported being averse to installing applications that have access to contacts or tracking location, but numerous Google applications which run the devices operating system, phone, and map do not leave the user with much of an alternative (Harris, Brookshire, Patten, and Regan, 2015).

Google's privacy and data policy details data tracking mechanisms across all of Google's applications (Google Chrome, Google Maps, YouTube, etc.) This policy identifies "things you search for, videos you watch, ads you click or view, your location, websites you visit, and apps, browsers, and devices you use to access Google services" (Google Privacy, 2018) as information tracked by Google. In addition to this data collection, content created by users (such as e-mails, photos, videos, documents, YouTube comments, contacts, calendar events, etc.) are also logged by Google (Google Privacy, 2018). People sign into their Google accounts on many devices and then the activity is used to build a profile which details what that person is doing on their personal device, work device, or any device where they access Google applications. Each device connection to Google services provides Google with a more detailed view into the user's personal life. This data logging can be used to identify gender, ethnicity, employment status and current/previous employer, political affiliation, sexual orientation, etc. and can also be used to tailor news and search results to a user's profile without their consent for content tailoring.

Facebook's privacy policy (revised April 19, 2018) explained data collection methods and how gathered information could be utilized. Facebook intentionally makes their policies vague and difficult to pin down. Facebook's data collection policy outlined data tracking of "people, pages, accounts, hashtags, and groups you are connected to and how you interact with them" across Facebook products (Facebook, 2018). Device information such as device attributes, device signals, data from device settings, and cookie data allow Facebook to obtain device operating system version, battery level, hardware and software version, available storage space, Bluetooth signals, nearby wireless access points, GPS location, camera and photo access, and cookie data from websites visited on paired devices (Facebook, 2018). Interactions outside of Facebook are gathered through Facebook Business Tools. Through this platform, website developers and third-party applications share user data from outside of Facebook to Facebook developers. "Profiles" are built for each Facebook user based on this collected data which allow for targeted advertisements based on website history, application activity, and GPS location (Facebook, 2018).

Lessard & Kessler, 2010 declared that 46.3% of global use of mobile devices originated from the United States. This prevalence reiterated a need for standardization in collection and analysis tools as an increase in mobile data storage is realized. Researchers present methods for obtaining stored mobile device data through enabling USB debugging, rooting the device, examining the memory, and recovering the contents using Access Data's Forensic Tool Kit (FTK) v1.81. The results of their analysis displayed all images, documents, plain-text passwords, web searches, a Google Maps database of previous locations, call and text records, social media information, e-mail transmissions, etc. (Lessard & Kessler, 2010).

In 2011, Vidas, Zhang, & Christin presented their findings regarding a uniform methodology for Android device data collection. Acknowledging the prevalence of the Android platform and its ability to yield substantial amounts of user data, their research identified a standard collection process for data acquisition while preserving data integrity. The author demonstrates a solution for best practices while acknowledging manufacturer variances (Vidas, Zhang, & Christin, 2011).

In 2012, Stirparo & Kounelis presented their findings regarding privacy and security concerns of Android devices. Researchers identified target applications, populated data on a test device, acquired test data through digital forensic software, and analyzed the collected data results. Digital forensic analysis was conducted which yielded user files, images, location information, application credentials, application activity, and other concerning data from tested applications (Stirparo & Kounelis, 2012).

This study examined suspicions that Google applications collect and store information about the device location. The research questions in this study were:

R1: Do Google applications on an Android mobile device store location information locally on the device?
R2: Do Google applications on an Android mobile device store location information in the cloud?
R3: Do Google applications on an Android mobile device record location data when location services are disabled by the user?
R4: Does Facebook track and store interactions within the Android Facebook application?

