Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.27, NO.6, Dec. 2017 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2017.27.6.1397

파티션 복구 도구 검증용 데이터 세트 개발 및 도구 평가

박 송 이,† 허 지 민, 이 상 진‡
고려대학교 정보보호대학원

Development of a Set of Data for Verifying Partition Recovery Tool and Evaluation of Recovery Tool

Songyee Park,† Gimin Hur, Sang-jin Lee‡
Center for Information Security Technologies(CIST), Korea University

요 약

손상된 저장매체에 대해서 디지털포렌식 조사를 진행할 때 복구 도구를 활용한다. 하지만 사용하는 복구 도구에 따라서 복구 결과가 다른 문제가 존재한다. 그러므로 정확한 조사를 위해서는 도구의 성능과 한계점을 파악하여 사용할 필요가 있다. 본 논문에서는 이러한 파티션 복구 도구의 성능을 검증할 수 있도록 MBR, GPT 디스크 인식 방식과 FAT32, NTFS 파일시스템의 구조적 특징을 고려한 검증 시나리오를 제시한다. 그 후 검증 시나리오를 바탕으로 제작한 데이터 세트를 통하여 기존 복구 도구에 대한 성능 검증을 진행한다.

ABSTRACT

When a digital forensic investigation is conducted on a damaged storage medium, recovery is performed using a recovery tool. But the result of each recovery tool is different depending on the tools. Therefore, it is necessary to identify and use the performance and limitations of the tool for accurate investigation. In this paper, we propose a scenario considering the disk recognition type such as MBR, GPT and the structural characteristics of FAT32 and NTFS filesystem to verify the performance of the partition recovery tool. And then We validate the existing tools with the data set built on the scenarios.
Keywords: Digital forensics, Data set, Digital forensics tool testing, Recovery tool

I. 서 론

조사 대상 저장매체가 손상되면 내부 데이터에 대해 접근하는 것이 어렵다. 그래서 디지털포렌식 조사 과정에는 손상된 저장매체를 정상적으로 인식시키기 위해서 복구 과정을 거친다. 그러나 복구 도구에 따라서 동일한 저장매체에 대해서 복구를 진행해도 복원되는 결과가 다르다. 이러한 문제는 각 도구에서 복구를 위해 참조하는 정보가 다르거나 도구의 오류로 인해서 나타난다. 따라서 조사관은 자신이 사용하는 도구들에 대하여 기능적 한계점과 오류의 존재 여부를 객관적인 기준으로 검증함으로써 보다 정확한 조사를 할 수 있다.

도구의 검증을 위해서는 검증용 시나리오와 시나리오에 따라 제작된 데이터 세트가 필요하다. 본 논문에서는 국내에서 주로 사용하는 윈도우 운영체제를 기준으로 디스크 인식 방식과 FAT32와 NTFS 파일 시스템을 고려한 복구 도구 검증용 데이터 세트를 제작한다. 이를 통해서 현재 사용되는 파티션 복구 도구에 대해서 검증을 진행한다.

Received(09. 26. 2017), Modified(10. 31. 2017), Accepted(11. 01. 2017)
† 주저자, songhasong@korea.ac.kr
‡ 교신저자, sangjin@korea.ac.kr(Corresponding author)

2장에서는 미국에서 디지털포렌식 도구 검증을 위해 공개하고 있는 데이터 세트에 대해서 알아본다. 다음 3장에서는 데이터 세트 제작의 기반이 되는 파티션 손상에 대한 시나리오에 대해서 제시한다. 4장에서는 제작한 데이터 세트를 기반으로 실제 도구에 대한 테스트를 진행한다. 5장에서 결론 및 향후 연구 계획을 소개한다.

II. 관련 연구

미국 상무성(Department of Commerce) 소속의 정부 기관인 표준기술연구소(National Institute of Standards and Technology, NIST)의 CFTT(Computer Forensic Tool Test) 프로젝트에서는 디지털포렌식 조사 시 사용하는 도구에 대한 검증 방법론을 수립하고자 도구 기능 명세, 검증 절차, 검증 기준, 검증 데이터 세트 개발 등을 하고 있다[1]. CFTT 프로젝트에서는 디스크 이미징, 파일 복구, 파일 카빙, 스트링 서치 등 총 8가지에 대해서 진행되었다. 도구에 대한 평가 결과 보고서는 미 국토안보부 홈페이지를 통해 공개하고 있으며 디지털포렌식 도구의 신뢰성을 검증하기 위한 자료로 활용되고 있다.

