2019학년도 제2학기

EMB000057404de0

연구논문/작품 중간보고서

|  |  |  |  |
| --- | --- | --- | --- |
| **제목**  **Title** | |  | | --- | | Application of Digital Forensic Methodologies in Epidemiological contact tracing | | ○ 논문(∨) 작품( ) |
| **GitHub URL** | |  | | --- | | https://github.com/dlsgk147/2013311425\_GraduateEssay | | |
| **평가등급**  **Grade** | **지도교수 수정보완 사항**  **Revision & Supplementation** | **팀원 명단**  **Team Member** |
| A,B,F중 택1  (지도교수가 부여) | ○  ○  ○ | 윤인하 (인)  (학번: 2013311425) |

2019년 9월 19일

지도교수 : Arthur Tang 서명

Abstract

In epidemiological situation, finding out epicenter and preventing diseases from spreading out is important. That is, contact tracing of infected people is necessary to search their whereabouts and contact. In crime scene, digital forensics is widely used to chase the suspect. Likewise, we can apply digital forensics in epidemiological contact tracing. In this paper, I’ll focus on making exhaustive list of all possible digital entities. Entities is selected in two perspectives; who the suspect contacted in the last 14 days and where the suspect went to in the last 14 days. Classification of entities and Determination of ranking will be done in final report.

Preface

a) Background & Necessity

Epidemics becomes more severe and makes serious casualty. In 2015, MERS caused 186 infected people and 38 death tolls in South Korea. It is second highest mortality in the world. Eradication took long time because secondary, tertiary infection occurs and caused propagation of disease. It ventilated the importance of tracing start point and blocking the contagion. Not only for the temporary epidemics, but also the ordinary ones like tuberculosis or influenza are also important in national prevention of epidemics.

Table 1. Statistics about infection occurrence of each epidemic, KOSIS [1]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Diseases | Year | | | |
| 2018 | 2017 | 2016 | 2015 |
| Hepatitis A | 2,437 | 4,419 | 4,679 | 1,804 |
| Varicella | 96,467 | 80,092 | 54,060 | 46,330 |
| Tuberculosis | 26,433 | 28,161 | 30,892 | 32,181 |
| AIDS | 989 | 1,008 | 1,060 | 1,018 |

These kinds of tracing are called contact tracing. It may accompany with analog method such as legwork or investigating real scene. However, in recent society, most people exposed to digital devices and they have lots of information about one person. Thus, it is important to deal with these sources and find out needed information.

Digital forensics are overall methodologies to find out data from digital entities. Entities can be any kind of sources; tangible or intangible, public or private, portable or stationary, etc. Its history is just about 30 years, but now digital forensics is one of most important investigation methodologies because almost every people are exposed to various digital devices. In daily routine, people use their cell phone, computer, credit card, etc. Except some rural area, CCTV is installed in every road and alley. Cars usually equips a black box. These kinds of digital devices contain lots of information. Naturally methods for investigating those data is studied and developed. It is now established with formal methodologies called digital forensics and numerous search algorithm or programs are developed to each specific source.

Table 2. Example of Digital Forensics Tools [1][2]

|  |  |
| --- | --- |
| Data Source | Forensics Tools |
| Cell Phone | XRY, TULP 2G |
| Desktop/Laptop | CrowdResponse, Volatility (RAM), Linux ‘dd’, HxD |
| Image | Exiftoll (EXIF metadata), |
| USB | DSi USB Write Blocker, USB Historian |
| File System | FTK Imager, Free Hex Editor Neo, |

It can be applied to various situation, but most commonly used part is crime scene investigation, especially to suspect tracking. People leave their trace more easily to various devices and it becomes important evidence to find out suspect, victim or witness’s whereabouts. These methods can be applied to epidemiological contact tracing also.

There are some differences between suspect tracing and epidemic case. For example, suspects try to hide the evidence, but people don’t matter of leaving their trace in epidemic situation. Thus, it is easier to find out their trace. In addition, epidemiological situations are mostly controlled in national level, therefore more entities are accessible.

Meanwhile, there are still problems for the application. First, experts or methodologies about digital forensics are still lacked in South Korea. Also, the research attaching digital forensics to epidemiological circumstances is not much compared to crime scene investigation. Because of that, it is needed to make guideline that applies to various problematic situations while using digital forensics methodologies to epidemics contact tracing.

b) Goals

Establishing whole methodologies from evidence identification to analysis, presentation is not easy procedure. Because of that, I’ll focus on evidence identification and collection. Among others, main goal will be making a checklist of all possible digital devices. The entities contain information and data about ⅰ) who this person contacted in the last 14 days, ⅱ) location this person went to in the last 14 days.

