http://dx.doi.org/10.3745/KIPSTC.2012.19C.1.001

클라이언트관점의 SaaS 사용 흔적 분석

강 성 림[†] · 박 정 흠[††] · 이 상 진[†††]

요 약

최근 초고속 통신망의 발달과 클라우드 컴퓨팅 기술을 활용한 온 디멘드(On-demand)형 소프트웨어 SaaS(Software as a Service)의 사용이 크게 증가하고 있다. 하지만 이러한 클라우드 컴퓨팅 환경에 대한 디지털 포렌식은 현재 학문적, 실용적 체계화 수준이 미비한 실정이다. 또한 클라우드 환경의 특성상 사용자 행위에 의한 데이터가 로컬 시스템에 거의 남지 않고, 원격지 서버에 분산 저장되기 때문에 디지털 포렌식 조사관점에서 많은 어려움이 따른다. SaaS는 기본적으로 웹을 통해 접속하는 서비스 형태이기 때문에, 클라이언트관점에서 History files, Cookie files, Temporary Internet Files, 물리메모리 등의 분석이 중요하다. 본 논문에서는 현재 널리 쓰이는 SaaS를 선정하여, 클라우드 컴퓨팅 서비스 타입 중 하나인 SaaS 사용 흔적 분석 방안을 제시한다.

키워드 : 클라우드 컴퓨팅, SaaS(Software as a Service), 디지털 포렌식, 인터넷 임시파일

The Trace Analysis of SaaS from a Client's Perspective

Sunglim Kang[†] · Jungheum Park[††] · Sangjin Lee[†††]

ABSTRACT

Recently, due to the development of broadband, there is a significant increase in utilizing on-demand Saas (Software as a Service) which takes advantage of the technology. Nevertheless, the academic and practical levels of digital forensics have not yet been established in cloud computing environment. In addition, the data of user behavior is not likely to be stored on the local system. The relevant data may be stored across the various remote servers. Therefore, the investigators may encounter some problems in performing digital forensics in cloud computing environment.it is important to analysis History files, Cookie files, Temporary Internet Files, physical memory, etc. in a viewpoint of client, since the SaaS basically uses the web to connects the internet service. In this paper, we propose the method that analysis the usuage trace of the Saas which is the one of the most popular cloud computing services.

Keywords : Cloud Computing, SaaS(Software as a Service), Digital Forensic, Temporary Internet Files

1. 서 론

클라우드 컴퓨팅이란 서로 다른 물리적인 위치에 존재하는 컴퓨터들의 리소스를 가상화 기술로 통합해 제공하는 기술이다. 이러한 클라우드 컴퓨팅 기술을 기반으로 제공되는 서비스는 IaaS(Infra as a Service), SaaS(Software as a Service), PaaS(Platform as a Service) 세 가지 타입으로 분류된다[1].

세 가지 서비스 타입 중 SaaS는 웹 기반의 소프트웨어로써, 소프트웨어의 여러 기능 중에서 사용자가 필요로 하는 서비스만 이용 가능하도록 한 온 디멘드(On-demand)형 소프트웨어이다.

현재 웹을 통해 언제 어디서든 접속이 가능한 사용자 접근성과 편의성을 장점으로 점차 오프라인 소프트웨어를 대체해 나가고 있다. 현재 가장 널리 쓰이는 웹 기반 문서편집도구인 'Google Docs'의 경우, 온라인상에서 이질감 없는 인터페이스와 빠른 속도, 다양한 포맷의 문서생성과 편집기능, 무료 데이터 저장 공간을 제공함으로써, 사용자들로부터 큰 호응을 얻고 있다.

하지만 디지털 포렌식 관점에서 이러한 웹 기반 소프트웨어를 사용한 용의자의 시스템을 분석해야하는 경우, 중요 사용자 문서 데이터가 대부분 해당 서비스 서버에 존재하기 때문에, 기존의 디지털 포렌식 기술을 활용한 데이터수집에 많은 제약이 따른다.

SaaS는 기본적으로 서버 클라이언트 구조로서, 웹을 통해 접속하는 형태이기 때문에, 클라이언트에 사용자가 저장

※ 본 연구는 한국연구재단을 통해 교육과학기술부의 바이오연구개발사업으로부터 지원 받아 수행되었음(2011-0027732).
† 준 회 원 : 고려대학교 정보보호대학원 석사과정
†† 준 회 원 : 고려대학교 정보경영공학전문대학원 박사과정
††† 종신회원 : 고려대학교 정보경영공학전문대학원 교수(교신저자)
논문접수 : 2011년 7월 7일
수 정 일 : 1차 2011년 10월 10일, 2차 2011년 11월 2일
심사완료 : 2011년 11월 2일

한 데이터가 남지 않는다. 하지만 웹 브라우저 사용흔적인 History files, Cookie files, Temporary Internet Files과 사용자의 활성데이터를 저장하고 있는 물리메모리 분석을 통해 중요한 사용자 데이터 수집이 가능하다.

본 논문에서는 클라이언트 관점에서 널리 쓰이는 SaaS를 선정하여 사용 흔적을 분석하고 SaaS 환경을 고려한 포렌식 조사방법론을 제시한다.