To answer the research questions, the researchers analyzed stored application data through natively installed, cloud-hosted, and social media applications to identify security and privacy issues. The researchers used a Samsung Galaxy Note 5 using the Android OS to examine what data existed for Google's applications (G-Mail, Calendar, Hangout, and Google Timeline.), Facebook, and Facebook Messenger. Location services were enabled for a week, to simulate a typical user interaction with the device and then disabled for a week. The researchers periodically interacted with Google Maps, Google Calendar, phone, and Facebook applications during both phases of the experiment. Following the data collection period, an analysis of the week with location services enabled vs. the week with location services disabled were compared. Regardless of whether the location settings were set to enabled or disabled, Google Timeline was still actively tracking location. Google Timeline yielded an exact location for every moment of every day with timestamps and movement types (walking, driving, etc.). Further analysis with Paraben detailed Google Maps links and GPS coordinates for each location tracked through Google Timeline. This research details the applications and accounts that recorded information from the Android device used in this study, specifically when analyzed using mobile forensics tools.

METHODOLOGY

In this study, a Samsung Galaxy Note 5 was prepared by performing a factory reset and then updates. The Android OS version was updated to Android 7.0. Google applications, Facebook, and Facebook Messenger were installed and updated. Then USB debugging and developer tools were enabled to root the device. The root process was completed utilizing ODIN, SuperSU, TWRP recovery, and associated Samsung Galaxy Note 5 drivers and verification through the RootChecker application. The device was rooted to gain root access to the physical storage. Samsung, Google, and Apple have memory, file, and partition encryption which means that the device must be rooted or jailbroken before it can be acquired (Magnet Forensics, 2017). Without the device rooted, the investigators would have been limited to acquire a logical acquisition or nothing at all. Google

Backup and Google Location initially used the default application access of enabled with Google Location later being changed to disabled. The device was used for two weeks, as a typical smartphone user would use applications, messaging, and maps.

Following the initial application setup, a series of controlled tests were conducted. For Google applications, the researcher created calendar events, sent and responded to e-mails through G-Mail, sent chat messages through Hangouts, viewed, liked and commented on YouTube videos, created and edited documents in Drive, uploaded photos in Drive, browsed Chrome in regular browsing mode and in incognito mode. For Facebook, the researcher added friends, created posts, commented on posts, shared photos, and links, joined multiple groups, created events on Facebook, joined events, shared location on Facebook, and communicated via Facebook Messenger. Several tests followed a pattern such as creating an event in Google Calendar, navigating to that event using Google Maps, and then posting about the event on Facebook. Numerous instances of each test scenario were created to ensure consistencies in the data findings and establish a "baseline" for standard data patterns.

After two weeks of using the smartphone, the researchers connected the device to a Windows computer with Paraben Electronic Evidence Examiner (E3) installed (Paraben, 2018). The researchers followed the basic process of an investigation including response, data gathering or seizure, acquisition, analysis, and reporting (Yerby, Hollifield, Kwak, & Floyd, 2014). Paraben E3, previously known as Device Seizure, can reliably acquire logical and physical forensic images from mobile devices according to the NIST guidelines for mobile devices (Ayers, Brothers, and Jansen, 2014). There was a passcode on the Android smartphone, but the passcode was entered prior to data acquisition. The forensics software established a connection with the device and the researchers made a physical and cloud acquisition of the device along with the attached storage, there was no SIM card. The forensic image was saved to an evidence file within the E3 case file. All the acquisition steps followed NIST mobile device guidelines. The evidence was parsed, and the researchers proceeded with forensic analysis.

RESULTS

The software calculated the total size and a hash value that would be used to verify the integrity of the image later. Figure 1 displays the Google and Facebook accounts located by the cloud acquisition. The Google and Facebook credentials were cached on the device, which allowed the investigators access to the cloud data. The combination of the physical acquisition and cloud import yielded a 399-page document that displayed content from file system architecture to exact locations recorded by Google Location. The investigator was able to review text message details, call records, voicemails, alarm clock data, Bluetooth configuration, audio files, photos, installed applications (and associated application access), and some cached passwords. The physical acquisition portion revealed capabilities that could be compromising to those concerned with data security and privacy. Each of the applications revealed substantial data retention that could prompt security concerns among recurrent users. Regarding research question 1, the Google applications did store information such as contacts, text messages, photos, and searches on the local device, but the location tracking data was stored in the user's cloud account as theorized by research question 2 in this study. Google Timeline was proven to store location tracking information in the cloud account for the smartphone in this experiment regardless of whether location services was enabled or disabled by the user.