Table 3. Example list of digital entities

|  |  |
| --- | --- |
| 1. Who this person contacted in the last 14 days | 2. Location this person went to in the last 14 days. |
| Call Records | GPS |
| SNS | Black Box |
| E-Mail | CCTV |
| Messenger | Wi-Fi Access Record |
| Image Content | Image EXIF Metadata |
| … | … |

One source can contain sort of data in it. For example, cell phone is most important entity and holds several kinds of information in it. Some of them like GPS or EXIF in image will be related to whereabouts while image content or call record to contacted people. In list, each data is documented separately.

Table 4. Example of data in one source

|  |  |
| --- | --- |
| Data Source | Data Type |
| Cell Phone | Call Record |
| Wi-Fi Access |
| Bluetooth |
| Image |
| NFC |
| SNS Record |
| Image | Content |
| EXIF Metadata |

List contains entities available in domestic only because things in abroad is difficult to deal with. It will be exhaustive list that includes every possible evidence. If several sources have the same data type (i.e. Images in mobile phone, Images in digital camera), they will be treated with different one.

Next, I’ll determine the ranking according to 2 standards; importance of evidence and level of difficulty to collect. Difficulty will be determined with cost of tools (open vs. closed), legal issues, quantity and several conditions. Consequentially, it will be arranged into a checklist. Researchers check the entity is available or not, and the quantity. This part will be mainly discussed in final report.

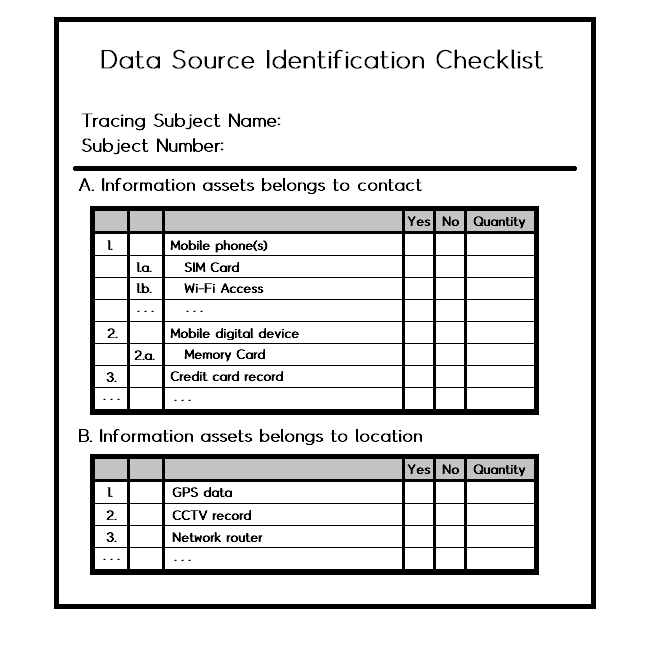


Figure 1. Example of data source checklist

c) Overview

In overall digital forensic models introduced in [4], this research can be helped from preparation to examination phase. It is difficult to suggest whole standards of digital forensics, so the direction will be small part of it, especially with evidence collection. Unlike physical evidence, digital entities include lots of intangible, invisible evidence, so it is likely to be missed than physical one. That’s why the checklist must include as many entities as possible. It also must be detailed rather than rough subscription. As I described in upper paragraph, some digital device like mobile phone can have several digital entities. Final output will be similar subscribed in goal paragraph, the exhaustive list of digital entities. Nonetheless, there are some missing sources due to lack of research or omission. Thus, the list will be used as an abstract and it will be revised, added or subtracted continuously in real scene.

In investigation, researchers start with first subject, broaden risk group according to subject’s contact and whereabout, and finally prevent further infection or find out the root of epidemics.