2. SaaS에 대한 디지털 포렌식 문제점

클라우드 컴퓨팅 서비스의 특성상 사용자 데이터는 클라우드 서버 어딘가에 분산된 형태로 존재하기 때문에, 디스크, 메모리, 네트워크 등의 하드웨어 자원을 공유하는 가상환경에서 개체를 물리적으로 수집하기 어렵다. 특히, 'Google Docs'나 'MS Live Office' 등과 같은 웹 기반 문서 편집도구의 등장으로 기존의 오프라인에서 이루어지던 문서작성, 스프레드시트, 프레젠테이션 제작 작업을 웹서비스를 통해 제공한다. 또한 온라인상의 무료 데이터 저장 공간을 제공함으로서, 사용자의 시스템에 문서파일을 저장할 필요 없이 온라인상에 저장하여 언제든지 단말기를 통해 접근 가능하다.

포렌식 관점에서 문서 파일은 중요한 정보가 저장되는 경우가 많기 때문에 반드시 수집해야 하는 데이터이다. 하지만 SaaS를 이용하여 문서 작성이 진행되면 클라이언트 상에서는 의미있는 데이터가 남아있지 않아 포렌식 조사에 어려움이 발생한다. 따라서 기존의 포렌식 조사절차와는 차별화된 포렌식 조사방법이 요구된다.

3. 클라이언트관점의 SaaS 포렌식 조사 방법

3.1 SaaS 포렌식 조사 방법론

클라우딩 컴퓨팅 기술을 활용한 서비스가 도입되면서, 기존의 디지털 포렌식 절차에서 추가적으로 고려해야 할 사항은 해당 조사 시스템에서 클라우드 컴퓨팅 서비스의 사용유무를 파악하는 것이다. 클라우드 컴퓨팅 서비스의 종류(IaaS, PaaS, SaaS)에 따라 분석방법론이 나누어지며, 사용자가 클라우드 컴퓨팅 서비스를 사용한 흔적이 발견된 경우, 사용자 관련 중요 데이터는 해당 서비스 서버에 존재할 가능성이 크다. 이를 간과할 경우, 핵심적인 증거 데이터들을 획득하지 못하고, 정밀한 증거분석을 할 수 없기 때문에 포렌식 분석에 차질이 발생될 가능성이 큰바, 기존 포렌식 조사방법과는 차별화된 조사방법론이 필요하다.

클라우드 서비스를 사용하게 되면 각 단말기에는 해당 서비스를 사용한 흔적이 남게 되어 클라우드 서비스를 사용했는지 여부를 판단할 수 있다. SaaS를 사용하기 위해서는 해당 서비스를 제공하는 사이트에 접속해야 하기 때문에 인터넷 히스토리에 사이트 접속 기록이 남게 된다. 또한 물리메모리에는 해당 서비스를 알 수 있는 고유한 문자열이 존재한다. <표 1>은 대표적인 SaaS에 대한 시그니쳐를 정리한 내역이다.

<표 1> 서비스별 SaaS 시그니쳐

종류	서비스	웹 브라우저 흔적	물리메모리 흔적
SaaS	Think Free	thinkfree.com	smb_user
	Zoho Ofiice	zoho	gmail.com& PASSWORD=
	MS Live Online	office.microsoft	hotmail.com&passwd=
	Google Docs	docs.google	gmail.com&Passwd=
	Glice OS	glidedigital	<tns:username>

클라우드 컴퓨팅 서비스 사용유무를 확인 후, 웹을 통해 해당 서비스 서버에 접속을 통해 데이터 수집을 할 수 있다. 하지만 이는 사용자의 아이디와 패스워드를 인정절차로 사용하기 때문에 분석 이미지 내에서 민감한 개인정보를 발견하기가 쉽지 않다. 온전한 사용자 데이터파일을 획득하기 위해서는 해당 클라우드 서비스에 접속을 통해 데이터를 수집해야 하지만 이는 클라우드 서비스 특성상 여러 가지 제약이 존재한다.

클라우드 서비스를 사용하면, 쿠키정보에서 사용자 아이디를 일부 서비스에서 확인가능하며, 인터넷 임시파일을 통해 사용자의 데이터를 일부 확인 가능하다. 또한 활성상태의 시스템분석일 경우 물리메모리 덤프이미지를 획득하면 사용자 계정정보와 작성중인 문서파일정보를 확인할 수 있다.

웹 기반 서비스 종류에 따라 남는 흔적파일은 다르지만, 이는 수사관 입장에서 인터넷 임시파일에 존재하는 일부 작성한 중요문서 정보나 사용자의 행위기반 정보추출을 통해 추후 사건분석 진행시에 상당한 도움이 될 수 있다.

3.2 주요 분석 요소

대부분의 클라우드 컴퓨팅 서비스를 이용하기 위해서는 웹 브라우저를 통해 해당 클라우드 서버에 접속해야 한다. 기존의 웹 포렌식 도구들의 주요 분석 데이터는 History files, Cookie files, Temporary Internet Files이며 이는 웹사이트 방문목록, 마지막 사이트 방문시간, 검색어 등의 데이터 추출을 통해 범죄행위를 밝히는데 결정적인 증거데이터를 제공한다. 특히, 주요 웹서비스에 대한 Temporary Internet Files 분석은 분석관에게 피의자가 사건 발생 시점 전후로 어떠한 행동을 했는지 입증해 주는 증거가 될 수 있다. 예를 들어 피의자가 웹 메일이나 SaaS를 사용하였다면 사용자 정보가 Temporary Internet Files에 저장되었을 가능성이 크다. 그리고 웹 사이트로부터 문서 파일을 열었다면 열어본 파일 또한 Temporary Internet Files에 저장되었을 가능성이 높다.