NIST에서는 CFTT 프로젝트를 통해서 개발한 검증 데이터 세트를 웹 사이트에 공개하고 있으며 이 데이터 세트를 CFReDS(Computer Forensic Reference Data Sets)라고 한다. CFReDS는 도구를 테스트할 수 있는 이미지와 디지털포렌식 조사관이 증거를 분석하는 방법에 대해서 학습할 수 있는 데이터 세트 등을 제공한다.

DFTTI(Digital Forensics Tool Testing Images)는 민간 주도로 디지털포렌식 도구를 검증하는 테스트 세트를 제작하는 프로젝트로 2005년부터 2010년까지 진행되었다[2]. DFTTI에서 제공하는 테스트 이미지 세트는 총 10개로 DD 파일 형태로 배포하고 있다. 제공하는 테스트 세트로는 확장 파티션 테스트, 파일시스템 별 키워드 서치 기능 및 삭제한 파일 및 폴더에 대해서 복구 가능성을 테스트할 수 있는 테스트 세트를 제공하고 있다. 그러나 CTFF, CFReDS, DFTTI에서는 디지털포렌식 도구의 모든 범주를 모두 다루고 있지 않다[4]. 저장매체 복구 도구의 경우 디지털포렌식 조사 과정에서 활용도가 높은 도구이나 도구 검증을 위한 검증 시나리오와 데이터 세트가 개발되어 있지 않다.

III. 파티션 복구 도구 검증 시나리오 개발

본 장에서는 저장 매체가 논리적으로 손상이 되었을 경우를 가정하여 도구가 저장매체를 적절하게 복구할 수 있는지를 검증할 수 있는 시나리오 제작 방법에 대해서 다룬다.

3.1 파티션 복구 도구

파티션 복구란 논리적으로 손상된 저장매체에 대해서 디스크 인식 방식과 파일 시스템 구조를 고려하여 손상된 부분을 재구성하는 것을 의미한다. 여기서 말하는 저장매체의 논리적 손상이란 시스템 충돌, 악성코드에 의한 공격, 범죄자에 의한 임의 삭제 등과 같은 이유로 디스크 인식 부분과 파일시스템에 문제가 생긴 경우를 의미한다.

디스크 인식 방식으로는 Master Boot Record(MBR) 방식과 GUID Partition Table(GPT) 방식이 존재한다. 두 가지 방식은 디스크의 내부 정보를 관리하는 방식으로 구조적 차이를 보인다. 따라서 복구 시에는 이러한 구조적 특징을 고려해야 정상적으로 복구가 가능하다. 또한 복구 시에는 파티션이 어떠한 파일시스템으로 포맷이 되어 있었는지에 대해서도 파악하여 복구해야 한다. 따라서 본 논문에서는 파티션 복구 도구의 신뢰성을 검증하기 위해서 두 가지의 디스크 인식 방식과 FAT32, NTFS 파일시스템의 구조별 특징을 고려하여 손상된 저장매체에 대한 시나리오와 데이터 세트를 제작한다.

3.2 디스크 방식에 따른 손상 시나리오

3.2.1 MBR 방식

MBR 영역을 복구하기 위해서는 MBR에 저장되는 파티션 정보와 크기 정보 등을 현재 디스크 상태에 맞게 복원해야 한다.

MBR은 Master Boot Code와 Partition Table로 구성된다. Master Boot Code에는 디스크의 부팅 관련 정보가 저장되어 있다. Partition Table에는 현재 디스크에 존재하는 파티션의 Boot Record(BR)에 대한 위치 정보가 저장되며 총 4개의 BR 정보를 기록할 수 있다. BR은 파티션의 시작 영역을 의미하며 파일시스템에 따라서 Boot Sector,

Super Block과 같은 이름으로도 불린다. Extend Boot Record(EBR)는 MBR 디스크 방식에서 4개 이상의 파티션으로 구성하는 경우에 생성된다. EBR의 구조는 MBR과 동일하지만 Boot Code가 존재하지 않으며 2개의 Partition Table을 사용한다. MBR 디스크의 전체적인 구조는 Fig.1과 같다.