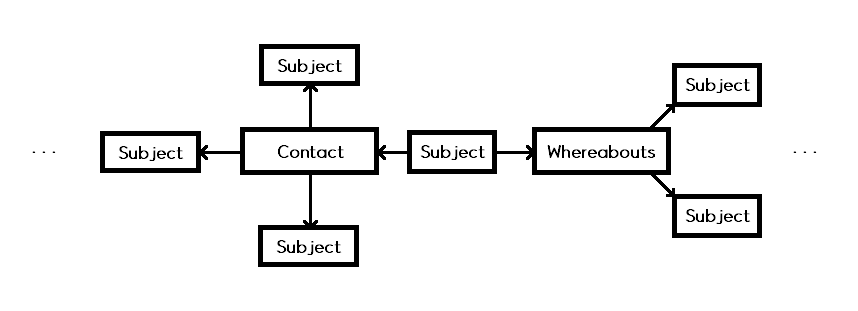


Figure 2. Subject broadening process

Related Research

[5] IACP introduces common electronic devices that generate digital evidence. There are personal digital devices like cell phone or laptop, and regional devices like CCTV. It doesn’t show further details about the types of evidence, but it can be used as first abstract of our goal.

Table 5. List of digital devices

|  |  |
| --- | --- |
| Device | Types of Evidence |
| Cell Phone | Call Records  Messenger  Applications |
| Desktop/Laptop | Internet Search History  Documents |
| CCTV | Videos |
| Drones | Videos |
| … | … |

Most important digital devices may be smartphone. Penetration rate of smartphone in South Korea already exceeds 90 percent (Gallup, 2017) and it is most portable among digital devices. Besides, smartphone becomes more versatile and contains more sensors, storages and applications collecting and storing data inside. That is, smartphone is most likely storing important information and data in it.

[6] In Willassen’s paper (2005), it introduces methods for imaging the internal memory of mobile phones including deleted memory. Usual phones use GSM system and it mandates SIM as a memory. In typical methods that using AT commands or analyzing live memory, there is some problems that deleted memory cannot be restored or memory can be overwritten while a phone is on. Thus, it is safe to analyze the phone in ‘dead’ status. Not only for smartphone, most memory devices tend to be analyzed with off status. In epidemic case, the probability of intentional memory corruption will not be occurred, but still the possibility of by accident exists. Thus, storage analysis may accompany with dead forensic and differences between live and dead forensic must be in calculation.

[7] Methods for collecting smartphone sensor data is introduced in Mylonas et al. (2013). Smartphone includes multimedia hardware and various sensors like location, but the most sensor data is not available because it’s volatile data. The commonly possible ways are collecting data as soon as the signal is occurred by the subject or from service provider. In our research, we focus on finding out the whereabouts of last 14 days so ad hoc data may not available, but system provider and hardware log file are still available.

[8] Agarwal et al. (2012) introduces Smartphone forensic investigation process model (SPFIPM). It claims the difference of computer forensics vs. mobile phone forensics and suggests new process model for the smartphone. Comparing to computer, mobile phone constantly alters its memory, so it is harder to get whole copy of the memory. Phone has more variable ways to connect to online than computer. That is, whether the phone is on or off is an issue in the process.

Major difference between traditional digital forensic model descripted in Reith’s paper (2002) [4] and communication shielding according to the PDA mode and subdivision of evidence collection. The important thing is remaining PDA mode as it is and preventing volatile memory from disappearing.

[9] Kesslar et al. (2010) subscribes Android forensic procedure. It shows several methods to analyze Android smartphone that acquiring dd image using FTK Imager, logical analysis and data extraction with the CelleBrite UFED. From the image, various digital data like call history, MMS/SMS message, web search history, e-mail, GPS data and video/photo can be found. Each of them can be considered as available digital entities in smartphone.

[10] Jeong et al. (2011) discusses about digital forensics of IaaS cloud computing service. In cloud computing, physical data storages places here and there, so it causes legal problem in collection. Also, in the public cloud service like AWS, numerous users share the service, memory or network thus it makes collection difficult further. This paper suggests its own forensic process for cloud service and find out evidence.



Figure 3. Cloud Service Digital Forensic Process [10]

Available digital evidence is virtual instance, network layer data and client system data. Instance can provide data location and drive position and client system works as an application in client site, so its data can be important evidence. Network layer provides communication data and protocol data, but it is better investigated with other network forensic process.

[11] Perumal et al. (2015) introduces IoT digital forensics investigation model. With IoT service, digital devices become more versatile. It makes each device can store various types of data. Therefore, like a smartphone, IoT devices become important in terms of digital forensics. Unlike typical computer or devices, IoT connects almost every device in it and composes large, complex system. Because of that, it is hard to apply typical forensic methods to IoT effectively. They divide entire system as 3 zones: internal network, middle (gateway, firewall, etc.) and external network (Internet, cloud, web services, etc.). With IoT digital forensic, investigators can get almost every kinds of digital evidence.