또한, 물리메모리는 현재 작업중인 문서정보 일부와 웹 브라우저를 통해 로그인한 사용자 아이디, 비밀번호 등의 당시 사용자 행위 기반 데이터를 저장하고 있다. 따라서 Temporary Internet Files과 물리 메모리는 SaaS 사용흔적 분석시 반드시 고려해야 하는 데이터이다.

3.2.1 Temporary Internet File

Temporary Internet File이란 웹 브라우저를 통해 웹 서버에 접속할 때 서버에서 전송하는 각종 HTML 파일, 쿠키, 이미지, 플래시 파일을 임시로 클라이언트에 저장한 파일이다. 인터넷 접속 시에 처음 방문한 웹 페이지는 재접속 시에 빠른 웹페이지 로드를 위하여 인터넷 임시 파일 폴더에 다양한 종류의 임시파일을 저장한다[2]. 마이크로소프트의 Internet Explorer의 경우, Temporary Internet File폴더내부의 'index.dat' 파일에서는 저장된 임시 파일의 URL, 임시파일의 저장위치, 임시파일의 이름, 접속 시간을 확인할 수 있다.

이러한 인터넷 임시파일은 디지털 포렌식 조사관점에서 사용자의 행위를 유추할만한 유용한 정보의 파편을 담고 있으며, 서비스 별로 생성되는 임시파일의 종류와 획득정보가 다르며, 특정위치에 빈번하게 삭제 또는 덮어 쓰여지는 특징이 있으므로 활성 데이터와 같이 최대한 빨리 수집해야 한다.

3.2.2 History files

History files은 사용자가 웹 브라우저를 통해 방문한 URL을 저장한다. 또한 최초 방문 시간, 마지막 방문 시간, 방문횟수, 웹 페이지의 제목과 같은 정보를 확인할 수 있다. 또한 사용자가 검색한 사이트에서 무엇을 검색했는지, 어떤 파일을 다운받았는지의 여부를 알 수 있다. Windows Internet Explorer의 경우, 'index.dat' 이라는 파일에 앞서 제시한 정보를 저장하고 있다. (그림 1)은 History files 내부에서 확인한 주요 웹 기반 문서편집도구 방문 URL목록이다.

3.2.3 Cookie files

Cookie files은 접속된 사이트에 접속한 사용자가 누구인지를 식별하기 위한 정보를 저장한 파일이다. 개인의 특정 정보나 방문한 사이트를 알 수 있다. 해당 사이트에 로그인 시 아이디 기억하기 옵션을 체크하면 해당 아이디가 쿠키내에 저장되어, 사용자 아이디 또한 추출가능하다. (그림 2)는 해당 쿠키 파일의 마지막 접근시간을 통해 해당 사이트에 마지막으로 방문한 시간을 파악한 것이다.

<표 2>는 MS Windows 운영체제의 버전별 Internet Explorer의 Temporary Internet Files, Cookies, History file의 저장위치이다.

〈표 2〉 윈도우 운영체제 버전별 웹 사용흔적파일 저장위치

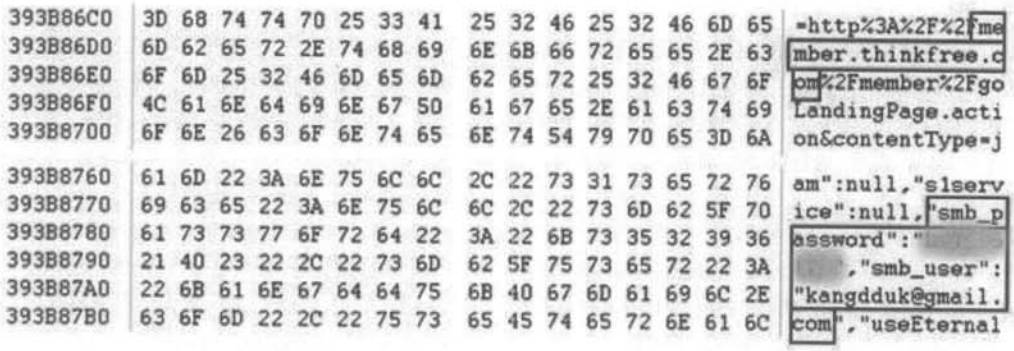
운영체제	저장위치
Windows XP	• Documents and Settings\<username>\Local Settings\TemporaryInternetFiles\Content.IE5\ • Documents and Settings\<username>\Cookies\ • Documents and Settings\<username>\Local Settings\History\History.IE5\
Windows Vista/7	• Users\<username>\AppData\Local\Microsoft\Windows\TemporaryInternetFiles\Content.IE5 • Users\<username>\AppData\Local\Microsoft\Windows\History\History.IE5 • Users\<username>\AppData\Roaming\Microsoft\Windows\Cookies