MBR 디스크에 대한 손상 시나리오는 5개의 파티션으로 구성된 EBR 디스크를 대상으로 제작했다. 개의 파티션으로 구성함으로써 MBR 디스크 방식에서 EBR을 생성하여 현재 EBR과 다음 EBR을 연결하는 EBR Table을 생성할 수 있다.

Master Boot Code 및 BR 손상 시나리오는 파티션의 시작 위치만을 알고 있을 때 파티션 복구 도구가 Master Boot Code와 파일시스템의 기본 정보에 대한 복구가 가능한 지를 판단할 수 있다. 시나리오 구성은 Fig.2.와 같다.

파티션의 시작 위치를 알 수 없을 때 파티션 복구 도구가 이미지를 전체 스캔하여 각 파티션의 시작 위치에 대해서 확인하는지를 판단할 수 있다. 시나리오 구성은 Fig.3.과 같다.

EBR 내부의 Current Entry와 Boot Record 손상 시나리오는 MBR과 구조는 동일하기 때문에 Master Boot Code를 손상시키지 않고 파티션 시작 위치와 파티션의 기본 정보가 모두 손상되었을 경우의 복구 여부를 확인할 수 있다. 시나리오 구성은 Fig.4.와 같다.

EBR Next Entry 손상 시나리오는 MBR 내부에서 EBR을 가리키는 4번째 Partition Table Entry와 EBR 내부의 Next Entry가 모두 손상되었을 경우 복구 가능한지를 판단할 수 있다. 시나리오 구성은 Fig.5.와 같다. MBR 디스크의 손상 시나리오를 정리하면 Table.1.과 같다.