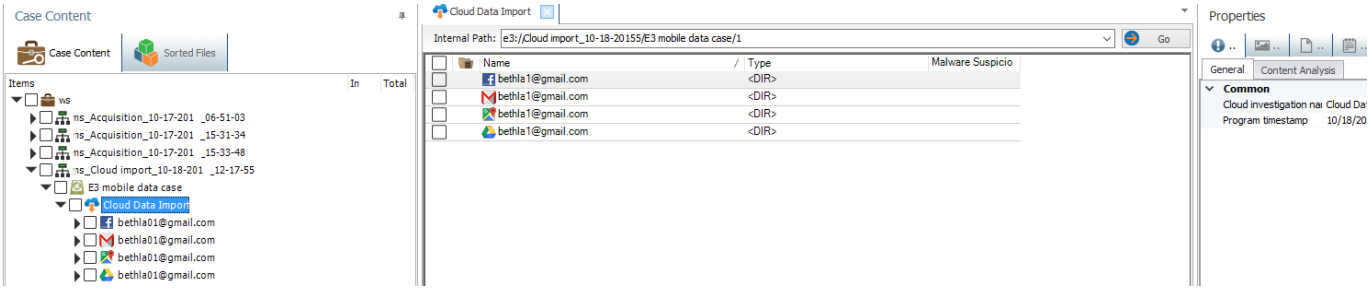

Figure 1. Cloud Acquisition displaying G-mail, Drive, Maps, and Facebook account

Facebook

A Facebook account was created on the smartphone to simulate user social interactions through posting, commenting, and chatting with other Facebook members. Following the two-week collection period, the researchers were able to retrieve every interaction that occurred on Facebook. The researchers could view detailed insight into profile attributes, posting times, and links to profiles of all users that had interactions with the Facebook account. Figure 2 details the associated file structure. Figure 3 displays detailed user interactions through the Facebook mobile application. All the reported information was searchable and could be exported.

Google Application Suite

A Google account was created on the Galaxy Note 5 device. The researchers sent and received e-mail that contained both text and images, created calendar events in Google Calendar, used Google Hangouts, created and edited files in Google Drive, and used Google Maps. Following the acquisition, details of every action were retrieved utilizing Paraben E3.

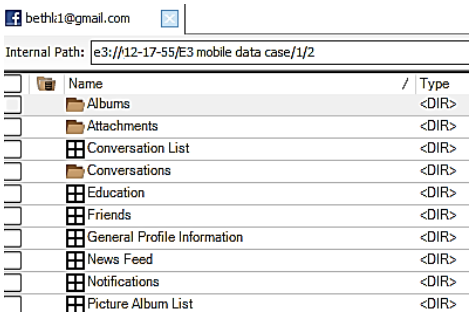

Figure 2. Facebook file structure

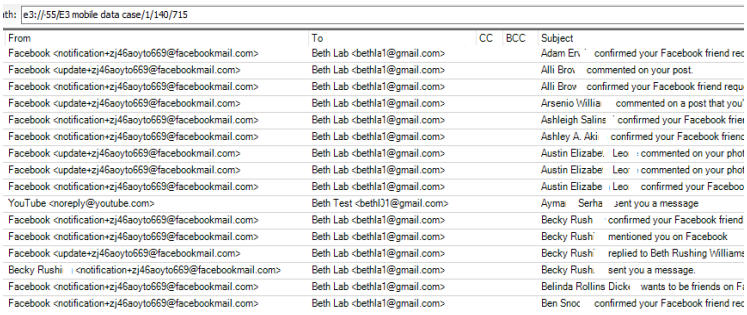

Figure 3. Facebook detailed interactions

Chat logs were aggregated from multiple accounts and services as shown in Figure 4. The chats were not explicitly saved to the device storage; instead they were captured from the cloud acquisition. Figure 5 shows that investigators were able to also uncover photos, pdfs, and other documents that were stored locally and in cloud accounts that had cached login credentials. Due to the cloud acquisition of the tested user account which yielded all conversations, user interactions, friend connections, etc., our findings for research question 4 indicate that Facebook does track and store all user interactions through their application.