Internet Explorer	gm...	https://docs.google.com/spreadsheet/auth/getcooki...	2011-10-03 22:19:46
Internet Explorer	htt...	https://0.docs.google.com/bind?id=ignored&sid=1...	2011-10-03 22:19:44
Internet Explorer	htt...	https://0.docs.google.com/test?id=ignored&sid=1c...	2011-10-03 22:19:41
Internet Explorer	htt...	https://docs.google.com/?pli=1	2011-10-03 22:19:41
Internet Explorer	htt...	http://office.microsoft.com/ko-kr/	2011-10-03 22:19:26
Internet Explorer	abo...	about:blank	2011-10-03 22:19:25
Internet Explorer	htt...	http://www.microsoft.com/office/webapps/demo/d...	2011-10-03 22:19:24
Internet Explorer	htt...	http://r.office.microsoft.com/r/rlidPowerPointEmbe...	2011-10-03 22:19:24
Internet Explorer	htt...	http://www.zoho.com/	2011-10-03 22:19:07
Internet Explorer	htt...	http://member.thinkfree.com/member/goLandingP...	2011-10-03 22:18:38

(그림 1) history files 내부에 저장된 방문 URL 목록

Browser	Host	Last Access Time	Name	Value	Expiry
Internet Explorer	google.com/	2011-10-03 22:33:01	SID	DQAAAKcAAA...	2021-09-30 22:31:39
Internet Explorer	google.co.kr/	2011-10-03 22:33:20	SID	DQAAAKcAAAB...	2021-09-30 19:58:06
Internet Explorer	thinkfree.com/	2011-10-03 22:33:28	__utma	50139961.1545...	2013-10-02 22:30:02
Internet Explorer	docs.google.com/	2011-10-03 22:33:47	WRITELY_SID	DQAAAKkAAAA...	2021-09-30 22:17:20
Internet Explorer	www.microsoft.com/	2011-10-03 22:34:42	WT_NVR	0=/:1=office:2=...	2021-09-30 22:17:20
Internet Explorer	c.live.com/	2011-10-03 22:34:57	PREFBIN	0	2021-01-01 09:00:00
Internet Explorer	login.live.com/	2011-10-03 22:34:57	MSPShared	1	2037-12-31 00:59:55
Internet Explorer	c.msn.com/	2011-10-03 22:34:57	ANONCHK	0	2011-10-06 16:12:57
Internet Explorer	office.microsoft.com/	2011-10-03 22:34:58	awsuserguid	guid=1111d38a...	2012-10-03 22:17:55

(그림 2) 쿠키파일내부의 정보



(그림 3) 물리메모리 이미지파일 내의 사용자 아이디, 패스워드 정보

3.2.4 물리 메모리

물리 메모리에는 프로세스 정보와 이름, 네트워크 세션정보뿐만 아니라, 현재 작업 중인 사용자 행위 기반 활성데이터를 저장하고 있다. 웹 접근 시 SaaS 사용자 정보(웹 브라우저를 통해 로그인한 아이디, 비밀번호, 작성한 문서 일부 정보)등이 역시 물리 메모리 영역에 남아 있을 여지가 있기 때문에 반드시 수집해야 할 데이터이다. 모든 서비스 사용 후 공통적으로 해당하는 것으로 앞서 제시한 서비스별 지정 시그니쳐를 통해 사용자 계정정보를 추출할 수 있다. (그림 3)은 Thinkfree 서비스 사용 후 물리메모리 덤프이미지에서 발견한 사용자 아이디와 패스워드 정보이다.

4. SaaS 주요 서비스 사용 흔적 분석

본 논문에서는 현재 널리 사용되는 SaaS를 선정하여 클라이언트 관점에서 사용 흔적을 분석한다. 웹 기반 문서 편집도구인 구글의 'Google Docs', 마이크로소프트의 'MS Live Office', Zoho의 'Zoho office', 국내 한컴의 'Thinkfree'와 'Web OS' 서비스인 'Glide OS'를 분석대상으로 선정하였다. 많은 클라우드 컴퓨팅 서비스 중 웹 기반 문서 편집도구에 중점을 둔 이유는, 디지털 포렌식 분야에서의 문서파일의 중요성 때문이다. 사용자가 작성한 문서파일이나 보유파일은 기업회계부정이나, 개인의 사적인 정보 등을 기록해 놓아 사건의 중요한 증거요소가 되기 때문에 이러한 사용자 데이터를 수집하는 것은 가장 중요한 포렌식 분석절차중 하나이다. 또한 'Web OS'는 지금은 많은 사람들이 인지하지 못하고 있지만, 웹을 통해 생성한 자신만의 새로운 공간을 통해 여러 가지 작업(이메일전송, 캘린더기능, 문서작성)이 가능하다. 악의적인 행위를 구상할 경우, 이러한 공간을 악용할 수 있기 때문에 고려해야 할 서비스이다.

본 논문에서는 클라이언트 측면에서 웹 브라우저는 MS Internet Explorer v.10 과 Windows 7 운영체제 환경에서 아래 <표 3>에 제시한 SaaS의 사용 흔적을 분석한다. 앞서 제시된 주요 분석 요소인 History file과 Cookie file, 물리메모리에서 획득정보는 동일하기 때문에, 다음 장에서는 인터넷 임시파일을 서비스 별로 분석한다.