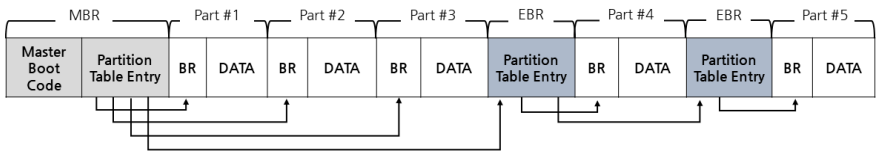


Fig. 1. MBR Structure

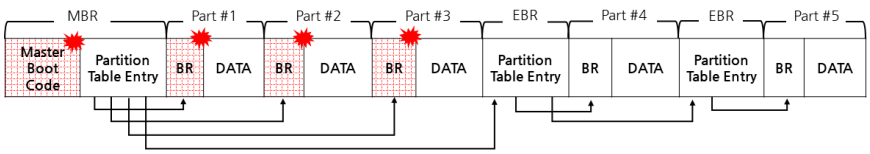


Fig. 2. Master Boot Code and BR within MBR

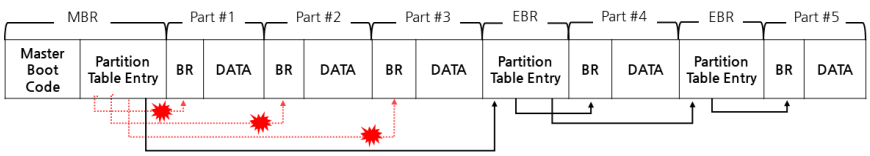


Fig. 3. Partition Table within MBR

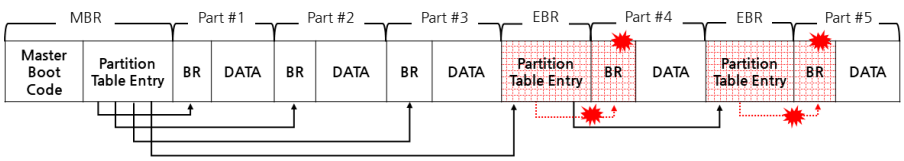


Fig. 4. Current Entry and BR within EBR

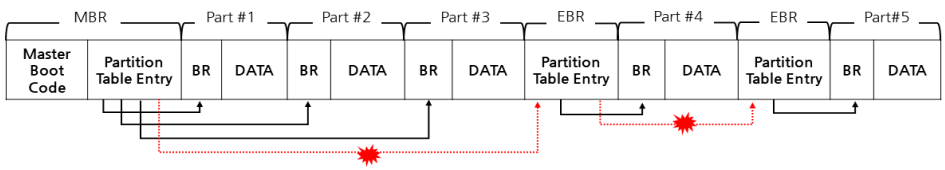


Fig. 5. EBR Next Entry

Table 1. MBR damage scenarios

Damage Scenarios	Detail
Master Boot Code and BR within MBR	Delete Master Boot Code and three Main BRs within MBR
Partition Table within MBR	Delete three BR Entries within MBR Partition Table
Current Entry and BR within EBR	Delete Current Entry and Extended BR within EBR
EBR Next Entry	Delete EBR Entry and EBR Table within MBR Partition Table

3.2.2 GPT 방식

GPT 디스크는 Protective MBR, 2개의 GPT Header와 2개의 Entries Table로 구성되어 있다. Protective MBR은 GPT 디스크의 0번째 섹터에 존재하며 컴퓨터에서 EFI(Extensible Firmware Interface)를 지원하지 않을 경우 GPT 디스크 방식으로 포맷되어 있다는 것을 알려주는 영역이다. 이 영역에는 GPT 파티션 영역 전체의 시작 주소와 끝 주소를 파티션 엔트리 0 번째에 저장한다. GPT Header는 주 GPT Header와 백업용 GPT Header로 총 두 가지 존재하며 GPT 디스크의 전반적인 설정 정보를 기록한다[10].

Entries Table는 128개의 Entry로 구성되며 각각의 Entry는 파티션 정보를 저장하고 있다. 따라서 GPT 방식으로 128개의 파티션을 표현할 수 있다. GPT Header와 마찬가지로 Entries Table에 대한 백업이 존재한다.

Primary GPT Header 손상 시나리오는 백업본인 Secondary GPT Header를 통해서 손상된 Header 영역을 복구할 수 있는지 확인 할 수 있다. 그리고 Primary GPT Header /Secondary GPT Header 손상 시나리오는 GPT Header에 대한 모든 정보가 손상되었을 경우에 Primary Entries Table 또는 저장매체를 전체 스캔하여 Header 영역을 재구성할 수 있는지를 판단할 수 있다. Primary Entries Table 손상 시나리오는 백업본인 Secondary Entries Table을 통해 복구 가능한지 여부를 알 수 있다. 그리고 Primary Entries Table/Secondary Entries Table 손상 시나리오를 통해서는 파티션의 정보가 모두 손상되었을 경우 저장 매체 전체를 스캔하여 위치 정보를 재구성 할 수 있는지에 대해 판단할 수 있다.

위의 GPT 디스크의 손상 시나리오를 정리하면 Table.2.와 같다.