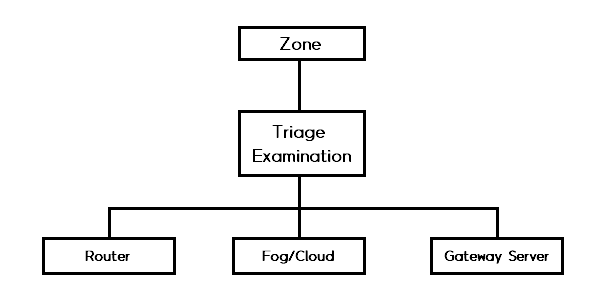


Figure 4. IoT based digital forensic model [11]

[12] Lee et al. (2008) suggests forensic analysis methodologies about embedded system. Each device has unique system according to their purpose and operation environment. Because of this diversity, it is nearly impossible to formulate general methods of embedded system forensic. Their methodology is not a whole process matching for every system, but provides comprehensive analysis techniques. It is composed with two phases: hardware and software analysis phase. With information provided by manufacturer, it checks specification, configuration, etc. about hardware/software. It emphasizes on revealing hidden or modified data, but it is less likely in our case, thus it is easier to find out required data with system analysis.

[13] Among IoT devices, Boztas et al. (2015) focuses on Smart TV forensics. Smart TV are usually IP based device, and can be connected to Internet, social network service and other digital devices. It introduces data acquisition method depending on the hardware type. Available data is similar to smartphone or computer.

Introduction

The topic is application of digital forensics for epidemiological contact tracing. I’m just a beginner in this area, so I decided to narrow down the goal. Real work is concentrated to the checklist for the digital forensics. Checklist consists of all kinds of digital evidence that can apply to the contact tracing. The type of evidence is not much different with the case of crime suspect tracing. Rather, the subject rarely tries eliminate or hide data compared to criminals. That is, each device can be investigated with typical methodologies without the concern about intentional concealment. It can widen the range of possible digital sources.

My research is made up of mainly two part. First is the complete list of all possible digital evidence. Evidence can be exposed from any kind of sources; thus the list must include digital sources as many as possible. Integrating all sources in one table is difficult and more likely to miss some entities. It is required to divide the entire list with several criteria according to their properties. It’ll be similar with categorization of devices. Representatively, portable vs. stationary, network device vs. stand-alone device or public vs. private can be categories. Not only their attributes, type of data (Image, Text, Video, etc.), which issues it is related to (location vs. contact vs. both of them) also can be the topic of list.