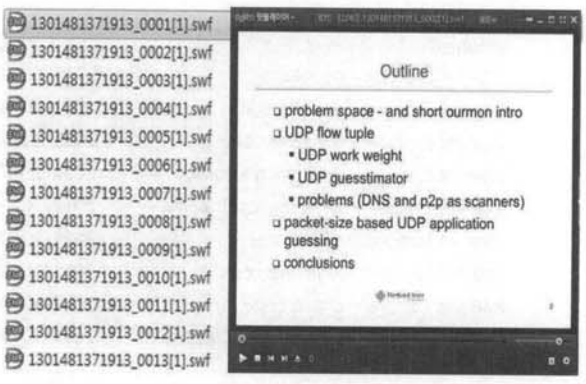
〈표 3〉 분석 대상 SaaS

서비스종류	벤더	서비스 이름
문서편집	한컴	ThinkFree
	Google	Google Docs
	Zoho	Zoho Office
	Microsoft	MS Live Office
Web OS	Transmedia	Glide OS

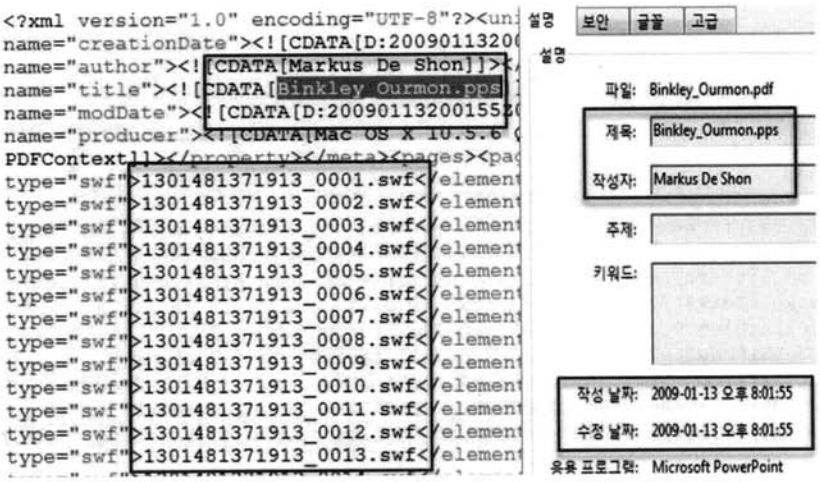
4.1 Thinkfree

국내 웹기반 편집도구인 'Thinkfree'는 'Java Web Start'란 기술을 사용한다. 'Java Web Start'란 'Java'로 작성된 많은 어플리케이션을 웹을 통해 배포, 실행 등을 할 수 있는 환경을 말한다. Java환경의 웹 어플리케이션의 경우, 클라이언트에서 웹 브라우저를 통해 문서를 로딩(loading)시 문서 각각의 페이지가 플래시 파일 형태로 변환되어 임시파일로 생성된다. 이러한 플래시 임시파일 데이터는 하나의 문서파일이 각각의 슬라이드형태로 구성되어 모든 문서의 본문내용을 저장하고 있다. 이는 분석관에게 사용자가 어떤 내용의 문서를 언제 열람했는지 여부를 파악할 수 있는 중요한 증거 데이터를 제공한다.

(그림 4)는 사용자가 열람했던 프리젠테이션 파일의 각 슬라이드가 플래시파일 형태로 로컬 시스템 인터넷 임시파일폴더 경로에 저장된 것을 보여준다.



(그림 4) 각 슬라이드의 플래시 파일



(그림 5) 파일별 플래시 구성 파일 확인

이 플래시 임시파일은 '[file upload time][random num]_[slide num][n].swf' 의 파일명명 규칙을 가진다. 파일명 맨앞 10자리 숫자는 파일 업로드 시각이 Unixtime 형식으로 저장되며, 서버가 구글과 연동되기 때문에 UTC 시각을 기준으로 한다. 그 이후 3자리 숫자는 랜덤하게 생성되는 넘버이며 '_'이후 4자리 숫자는 각각의 슬라이드 번호를 의미한다.

또한 열람한 문서가 전체가 몇장의 슬라이드인지, 어떤 인터넷 임시파일로 구성되어 있는지 전체구조를 파악할 수 있는 '[filename]_unipaper[n].xml' 파일이 생성된다. (그림 5)에서 'Bikley_Ourmon.pdf'파일이 2009년 01월 13일 오후 08:01:55'의 시각에 생성되었고, 13장으로 구성된 pdf파일을 열람했음을 짐작할 수 있다. 또한 클라이언트에 생성된 각각의 플래시 임시파일명을 확인할 수 있고, 해당 문서의 초기 생성시간과 최종 수정시간을 알 수 있다.

텍스트 파일의 경우 웹 브라우저 상에서 열람하는 순간 클라이언트에 본문정보를 저장한 '[파일명][n].txt' 파일이 생성된다. 이 텍스트파일은 모든 본문의 내용을 저장한 인터넷 임시파일이다.

또한, 웹 브라우저를 통해 임의의 문서편집기능을 수행할 경우 (그림 6)과 같이 파일의 기본정보와 사용자정보가 저장된 'goPowerEdit[n].htm', 'goSharedView[n].htm' 라는 임시파일이 클라이언트에 생성된다. 이 두 개의 파일에서 사

```
<script type="text/javascript>
        var baseurl='member.thinkfree.com';
        var memberurl = "http://member.thinkfree.com/member/";
        var signinurl = "http://member.thinkfree.com/member/gosignin.action";
        var staticresourcebaseurl = "/static/";
        var userid ="kangdduk@gmail.com";
        var uid = 10704200;
        var did = 1;
        var vkey = 1300928104890;
        var fileindex = 6458203;
        var app = 'show_editor';
        var filename = "test.pptx";
        var createtime = "1302711035074";
        var editor = 'v';
        var roleindex = 1;
        var shared = false;
        var published = false;
        var tmp = null;
```

(그림 6) 'goPowerEdit[n].htm' 파일내의 주요 정보

용자 아이디, 작업 중인 파일명, 편집상태, 공유상태, 생성시간 등을 알 수 있다.