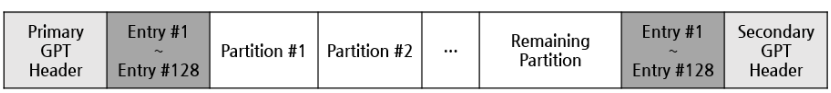


Fig. 6. GPT Structure

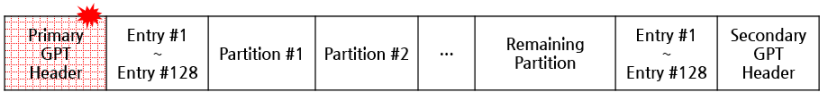


Fig. 7. Primary GPT Header

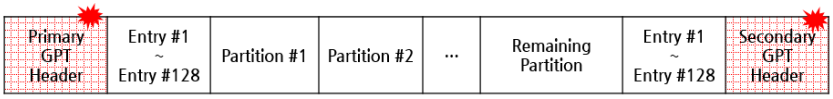


Fig. 8, Primary GPT Header /Secondary GPT Header

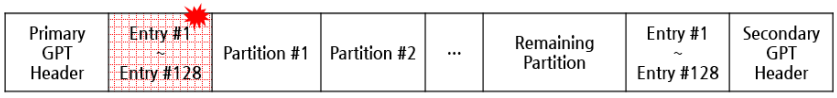


Fig. 9. Primary Entries Table

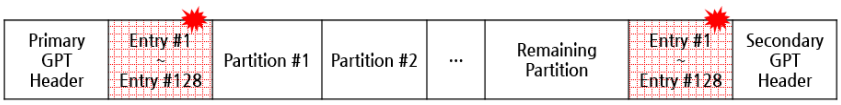


Fig. 10. Secondary Entries Table

Table 2. GPT damage scenarios

Damage Scenarios	Detail
Primary GPT Header	Delete Main GPT Header
Primary GPT Header /Secondary GPT Header	Delete Main GPT Header and Backup GPT Header
Primary Entries Table	Delete Main Entries
Primary Entries Table/Secondary Entries Table	Delete Main Entries and Backup Entries

3.3 파일시스템에 따른 손상 시나리오

파일시스템에 따른 손상 시나리오를 구성할 때에는 파일시스템 내에서 부팅과 관련된 정보를 저장하는 곳이 있는 지 파악하는 것이 중요하며 이를 손상시킨 후 도구에서 복구 기능을 제공하는지 확인가능하다.

FAT32의 구조에 대한 정보는 Reserved Area 영역에 기록된다. Reserved Area는 FAT32 파일 시스템에서 가장 앞쪽에 위치하며 내부의 BR(Boot Record)에는 운영체제 부팅을 위한 부트 코드와 파일시스템에 대한 정보를 저장하고 있다. Reserved Area의 구조는 Fig. 11과 같다.

BR영역은 Reserved Area 내부의 Boot Sector와 동일하다. 파티션의 0 번째 섹터가 주 Boot Sector이며 6 번째 섹터는 백업된 Boot Sector이다. 따라서 FAT32 파일시스템 안에 존재하는 모든 Boot Sector를 지움으로써 파티션 복구 도구가 백업된 Boot Sector 영역을 통해서 유실된 Boot Sector를 재구성하는 지를 확인할 수 있다.

NTFS 파일시스템으로 포맷하여 파티션이 최초에 생성되는 경우 파티션 정보가 Volume Boot Record(VBR)에 위치하게 된다. VBR은 파티션의 첫 번째 섹터에 위치하며 FAT 파일 시스템의 Reserved Area와 비슷한 형태이다. VBR은 Boot Sector와 Boot Code로 구성된다. NTFS로 포맷한 경우 Fig.12.와 같다.

VBR의 Boot Sector 백업본은 파티션의 가장 마지막 섹터 부분에 생성된다. 따라서 NTFS 파일

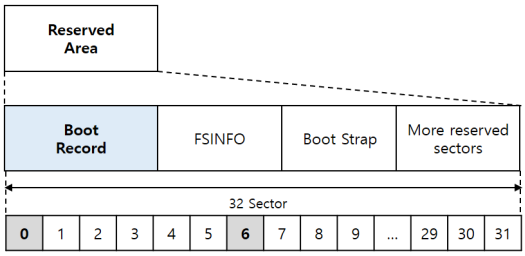


Fig. 11. Reserved Area in FAT32 Filesystem

Fig. 12. VBR of the NTFS Filesystem

시스템 안에 존재하는 모든 Boot Sector를 지움으로써 파티션 복구 도구가 백업된 Boot Sector를 사용해서 복구하는지 Boot Sector를 재구성하는지 구별 할 수 있다.

FAT과 NTFS 파일시스템에 따른 손상 시나리오는 Table.3.과 같다.