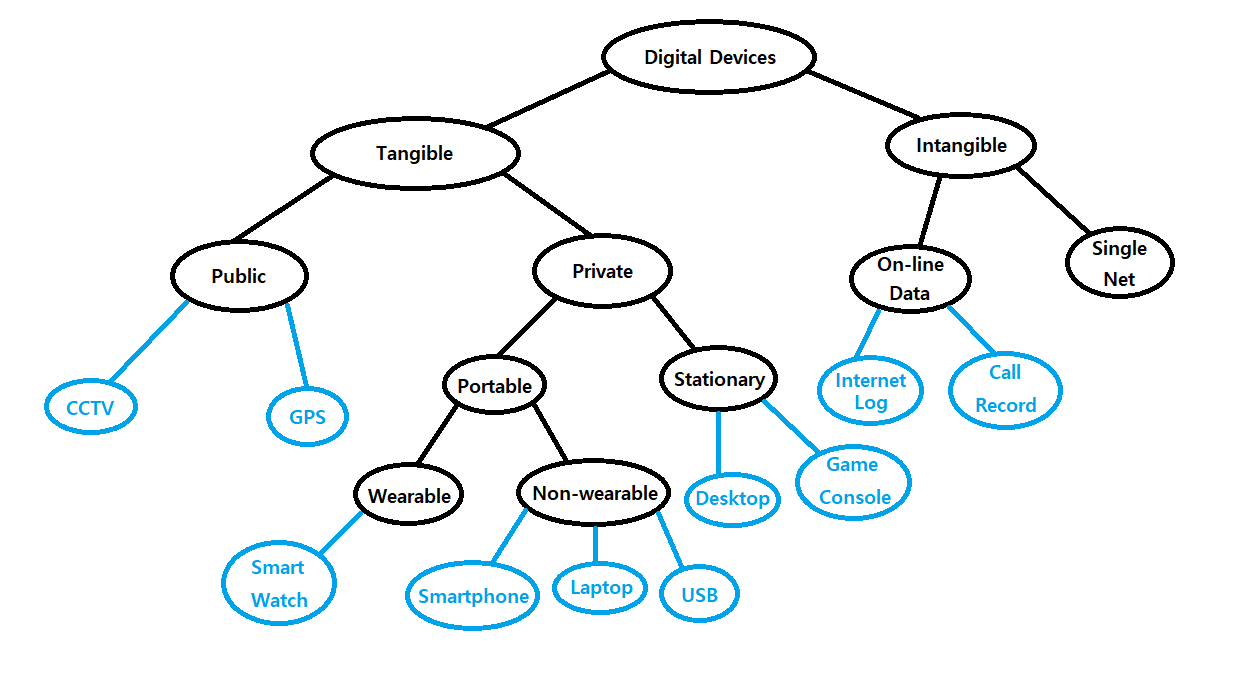


Figure 5. Classification of digital devices

Each device can be included in multiple lists. For example, smartphone is portable, personal and on-line device. Meanwhile, surveillance camera is fixed, public and local network device. Also, data in one source can be split with their purpose. For instance, in smartphone, GPS data or Wi-Fi access record are related to the place that this people went while SNS or call record to the people that this person met. List mainly consists of typically used devices because those are more important and accessible.

Second, entire list will be ranked by two criteria. First is the importance of the evidence. It can include several standards. For instance, its distribution rate seriously affects to the availability. The portability also can influence to the likelihood of the access. Second is the difficulty of acquirement. If this source is worthy and contains much data but it’s impossible to get it, it becomes useless. Analysis methods are still advancing, but still the problems like obsolete software/hardware or lack of research about new technologies make it harder to find out useful data from it. Also, contact tracing is related seriously to privacy invasion, so there are several lawful and moral issues. Each attribute can occur some issues while get rid of it. For example, public facilities can be researched more easily than personal one in terms of lawful issue. In the same private things, smartphone or desktop are more likely to have personally sensitive information than other devices and raise hostility to the investigation. Because of these problems, the researchers still endeavor to establish legal process of digital forensics. In our situation, some problem will be released in some level. In South Korea, there is law about disaster & safety management [14]. According to that, every investigation must be executed with executive order and every safety management facility must provide required data without specific reason. Civilians also must be cooperative to disaster response. That is, we can put legal problems as secondary and focus on technical issues to measure difficulty.

For each device, we can measure the approximate values about its importance and difficulty in the form of card. It contains the name, data type, evaluation, pros and cons, etc. Based on that, we can make ranking among sources.

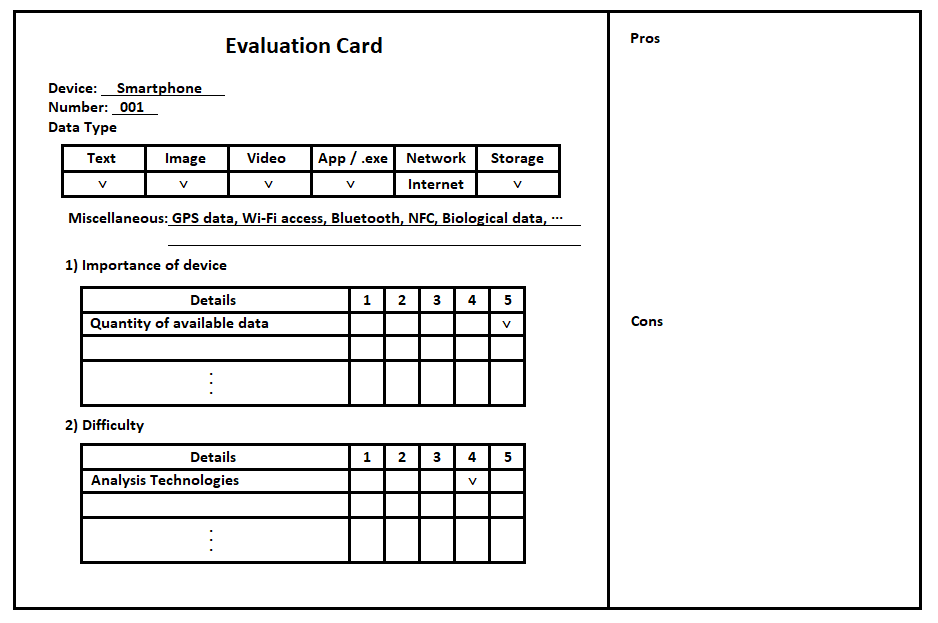


Figure 6. Example of evaluation card

Scoring is rough method and maybe hard to distinguish delicate difference among them, but it is more comfortable to measure its importance. Additional issues that cannot evaluated in previous items can be written in pros and cons section. It is temporary version and may be fixed.