이 파일들은 편집 창을 종료할 경우, 파일이 삭제되지만 복구도구를 통해 정상적으로 복구가능하다.

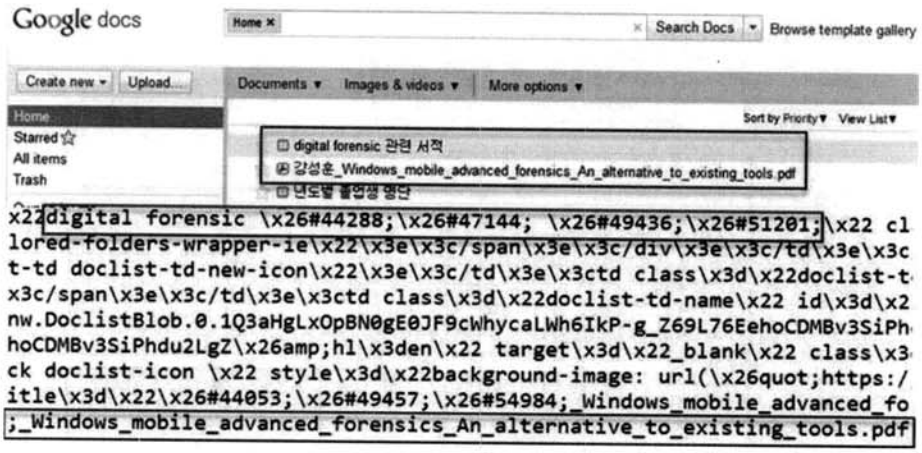
<표 4>는 사용자가 Thinkfree를 사용했을 경우, 클라이언트에 생성되는 주요 임시파일의 종류와 획득가능 정보이다.

4.2 Google Docs

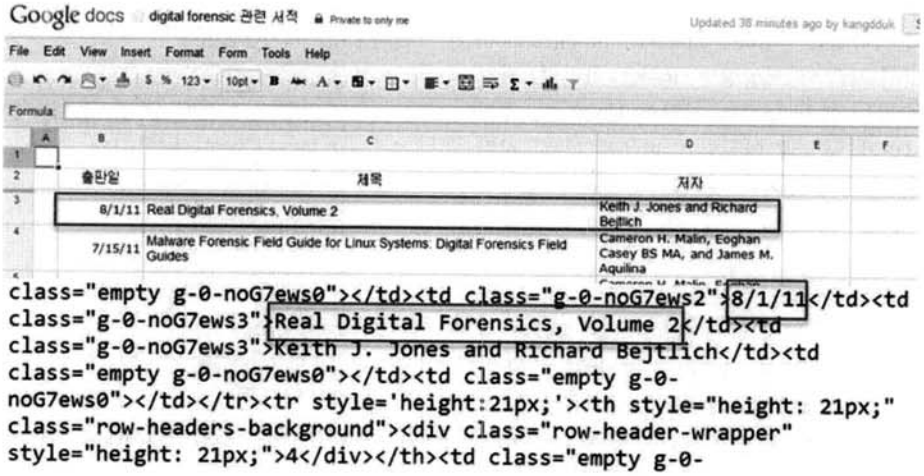
'Google Docs'는 처음 접속했을 때, (그림 7)과 같이 사용자가 서버에 저장한 문서목록 정보을 저장하고 있는

〈표 4〉 Thinkfree 사용 흔적 임시파일

사용자 행위	문서파일종류	흔적이 남는 임시파일	비고
문서편집	txt, ppt, doc, xls, hwp	goPowerEdit[n].htm, goSharedView[n].htm	User ID, 편집중인 파일명, 파일생성시간, 설정정보
문서열람	txt	[filename][n].txt	원문형태 그대로 텍스트 파일 형태로 존재
	ppt, pdf, doc, xls, csv, hwp	[file upload time][random num]_[slide num][n].swf	각 슬라이드가 .swf 파일형태로 존재



(그림 7) 사용자 문서목록 파일



(그림 8) xls 문서내의 본문 저장 문자열 확인

'Docs_google_com[n].htm' 파일이 임시 생성된다. 사용자 문서목록 중 한글문자열의 경우 유니코드로 인코딩되어 'x26#[10진수]' 형태로 값이 저장된다.

'Google Docs' 문서편집 기능을 실행했을 경우 'edit[n].htm'라는 파일명을 가진 임시파일이 생성된다. 이 HTML 파일 내부에는 문서편집을 위해 열람했던 문서상의 한 페이지 문자열만 저장되어 있다. 이는 구글이 빠른 웹페이지 로딩속도를 위해 일부 페이지만 로드한 후, 필요 요청시에 다음 페이지를 불러오는 방식으로 웹서버와 통신하기 때문이다.

엑셀파일의 경우, 문서편집 기능을 수행했을때 'ccc[n].htm'파일이 생성된다. 이 파일은 최대 101셀까지의 정보를 저장하고 있다. (그림 8)은 'ccc[n].htm' 파일에서 엑셀파일의 문자열 데이터 일부를 확인한 화면이다.