Table 3. Filesystem damage scenarios

Damage Scenarios	Detail
〔FAT Filesystem〕 BootSector	Delete all Boot Sector
〔NTFS Filesystem〕 Boot Sector	Delete all Boot Sector

IV. 데이터 세트 개발 및 검증 결과

4.1 검증용 데이터 세트

데이터 세트 개발은 하드디스크의 크기를 PC3000 도구로 8 GB로 축소하여 진행했다. 실제 하드디스크를 사용하여 데이터 세트를 제작함으로써 디스크 인식 영역의 손상에 대한 데이트 세트를 제작할 수 있도록 구성하였다.

데이터 세트는 각각의 디스크 인식 방식을 설정한 다음 5개의 파티션으로 나눴다. 모든 파티션은 동일한 파일시스템으로 포맷하였다. 각각의 파티션에는 10개의 서로 다른 확장자의 파일을 넣었다. 정상적인 영역에 'DUMMY'라는 단어를 연속적으로 덮어씌어서 파티션에 대한 정보를 손상시켰다.

제작한 검증용 데이터 세트는 웹 사이트를[1] 통해 공개한다.

4.2 파티션 복구 도구 평가

테스트 대상이 되는 파티션 복구 도구는 데이터 복구 도구 성능 순위 사이트를 참조하였으며 그 중에서 자동화된 파티션 복구 기능을 지원하는 것으로 선정하였다〔3〕. EnCase와 같이 조사관이 직접 손상된 영역을 확인하여 수동으로 복구를 진행해야하는 도구는 본 검증에서 제외하였다.

복구 도구 목록은 Table.4와 같다.

1) http://forensic.korea.ac.kr/tooltest/testset.html

Table 4. Tools to be validated

No	Tools
1	DATA RESCURE PC3
2	EaseUS Partition Recovery v5.6.1
3	GetData Recovery My Files Pro v5.1.0.1824
4	MiniTool Power Data Recovery 6.6

실험에서는 논리적인 손상이 생긴 파티션 내부의 파일을 복구하는 것으로 파일의 복구 상태이나 데이터 카빙의 결과는 포함하지 않았다. 실험 결과는 모든 파티션이 정상적으로 복구되었을 경우에 "O", 1개 파티션이 정상적으로 복구되지 않은 경우에 "△", 2개 이상의 파티션이 온전하게 복구되지 않은 경우는 "X" 로 표기하였다. D, E, G, M은 도구명의 시작 알파벳이다. 각각의 데이터 세트별 복구 도구 검증 결과는 Table.5.와 Table.6.과 같다.

검증 결과를 통해 FAT32로 포맷 되어있는 경우는 DATA RESCURE PC3가 모든 손상 시나리오에 대해서 정확한 복구를 결과를 보였다. 그러나 그 외의 도구는 EBR 영역이 손상이 되었을 경우 복구가 되지 않는 한계를 보였다. NTFS로 포맷이 되어 있는 경우에는 EaseUS Partition Recovery v5.6.1과 MiniTool Power Data Recovery 6.6이 좋은 결과를 보였다.

Table 5. Validation Result(FAT32)

Disk	Scenario	D	E	G	M
MBR	Delete Master Boot Code and three Main BRs within MBR	O	O	X	△
	Delete three BR Entries within MBR Partition Table	O	O	O	O
	Delete Current Entry and Extended BR within EBR	O	O	X	△
	Delete EBR Entry and EBR Table within MBR Partition Table	O	O	O	O
	Delete all Boot Sector	O	X	X	X
GPT	Delete Main GPT Header	O	O	O	O
	Delete Main GPT Header and Backup GPT Header	O	O	O	O
	Delete Main Entries	O	O	O	O
	Delete Main Entries and Backup Entries	O	O	O	△

Table 6. Validation Result(NTFS)

Disk	Scenario	D	E	G	M
MBR	Delete Master Boot Code and three Main BRs within MBR	O	O	X	O
	Delete three BR Entries within MBR Partition Table	O	O	△	O
	Delete Current Entry and Extended BR within EBR	O	O	O	O
	Delete EBR Entry and EBR Table within MBR Partition Table	△	△	X	△
	Delete all Boot Sector	X	O	X	O
GPT	Delete Main GPT Header	O	O	△	O
	Delete Main GPT Header and Backup GPT Header	O	O	△	O
	Delete Main Entries	O	O	O	O
	Delete Main Entries and Backup Entries	O	O	O	O