Every action taken on the phone had been logged, including location and was collected in the forensic analysis of this device. The Note 5 smartphone had location services turned on for the first half of the data logging period and turned off for the second half. Regardless of whether the investigator was reviewing the enabled or disabled logging period, Google Timeline collected real-time location points throughout the entirety of the two-week process. Figure 6 illustrates how the Timeline information is parsed by the mobile forensic software.

Time	Sender	Text Preview
10/8/2017 11:33:08 PM	Beth Test <bethl 01@gmail.com>	Are you ever online?
10/8/2017 11:37:37 PM	Beth W <ew01920@gec.edu>	Hey. Sorry. I've been busy. What's up
10/8/2017 11:38:14 PM	Beth Test <bethl 01@gmail.com>	Nothing. Getting ready for bed.
10/8/2017 11:38:27 PM	Beth Test <bethl 01@gmail.com>	Do you work tomorrow?
10/8/2017 11:39:06 PM	Beth Wi <ew01920@gec.edu>	Yeah. Until 5 and then class
10/8/2017 11:39:15 PM	Beth W <ew01920@gec.edu>	I'm going to bed now too
10/8/2017 11:39:36 PM	Beth Test <bethl 01@gmail.com>	Ok. I'll message you tomorrow

Figure 4. All chat activities

Name	Type
beyonce-vmasperf2.jpg	JPEG image data
Getting started	PDF document
Hello.docx	Microsoft Office Open XML DOCX
October 12th.docx	Microsoft Office Open XML DOCX
Second.docx	Microsoft Office Open XML DOCX
The First.docx	Microsoft Office Open XML DOCX

Figure 5. Documents stored on device and cloud

Research question 3 asked if location data was recorded when services were disabled by the user. In this study, the researcher did find that location was still recorded, regardless of the user disabling location services. The investigators could see when and where the tracking started and ended, addresses, GPS coordinates, distance traveled in each recorded segment, and even a link to Google Maps. One observation made through this data logging and analysis period was the inability for Google Location to pinpoint accurate location points on military installations. The test device traveled to Fort Gordon (Home of the U.S. Army Cyber Center) and location services pinged a nearby business but did not ping any location within the military installation. Further research could be observed regarding location tracking and signal blocking as it relates to military operations.

From (Time)	To (Time)	Location Name	Location Address	Coordinates	Distance Travelled (m)	Google Maps URL
10/10/2017 13:23	10/10/2017 15:37	Work (Georgia University)	13 Son Dr, Kennedy, GA 30432	32.42054890000001, -81.78653469999999	0	http://maps.google.com/maps?q=32.42054890000001,-81.78
10/10/2017 15:37	10/10/2017 16:41	Driving		32.394824, -81.818542	8888	http://maps.google.com/maps?q=32.394824,-81.818542&am
10/10/2017 15:37	10/10/2017 16:41	Driving		32.4079049, -81.7642387	8888	http://maps.google.com/maps?q=32.4079049,-81.7642387&a
10/10/2017 16:41	10/10/2017 17:30	Technical College	1 Joe Blvd, Kennedy, GA 30432	32.394824, -81.818542	0	http://maps.google.com/maps?q=32.394824,-81.818542&am
10/10/2017 17:30	10/10/2017 18:51	Driving		32.358551399999996, -81.7543242	10202	http://maps.google.com/maps?q=32.358551399999996,-81.7
10/11/2017 12:53	10/11/2017 13:29	Driving		32.42643041765729, -81.78199398405258	1491	http://maps.google.com/maps?q=32.42643041765729,-81.78
10/11/2017 13:29	10/11/2017 17:27	AndHall	18 Son Dr, Kennedy, GA 30432	32.4264548, -81.7819916	0	http://maps.google.com/maps?q=32.4264548,-81.7819916&a

Figure 6. Google Timeline tracking regardless of location services being on/off

DISCUSSION

The results of this study uncovered hidden Google Timeline tracking location on an Android device, regardless of location services being turned on or off. This discovery prompted the question of why this service was logging location and what other

stored data could be pulled from the mobile device. The results also showed a more in-depth tracking of chat logs and every activity through Facebook and aggregated across multiple accounts and the ease of accessing cloud accounts. The forensics analysis was supposed to find the information that it found, but the applications were not supposed to track and store every piece of data that was recovered. The results of these findings have implications relevant to every "smart" device user in the world.