Based on the evaluation, we can measure the rank of entities. So far from charge the rank one by one, it will be the sort of groups like most primary, secondary, ternary, etc. For example, most important devices like smartphone or desktop will be primary group, while intermediate ones like drone, smart watch, etc. will be secondary group. It may not be an exact priority, but it can be a guide to decide the order of investigation.

Implementation & Result Analysis

Table 6. List of Portable Items

|  |  |
| --- | --- |
| Device | Entities |
| Digital Camera | Images  Videos  Storage (SD Card) |
| Smartphone | Text Messages  Call Logs  Application  SNS accounts  GPS  Wi-Fi Access  Bluetooth  NFC |
| Credit Card | Credit Card Record |
| Smartwatch | Similar to smartphone |
| Laptop | Internet Search History  Storage (HDD, SSD)  SNS accounts |
| External Hard Disk | All kinds of digital files |
| Drone | Videos |
| USB Memory | All kinds of digital files |
| SD Card | All kinds of digital files |

Table 7. List of Stationary Items

|  |  |
| --- | --- |
| Device | Entities |
| Subject’s Home | |
| Desktop | Internet Search History  Storage (HDD, SSD) |
| Game Console | Images  Videos  Documents  Operation Log |
| Internet of Things [11] | Every type of digital file |
| Embedded System [12] | Every type of digital file |
| Smart TV [13] | Application  Web Browsing  Images and Multimedia Files  External Media (USB, Flash drive, etc.)  Cloud Service  E-mail |
| Subject’s Office | |
| Desktop | Internet Search History  Storage (HDD, SSD) |
| Network Switch / Router | Access Records |
| CCTV | Videos |
| RFID | Access Records |
| At Outside | |
| Surveillance Camera (CCTV) | Videos |
| HiPass | Through Time |
| Blackbox | Videos |
| ATM | Credit Card Record |
| Wi-Fi | Access Records |
| Network Switch / Router | Access Records |
| Airport Immigration Office | Immigration Records |

Table 8. List related to subject’s whereabouts

|  |  |
| --- | --- |
| Device | Entities |
| Smartphone | Application  SNS  GPS  Wi-Fi Access  Bluetooth  NFC |
| Desktop | Internet Search History  Storage (HDD, SSD) |
| Digital Camera | Images  Videos  Storage (SD Card) |
| CCTV | Videos |
| Network Switch / Router | Access Records |
| Blackbox | Videos |
| Smartwatch | GPS  Wi-Fi Access |
| Drone | Video |

Table 9. List related to subject’s contact

|  |  |
| --- | --- |
| Device | Entities |
| Smartphone | Text Messages  Call Logs  Application  SNS accounts |
| Desktop | Internet Search History  Storage (HDD, SSD) |
| CCTV | Videos |
| Blackbox | Videos |

Table 10. List divided into data type (Except storage device)

|  |  |
| --- | --- |
| Data type | Device |
| Text | Smartphone  Desktop / Laptop |
| Image | Smartphone  Digital Camera  Desktop / Laptop  Game Console |
| Video | Smartphone  Digital Camera  Desktop / Laptop  Drone  CCTV  Blackbox  Game Console |
| Sensor Data | Smartphone  RFID  GPS  Airport Immigration  Network Device  Wi-Fi |

It’s the brief list divided with several conditions. Priority is not calculated yet. These conditions are for finding out all stuffs without missing, so it can be added afterward. At now, the condition is portable vs. fixed, related topic. I think there can be a case that researchers want to find out some specific data type, so I added the list according to the data type it has. Storage devices (HDD, SSD, USB, etc.) can save almost every sort of data, so I excluded it. Additionally, I’ll make the list by whether it is private or public device.