V. 결 론

디지털포렌식 조사에 사용되는 도구의 신뢰성을 보장하기 위해서는 디지털포렌식 도구의 기능을 명세하고 기능을 검증할 수 있는 데이터 세트가 필요하다. 현재 공개되어 있는 검증용 데이터 세트의 경우 검증할 수 있는 디지털포렌식 도구의 종류가 한정적이며 파티션 복구 도구 검증을 위한 데이터 세트는 제공하지 않는다. 따라서 본 논문에서 파티션 복구 도구의 기능을 검증하고 파티션 복구 도구가 타당한 정보를 통하여 복구를 진행하는 지에 대해서도 확인할 수 있는 시나리오를 제작했다. 시나리오는 크게 디스크 인식 방식손상과 파일시스템 영역 손상으로 나눠서 제작했으며 이를 기반으로 데이터 세트를 개발했다. 그 후 데이터 세트를 기반으로 파티션 복구 기능을 제공하고 있는 도구를 검증했다.

본 논문에서는 PC 환경의 윈도우 운영체제에 맞춰서 시나리오를 제작했다. 그러나 최근에는 안드로이드 계열의 스마트 폰 및 MacOS를 활용하는 PC 등 다양한 운영체제의 사용이 증가하고 있으므로 그에 따른 복구 도구 검증을 위한 연구를 진행할 예정이다.

References

[1] https://www.cftt.nist.gov/

[2] http://dftt,sourceforge.net/

[3] http://www.toptenreviews.com/software/backup-recovery/best-data-recovery-software/

[4] Min-Seo Kim and Sang-jin Lee, "Development of Windows forensic tool for verifying a set of data," Journal of the Korea Institute of Information Security & Cryptology, Vol. 25, No. 6, pp. 1421-1433, Dec, 2015.

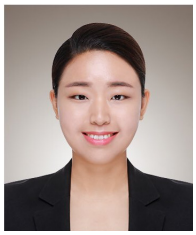
[5] Jaeung Namgung, Ilyoung Hong, Jungheum Park and Sangjin Lee, "A research for partition recovery method in a forensic perspective," Journal of the Korea Institute of Information Security & Cryptology, Vol. 23, No. 4, pp. 655-666, Aug, 2013.

[6] Guo, Yinghua, Jill Slay, and Jason Beckett. "Validation and verification of computer forensic software tools—Searching Function," Digital investigation, Vol. 6, pp. 12-22, Sep, 2009.

[7] Beckett, Jason, and Jill Slay. "Digital forensics: Validation and verification in a dynamic work environment," System Sciences 2007 HICSS 2007 40th Annual Hawaii International Conference on IEEE, pp. 266-266, Jan, 2007.

[8] Nikkel, Bruce J. "Forensic analysis of GPT disks and GUID partition tables." Digital Investigation, Vol. 6.1, pp. 39-47, Sep, 2009.

〈저자소개〉



박 송 이 (Songyee Park) 학생회원
2016년 2월: 서울여자대학교 정보보호학과 공학사
2016년 3월~현재: 고려대학교 정보보호대학원 정보보호학과 석사과정
〈관심분야〉 디지털 포렌식, 역공학, 딥러닝

허 지 민 (Gimin Hur) 학생회원
2015년 2월: 한양대학교 컴퓨터공학과 공학사
2015년 3월~현재: 고려대학교 정보보호대학원 정보보호학과 석박사통합과정
〈관심분야〉 디지털 포렌식, 역공학

이 상 진 (Sang-jin Lee) 종신회원
1989년 10월~1999년 2월: ETRI 선임 연구원
1999년 3월~2001년 8월: 고려대학교 자연과학대학 조교수
2001년 9월~현재: 고려대학교 정보보호대학원 교수
2008년 3월~현재: 고려대학교 디지털포렌식연구센터 센터장
2017년 3월~현재: 고려대학교 정보보호대학원 원장
〈관심분야〉 디지털 포렌식, 심층 암호, 해쉬 함수