macOS 파일시스템의 B-tree분석 디지털 포렌식 도구의 개발

조규상O
O동양대학교 컴퓨터학과
e-mail: cho@dyu.ac.krO

Development of B-tree Analyzing Tool for macOS Filesystem

Gyu-Sang ChoO
ODept. of Computer, Dongyang University

● 요 약 ●

본 논문에서는 macOS의 파일시스템인 HFS+의 B-tree구조를 디지털 포렌식의 관점에서 분석할 수 있는 기능을 갖춘 도구의 구현에 대하여 다룬다. HFS+ 파일시스템의 파일과 디렉토리에 대한 메타정보를 카탈로그 B-tree에서 구하여 디지털 포렌식 정보로 활용한다. HFS+파일시스템 포렌식 분석도구는 C/C++언어로 구현된다. 텍스트 기반의 명령행 프로그램으로 구현되며 macOS/Windows에서 터미널/명령프롬프트에서 각각 실행될 수 있도록 제작된다. 타임스탬프/파일크기/위치 등의 메타데이터의 파싱기능, 리프노드에 저장된 데이터를 이용한 파일/디렉토리 트리 구조의 재구성, B-tree구조에 의한 키워드 탐색 기능, 인덱스 노드 없이 B-tree 리프노드의 구성에 의한 파일/디렉토리 파싱/검색 기능 등이 구현된다.

키워드: 디지털 포렌식(digital forensics), macOS, HFS+ 파일시스템(filesystem), B-tree 분석(B-tree analysis)

I. Introduction

B-tree는 인덱스를 빠르게 검색할 수 있는 균형트리이다. 키값의 수가 증가하여도 깊이의 레벨이 많이 증가하지 않는 특징을 갖는다. NTFS, HFS+, APFS, Ext4등의 많은 파일시스템에서 B-tree를 파일/디렉토리를 관리하기 위한 방법으로 채택하고 있다[1].

디지털 포렌식 관점에서 B-tree에 관한 많지 않은 연구 중에서 Cho[2]의 연구에서는 일반 B-트리와 NTFS B-트리를 디지털 포렌식 관점에서 비교하여 이론적으로 구현된 것과 파일시스템에 현실적으로 구현된 것의 차이점에 관한 연구를 수행하였다. 리눅스 B-tree에 관한 연구[4]는 운영체제의 관점에서 BTRFS에 관한 분석을 하였다. HFS+의 B-tree에 관한 내용들은 여러 블로그나 자료를 통해서 발견할 수 있지만 디지털 포렌식을 위한 분석적인 연구는 상대적으로 발견하기 어렵다.

이 논문은 macOS에서 사용되는 HFS+ 파일시스템의 B-tree 분석기 도구 개발에 관한 것이다. 도구는 macOS/Windows에서 실행되도록 구현되며 타임스탬프/파일크기/클러스터 위치 등의 메타데이터의 추출 기능이 구현된다. 리프노드에 저장된 데이터를 이용한 파일/디렉토리 트리 구조의 재구성, B-tree구조에 의한 키워드 탐색 기능, 인덱스 없이 b-tree 리프노드의 구성에 의한 파일/디렉토리 파싱/검색 기능 등이 구현된다.

II. Implementation of B-tree Tool

2.1 Basic B-tree Structure of HFS+

HFS+ 파일시스템에 관련된 많은 구조들에 분석은 참고문헌[3]에서 제공하고 있다. HFS+의 파일시스템은 Volume Header, Allocation File, Extents Overflow File, Catalog File, Attributes File, Startup File, Alternate Volume Header, Reserved Area등으로 구성되어 있다. 이 중에서 Catalog File은 B-tree로 구조로 구현되어 있으며 파일/폴더에 대한 메타정보를 갖고 있다.

B-tree구조를 위한 노드디스크립터와 헤더레코드는 각각 BTNodeDescriptor, BTHeaderRec구조를 사용하며 카탈로그 파일 키는 HFSPlusCatalogKey구조, 카탈로그로그 파일 구조는 HFSPlusCatalogFile과 카탈로그 폴더는 HFSPlusCatalogFolder 구조, 카탈로그 스레드는 HFSPlusCatalogThread 구조를 사용한다[3]. 이 구조들을 사용하여 카탈로그 b-tree 분석도구를 구현한다.