Recently, more devices become versatile and have various function. Even it is the same source, available data will be different seriously. For example, unlike past feature phone, smartphone contains much more functionalities. Drones equips various sensor according to their purpose, so available data can be diverse. Also, some sources are newly appeared as data source. Television just has one functionality; Receiving electric wave and transmit the broadcast. But now, smart TV have some functions and applications same as smartphone. However, it also means that some entities not in the list are available actually while other in the list are not. It is general purpose checklist; thus, the entity will be composed of universal one.

Conclusion & Opinion

In interim report, I mainly focused on finding out all possible digital devices. In a month, I’ll summarize the importance and difficulties of each device and organize a ranking among them. The category and criteria of evaluation card are not standardized yet, thus it will be the beginning of the research. Additionally, I’ll try to add missed entities. I’m just a beginner for this topic and it’s hard to find out digital evidence in detailed. This list may not be available for real investigation with this research and must revised further with follow-up researches.

It was first time to write down the essay. I didn’t have any prior knowledge and interest, so it was hard to research and make up of complete report. My essay doesn’t include any programming but still the research was not an easy-going. While preparing employment and attending lectures, I couldn’t concentrate to the essay wholly and it makes entire process delayed seriously. I cannot convince it will be helpful to me because I couldn’t do the work properly and process roughly here and there. A series of work make me feel about the difficulty of making up of single essay and research and attitude for completing research. I’ll try to make the research at some level, but I wonder it can achieve valuable research result. If I have a chance to do some research after, I must have much time and concentrate only to it. The quantity of interim report is at least 20 pages, but it is too much for the undergraduate student that writes essay for the first time. I spent a lot of time to fill in the quantity, but I couldn’t prepare for real research. I wonder that the report must be in that much quantity. I think writing the report compactly and focusing on the implementation is better.

References

[1] http://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=DT\_117N\_A00406&vw\_cd=MT\_ZTITLE&list\_id=D19&seqNo=&lang\_mode=ko&language=kor&obj\_var\_id=&itm\_id=&conn\_path=MT\_ZTITLE

[2] https://techtalk.gfi.com/top-20-free-digital-forensic-investigation-tools-for-sysadmins/

[3] THING, Vrizlynn LL; NG, Kian-Yong; CHANG, Ee-Chien. Live memory forensics of mobile phones. *digital investigation*, 2010, 7: S74-S82.

[4] REITH, Mark; CARR, Clint; GUNSCH, Gregg. An examination of digital forensic models. *International Journal of Digital Evidence*, 2002, 1.3: 1-12.

[5] Common electronic devices that generate digital evidence, https://www.iacpcybercenter.org/officers/cyber-crime-investigations/common-electronic-devices-that-generate-digital-evidence/

[6] WILLASSEN, Svein. Forensic analysis of mobile phone internal memory. In: *IFIP International Conference on Digital Forensics*. Springer, Boston, MA, 2005. p. 191-204.

[7] MYLONAS, Alexios, et al. Smartphone sensor data as digital evidence. *Computers & Security*, 2013, 38: 51-75.

[8] GOEL, Archit; TYAGI, Anurag; AGARWAL, Ankit. Smartphone forensic investigation process model. *International Journal of Computer Science & Security (IJCSS)*, 2012, 6.5: 322-341.

[9] LESSARD, Jeff; KESSLER, Gary. Android Forensics: Simplifying Cell Phone Examinations. 2010.

[10] 정일훈, et al. IaaS 유형의 클라우드 컴퓨팅 서비스에 대한 디지털 포렌식 연구. *정보보호학회논문지*, 2011, 21.6: 55-65.

[11] PERUMAL, Sundresan; NORWAWI, Norita Md; RAMAN, Valliappan. Internet of Things (IoT) digital forensic investigation model: Top-down forensic approach methodology. In: *2015 Fifth International Conference on Digital Information Processing and Communications (ICDIPC)*. IEEE, 2015. p. 19-23.

[12] LIM, Kyung-Soo; LEE, Sangjin. A methodology for forensic analysis of embedded systems. In: *2008 Second International Conference on Future Generation Communication and Networking*. IEEE, 2008. p. 283-286.

[13] BOZTAS, Abdul; RIETHOVEN, A. R. J.; ROELOFFS, Mark. Smart TV forensics: Digital traces on televisions. *Digital Investigation*, 2015, 12: S72-S80.

[14] 재난 및 안전관리 기본법, <http://www.law.go.kr/>법령/재난및안전관리기본법