On December 11, 2018, Google's CEO, Sundar Pichai, testified at a Congressional hearing regarding Google's privacy policies amid concerns of consumer data collection. U.S. Congressman Bob Goodlatte provided opening statements with the comment, "Google is able to collect an amount of information about its users that would even make the NSA blush. Americans have no idea the sheer volume of information that is collected" (C-SPAN, 2018).

On Tuesday, September 21, 2017, an article was released on Quartz that revealed Google's use of Android location services to collect Android user locations, even when the location services option was disabled (Collins, 2017). The findings of this report were identical to the results yielded through the Google Location experimentation completed through this research. Following the announcement of the findings by Quartz, Google changed their user privacy policy to list "Your Location" as data they actively collect. Google also indicated that they track location services to more accurately tailor advertisements based on the user's location history. By Wednesday, November 22, 2017, Google responded to Quartz and addressed the allegations by stating that they would end location tracking of users with disabled location services by the end of November. There was no comment made about ending user tracking and reporting for those with location services turned on.

On August 13, 2018, the Associated Press released the results of their own investigation into Google's location tracking through a series of tests performed with location history turned off (Nakashima, 2018). Their results were identical to that of this research which indicated continued location tracking even after Google's 2017 announcement to cease location tracking on users with location settings disabled. Since the location tracking exposure in 2017, Google has restricted the controls of the easily accessible location history setting and redistributed micro-location attributes through hidden settings across numerous operating system level menus (Rash, 2018). The change from one central location setting to numerous nested location settings has allowed for Google's continued location tracking without users being made aware of this variation in policy (Google Privacy, 2018). This action has prompted concern over Google's compliance with the European General Data Protection Regulation (GDPR) (GDPR, 2018). Congressman Ted Poe, from the U.S. House Judiciary Committee, was frustrated with Google's C.E.O.'s lack of ability or transparency to answer if Google is tracking when his phone moves from one place to another. Sundar Pichai replied "not by default, but there may be a Google service which you've opted-in to use" (Wakabayashi & Kang, 2018). Since May 25, 2018 the GDPR is in force. The act requires that consumers have the right to access and right to be forgotten. If Google's location tracking is found to breach GDPR requirements, Google can be fined up to 4% of their annual global yield (Bridge, 2018).

LIMITATIONS

The study was conducted with a single Android Galaxy Note device with Google cloud application usage, picture/document storage, chat messaging. The device was rooted prior to being acquired and analyzed. Newer devices with enabled security protections may not yield the physical acquisition results that the researchers were able to obtain in this study.

CONCLUSION

At the time of examining this phone, Google and Facebook were not transparent regarding what data is tracked and how it is obtained. The findings of this study give users more information about what the device in their pocket is doing regardless of the settings that they believe they control. While data recording could benefit researchers and law enforcement, there is debate over what data should be collected by private entities. Google has made numerous updates to their privacy policies but only in response to consumers such as Quartz identifying location tracking inconsistencies and making these revelations public. These changes both in policy and software have made it more difficult for end users to understand and control what data is being collected. The most recent update to Google Timeline has made the application easily viewable and able to be corrected or deleted by the end user, while other parts that appear as if they were never tracked, are actually stored and accessible by professionals with expertise to perform forensic analysis on electronic devices. The implications regarding data storage by private companies without oversight is a serious concern for consumers that utilize these platforms. Regulations such as the GDPR assist in keeping companies honest about their data collection methods and how this stored data is manipulated in an effort to protect the privacy and security of consumers. As this research identified data privacy and security concerns related to application data storage, future iterations may focus on micro-location settings through Android OS, educating users of natively stored data, and how to ensure privacy in a technological age.

ACKNOWLEDGMENTS

This research began at Georgia Southern University in a graduate course, then continued at Middle Georgia State University.