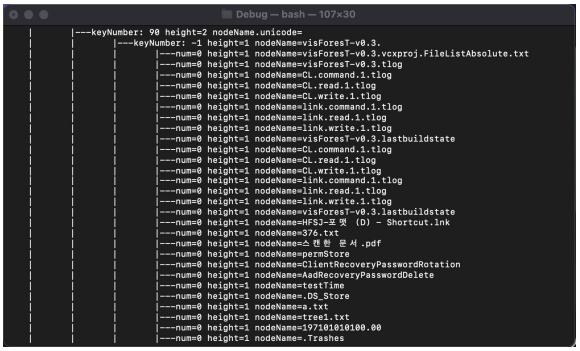

Fig. 1. Screen shot of the B-tree Analyzing Tool

2.2 Functions of B-tree Analyzing Tool

macOS의 HFS+파일시스템을 위한 B-tree 분석도구의 기능은 다음과 같이 구성된다. Fig. 1은 HFS+ 파일시스템 분석도구의 실행화면을 스크린 캡처한 것이다.

1) 메타데이터 파싱기능(생성시간, 수정시간, 시스템시간, 접근시간, 백업시간 등의 타임스탬프 추출)
2) 파일/디렉토리에 대한 정보 추출
3) 리프노드에 저장된 데이터를 이용한 파일/디렉토리 트리 구조의 재구성 기능. (부분적으로 손실이 있는 경우에 대하여 리프노드에 들어 있는 정보만으로 디렉토리 구조를 복구할 수 있는 기능)
4) B-tree구조에 의한 빠른 키워드 탐색 기능
5) 인덱스 노드 없이 B-tree 리프노드의 순서에 의한 재구성

2.3 Development Environment

이 도구를 실행하기 위해서는 관리자 기능으로 실행되어야 한다. 이 도구는 Windows와 macOS에서 각기 실행될 수 있도록 개발되었다. macOS에서는 장착되어 있는 디스크에서 바로 프로그램으로 읽어서 실행하는 방식으로 구현되었다.

Windows에서는 HFS+가 파일시스템으로 채택되어 있지 않지만 많은 포렌식 도구들이 Windows용으로 제공되고 있어서 호환을 위하여 기능을 구현한 것이다. HFS+의 이미징 파일을 읽는 방식과 mac Drive 10 프로그램을 실행하여 드라이브를 읽을 수 있는 방식으로 실행할 수 있도록 개발되었다.

Table 1. Run/Development Environment

	macOS	Windows
OS Ver.	Catalina ver. 10.15.4	Windows 10 Pro
Run Env.	Terminal (Root Privilege)	Command Prompt (Administrative Privilege)
Lang./ Dev. Tools	C/C++ xCode ver. 12.3	C/C++ Visual Studio 2019

III. Conclusions

이 논문에서는 macOS에서 사용되는 HFS+ 파일시스템의 B-tree 분석도구를 구현하였다. C/C++언어로 구현된 이 도구는 macOS/Windows에서 실행되며 HFS+의 메타데이터를 이용한 일반적인 정보추출과 리프노드에 저장 데이터만을 이용하여 구현한 기능들이 제공된다. 추후로 GUI 환경에서 동작하는 프로그램으로 구현하는 것이 필요하다.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (NRF-2019R1F1A1058902)

REFERENCES

[1] Wikipedia, “B+ tree”, https://en.wikipedia.org/wiki/B+tree.
[2] G.-S. Cho, "Ordinary B-tree vs NTFS B-tree: A Digital Forensics Perspectives," Journal of The Korea Society of Computer and Information, Vol. 22, No. 8, pp. 73-83, August 2017.
[3] Technical Note TN1150, HFS Plus Volume Format, https://developer.apple.com/library/archive/technotes/tn/tn1150.html
[4] Ohad Rodeh et. al, "BTRFS: The linux B-tree filesystem", ACM Transactions on Storage, Vol. 9, No. 3, August 2013.