또한 슬라이드에 삽입된 그림은 'viewer[n].png' 혹은 'viewer[Random string].png' 라는 파일명을 가진 이미지 파일이 클라이언트에 임시파일로 생성된다.

<표 5>는 사용자가 Thinkfree를 사용했을 경우, 클라이언트에 생성되는 주요 임시파일의 종류와 획득가능 정보이다.

〈표 5〉 Google Docs 사용 흔적 임시파일

사용자 행위	문서파일 종류	흔적이 남는 임시파일	비고
문서 목록	모든 문서 파일목록	Docs_google_com[n].htm	
문서 편집	doc | docx, ppt, rtf, txt	edit[n].htm	윈도우 편집창을 닫으면 삭제됨 doc문서의 경우 한페이지만 저장
	xls | xlxs, csv	ccc[n].htm	101열까지 저장
문서 열람	pdf	viewer[n].txt	pdf 내의 문자열 저장
		viewer[n].png, viewer[Random string].png	각장의 슬라이드가 그림파일로 남음
	pptx	viewer[n].png, viewer[Random string].png	

4.3 Zoho office

'Zoho office'에서 ppt, doc파일을 열람했을 경우, 'viewdoc[n].htm' 와 'docview[n].htm' 이란 임시파일이 생성되며 해당 열람문서의 모든 문자열이 저장 되어있다. 또한 문서 내에 삽입된 이미지의 경우, 'ImageView[n].png' 혹은 'ImageViewCA[Random String].png' 이미지 파일로 클라이언트에 임시파일로 남는다. 이러한 이미지 임시파일의 경우, 새로운 문서를 열람시 마다 같은 동일한 파일명으로 덮어쓰여지므로 파일복구가 어렵다.

<표 6>는 사용자가 'Zoho office'를 사용했을 경우, 클라이언트에 생성되는 주요 임시파일의 종류와 획득가능 정보이다.

〈표 6〉 zoho office 사용 흔적 임시파일

사용자 행위	문서파일 종류	흔적이 남는 임시파일	비고
문서열람	ppt, doc	viewdoc[n].htm docview[n].htm	본문내용 저장
		ImageView[n].png ImageViewCA[Random String].png	문서내 삽입된 그림파일
	txt	documentView[n].im	본문내용 저장

4.4 MS Live Office

MS Live Office의 경우, 사용자의 문서목록정보를 가진 'Office_live_com[n].htm'파일이 임시생성 된다. 문서목록 중 한글의 경우, 유니코드로 인코딩되어 '&#[10진수]'로 변환된 값이 저장된다.

그 외의 흔적은 클라이언트에서 발견할 수 없었다. 이는 해당 문서편집이나 열람을 했을 때 생성된 파일이 화면에 로드 후 바로 삭제되기 때문에 로컬 시스템에 임시파일로 저장되지 않는다.

4.5. Glide OS

웹기반 데스크탑 서비스는 간단한 문서작성이나 일정관리, 주소록, 이메일, 가상 하드디스크 공간 등을 제공한다. 대표적 웹기반 데스크탑 서비스인 'Glide OS'에 로그인 인증을 통해 접속했을 경우 'Session[n].xml'과 'FileSystem[n].xml' 라는 임시파일이 생성된다. 생성된 파일내부에서 서비스 타입(유료/무료) 디스크 할당량, 사용자(이름, 이메일, 아이디)정보를 확인할 수 있다.

웹 기반 데스크탑 서비스는 또한 웹 하드형식으로 가상의 사용자가 데이터를 저장할 수 있는 공간을 제공한다. 파일을 업로드 하면, 'update_file[n].xml'과 'download[n].htm'이란 임시파일이 생성되며 이 두 파일에는 업로드 하는 서버에 위치한 파일의 URL, 파일크기, 생성시간, 최종 수정시간, 송수신 IP 등의 정보가 저장되어 있다.

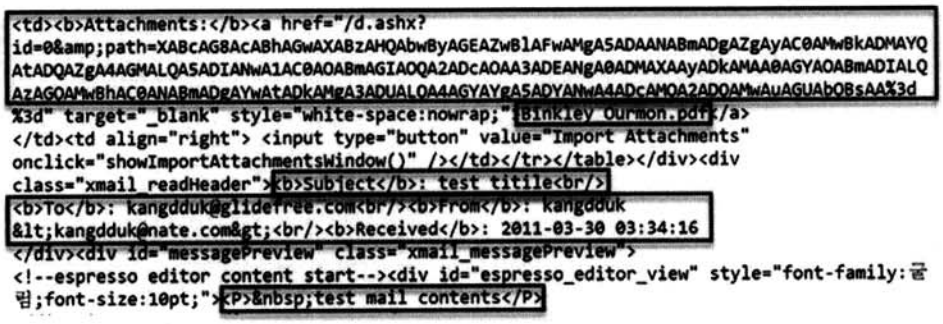
'Glide OS'의 또 다른 기능으로 일정관리를 할 수 있는 Calender 기능을 제공한다. Calender창을 활성화 시켰을 때, 사용자의 일정정보를 확인할 수 있는 'Calender[n].xml' 파일이 클라이언트에 생성된다. 'Calender[n].xml' 파일에서 획득가능한 정보는 일정시간, 스케줄, 간단한 메모정보 등을 확인할 수 있다.