REFERENCES

1. Ayers, R., Brothers, S., Jansen, J. (2014) *Guidelines on mobile device forensics.* (NIST Special Publication 800-101, 2014 Edition). Retrieved from https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-101r1.pdf
2. Bridge, M. (2018) Google at risk of £3bn fine over location tracking, *The Times.* Retrieved from https://www.thetimes.co.uk/article/google-at-risk-of-3bn-fine-over-location-tracking-tvq9fc97k
3. C-SPAN. (2018) *Google CEO Sundar Pichai testifies over data privacy and bias concerns* [Video File]. Retrieved from https://www.c-span.org/video/?455607-1/google-ceo-sundar-pichai-testifies-data-privacy-bias-concerns
4. Collins, K. (2017, November 21) Google collects Android users' locations even when location services are disabled, *Quartz.* Retrieved from https://qz.com/1131515/google-collects-android-users-locations-even-when-location-services-are-disabled/
5. Facebook. (2018) Facebook data policy. Retrieved from https://www.facebook.com/policy.php
6. General Data Protection Regulation (GDPR)(2018) Retrieved from https://eugdpr.org/
7. Google. (2018) General Data Protection Regulation (GDPR) | Google Cloud. Retrieved from https://cloud.google.com/security/gdpr/
8. *Google privacy* (2018). Retrieved from https://safety.google/privacy/data/
9. Harris, M. A., Brookshire, R., Patten, K., & Regan, B. (2015) Mobile application installation influences: Have mobile device users become desensitized to excessive permission requests, In *Proceedings of the Twentieth Americas Conference on Information Systems (AMCIS 2015)*.
10. Khatibloo, F. (2018, Nov. 14) Ethics and consumer action will transform privacy, *Forbes*. Retrieved from https://www.forbes.com/sites/forrester/2018/11/14/ethics
11. Lessard, J., & Kessler, G. (2010) Android forensics: Simplifying cell phone examinations, *Small Scale Digital Device Forensics Journal*, 1-12.
12. Magnet Forensics (2017) *Android acquisition methods from root to recovery.* Retrieved from https://www.magnetforensics.com/white-papers/download-white-paper-android-acquisition-methods/
13. Nakashima, R. (2018) AP Exclusive: Google tracks your movements, like it or not, *Associated Press News.* Retrieved from https://apnews.com/828aefab64d4411bac257a07c1af0ecb
14. Paraben (2018) E3 Digital Forensic Software - Paraben Corporation. Retrieved from https://paraben.com/digital-forensic-software-e3/
15. Rash, W. (2018) Probe finds google tracks your location even when you tell it to stop, *eWeek.* Retrieved from http://www.eweek.com/cloud/probe-finds-google-tracks-your-location-even-when-you-tell-it-to-stop
16. Schilit, B., Hong, J., & Gruteser, M. (2003) Wireless location privacy protection, *Computer*, 36, 12, 135-137. doi: 10.1109/MC.2003.1250896
17. Stirparo, P., & Kounelis, I. (2012) The MobiLeak Project: Forensics methodology for mobile application privacy assessment, *2012 International Conference for Internet Technology and Secured Transactions*, 297-303.
18. Valentino-DeVries, J., Singer, N., Keller, M., & Krolick, A. (2018) Your Apps Know Where You Were Last Night, and They're Not Keeping It Secret, *The New York Times.* Retrieved from https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html
19. Vidas, T., Zhang, C., & Christin, N. (2011) Toward a general collection methodology for Android devices, *Digital Investigation*, 8, 14-24.
20. Wakabayashi, D., Kang, C. (2018, December 11) Sundar Pichai, Google's C.E.O., testifies on Capitol Hill, *The New York Times.* Retrieved from https://www.nytimes.com/2018/12/11/business/sundar-pichai-google-house-hearing.html
21. Yerby, J., Hollifield, S., Kwak, M., & Floyd, K. (2014) Development of serious games for teaching digital forensics. *Issues in Information Systems*, 15, 2.
22. Yerby, J. (2013) Legal and ethical issues of employee monitoring. *Online Journal of Applied Knowledge Management*, 1, 2, 44-55.