또한 사용자가 'Glide OS' 회원가입 시 '[사용자]@glidefree.com' 메일계정이 자동 생성되어 이메일을 송수신할 수 있다. 사용자가 송수신한 메일 목록은 'Glide_box[n].htm'과 'Inbox[n].htm'에 저장되어 있다. 그리고 메일을 수신했을 경우, (그림 9)과 같이 'Read[n].htm'파일과 'Read_send[n].htm'에 송수신자의 첨부파일의 경로, 첨부파일명, 메일제목, 송신자, 수신시각, 본문내용등의 정보가 저장되어 있다.

<표 7>는 사용자가 'Glide OS'를 사용했을 경우, 클라이언트에 생성되는 주요 임시파일의 종류와 획득가능 정보이다.

〈표 7〉 Gilde OS 사용 흔적 임시파일

주요 기능	생성되는 임시파일	획득 가능한 사용자정보	비고
로그인	Session[n].xml FileSystem[n].xml	디스크할당량, 서비스타입, 사용자 정보	
웹하드	update_file[n].xml download[n].htm	파일의 경로, 사이즈, 생성시간, 수정시각, ip정보	
일정 관리	Calender[n].xml	일정내용, 시각	
이메일	Read[n].htm Read_send[n].htm	송수신자 메일주소, 첨부파일 경로, 발송시각, 메일본문	메일 송수신
	Glide_box[n].htm Inbox[n].htm	발신자, 메일제목, 첨부파일명, 발송시각	수신 메일 목록함
문서 열람	[filename][n].htm	문서내용	



(그림 9) 'Read[n].htm' 파일 내부 사용자 정보

5. 결 론

최근 클라우드 컴퓨팅 기술을 활용한 SaaS가 점차 보편화됨으로써, 이러한 환경에 대한 디지털 포렌식 연구의 중요성이 더욱 증대되고 있다, 본 논문에서는 클라이언트 관점에서 현재 널리 사용되는 대표적인 SaaS를 대상으로 사용 흔적을 분석하였다. 클라우드 컴퓨팅 기반의 SaaS 사용 흔적 분석은 Cloud Service Provider에 의존적인 특징이 있지만 Temporary Internet Files, History files, Cookie files, 물리메모리 등의 웹 브라우저 흔적은 디지털 포렌식 분석에서 중요한 분석 요소임을 실험을 통해 확인했다.

기존 웹 포렌식 분석에서는 History files의 웹사이트 방문기록, 검색어 등을 중점으로 분석하는 경향이 많았다. 하지만 본 논문에서 제시한바와 같이 Temporary Internet Files와 물리메모리에도 수사의 결정적인 증거를 제공할 여지가 있으며, 분석시 반드시 고려해야하는 필수요소이다. 또한, 최근 클라우드 서비스의 대중화로 인해, 용의자 PC에서 클라우드 서비스를 사용했는지의 여부를 판별하는 것은 상당한 의미가 있다.

따라서 향후 디지털 기기를 조사할 때에는 반드시 본 논문에서 제시한 흔적에 기반하여 클라우드 서비스를 사용했는지 확인할 필요가 있다.

참 고 문 헌

[1] NIST, Cloud computing, URL: http://csrc.nist.gov/publications/drafts/800-145 /Draft-SP-800-145_cloud-definition.pdf, 2011.

[2] 이승봉, 권혁돈, 임경수, 이상진, "웹 브라우저 사용정보 분석을 위한 도구 설계 및 구현", 한국디지털포렌식학회, pp.18-19, 2008.

[3] Roussev R, "A cloud computing platform for large-scale forensic computing", Advances in Digital Forensics, pp.201-215, 2009.

[4] Keyun Ruan, "Cloud forensics: An overview", Vol.14, No.1, 2010.

[5] Dominik Brik, "Techical Challenges of Forensic Investigations in Clould Com puting Environments", pp.1-3, 2011.

[6] M. Taylor, J. Haggerty, D. Gresty, R. Hegarty, "Digital evidence in cloud computing systems", Digital Investigation, pp.2-3, 2010.

[7] Junghoon Oh, Seungbong Lee, Sangjin Lee, "Advanced Evidence Collection and Analysis of Web Browser Activity", Digital Investigation, pp.2-9, 2010.

[8] SaaS(Software_as_a_service), URL : http://en.wikipedia.org/wiki/Software_as_a_service

강 성 림

e-mail : kangdduk@korea.ac.kr
2010년 2월 강원대학교 컴퓨터과학과
2010년 3월~현 재 고려대학교 정보보호 대학원 석사과정
관심분야 : 디지털 포렌식, 정보보호

박 정 흠

e-mail : junghmi@korea.ac.kr
2007년 2월 한양대학교 정보통신대학 컴퓨터전공(공학사)
2007년 3월~2009년 2월 고려대학교 정보경영공학전문대학원(공학석사)
2009년 3월~현 재 고려대학교 정보경영공학전문대학원 박사과정
관심분야 : 디지털 포렌식, 안티-안티 포렌식

이 상 진

e-mail : sangjin@korea.ac.kr
1987년 2월 고려대학교 수학과
1989년 2월 고려대학교 수학과(이학석사)
1994년 8월 고려대학교 수학과(이학박사)
1989년 10월~1999년 2월 ETRI 선임연구원
1999년 3월~2001년 8월 고려대학교 자연과학대학 조교수
2001년 9월~현 재 고려대학교 정보경영공학전문대학원 교수
관심분야 : 디지털 포렌식, 모바일 포렌식, 심층 암호, 해쉬 함수