情 報 保 護 學 會 誌
第21卷 第6號, 2011. 10

특수유형 OSP(Online Service Provider)의 대용량 데이터베이스 포렌식 분석 방안 연구

이 동 영*, 전 완 근**, 김 홍 윤***

요 약

다수의 사람과 파일을 공유할 수 있는 웹하드 서비스의 이점을 이용하여 각종 불법복제물 등의 업로드를 유도하고 다운로드를 통해 이득을 취하는 특수유형의 OSP(Online Service Provider, 온라인 서비스 제공자)들이 출현하게 되었다. 이런 범죄가 일어나는 업체의 데이터베이스에는 모든 이용자들의 관련 기록을 담고 있어 헤비업로더의 활동내역뿐만 아니라 업체측의 방조혐의 등의 증거를 추출할 수 있다. 본 논문에서는 특수유형 OSP들의 대용량 데이터베이스를 신속하고 정확하게 무결성을 유지하며 데이터베이스의 데이터를 수집할 수 있는 방법에 대해 연구해보고, 수집한 데이터 또한 신속하게 분석하는 방법을 제안하였다.

Ⅰ. 서 론

인터넷과 스토리지의 발달로 일정한 용량의 저장공간인 스토리지를 확보해 이동형 저장매체 없이도 어느 곳에서나 인터넷을 통해 자신이 작업한 문서나 파일을 저장·열람·편집하고, 다수의 사람과 파일을 공유할 수 있는 웹하드 서비스가 생겨나면서 현재 이런 종류의 서비스는 여러 업체에서 제공하고 있다.

파일 올리고 내리기, 파일 및 폴더의 생성·변경·이동·삭제·복사, 메모장 작성, 간편한 자동 백업 따위의 다양하고 편리한 기능을 제공해 많은 가입자를 확보하고 있다. 자유롭고 편리한 파일 공유·전송·저장, 저장매체의 파손·분실·도난방지, 파일 전달 기능을 활용한 공동 연구 및 팀·부서 간의 공동업무 수행, 안전한 데이터 백업 및 복구, 저렴한 비용과 인터넷 연결만으로 즉시 제공되는 서비스, 각종 보안장치를 통한 외부의 불법접근 차단, 전세계 어디서나 이용 가능한 서비스 제공, 대용량 자료의 빠르고 정확한 전달 등이다.

서비스 대상은 보통 개인, 사업자·법인, 독자적 웹하드운용을 필요로 하는 개인·단체·기업 등으로 구분된다. 운용 용량·속도, 서비스 방식 등에 따라 가입비가 다른데, 가입 비용은 저렴한 편이다.

하지만 이러한 이점을 이용하여 각종 불법복제물 등의 업로드를 유도하고 다운로드를 통해 이득을 얻는 특수유형의 OSP들이 우후죽순 생겨나고 있다. 업로더들에게는 업로드한 컨텐츠의 다운로드 횟수만큼 보상금을 지급하고, 업로드된 컨탠츠에 대해서는 불법복제물 식별 여부, 검색어 필터링 등 최소한의 기술적보호조치를 취하지 않음으로써 저작권침해가 심각해지고 있다.

이런 업체들을 조사하기 위해서는 웹과 연동한 데이터베이스를 수집 및 분석해야 하지만 이들의 데이터베이스는 회원정보, 업로드 정보, 다운로드 정보, 결제정보, 보상정보, 게시물 정보 등 대용량의 데이터를 담고 있다. 그러다보니 수많은 테이블과 컬럼의 구조를 파악하고 방조 및 범죄수익금 산출에 필요한 데이터를 분석하는데 많은 시간이 소요된다.

Ⅱ. DBMS 유형

데이터베이스는 그 내용을 쉽게 접근하여 처리하고 갱신할 수 있도록 구성된 데이터의 집합체이다. 가장 널리 보급된 데이터베이스는 데이터를 다양한 방법으로

* 한국저작권위원회 디지털정보보호팀(dylee@copyright.or.kr)
** 대검찰청 디지털포렌식센터 검찰수사관(iwisky@spo.go.kr)
*** 한서대학교 대학원 디지털포렌식학과 교수(hykim@hanseo.ac.kr)

〔표 1〕 DBMS 종류 및 사용 예

종 류	사용 예
RDBMS	MS-SQL, MY-SQL, ORACLE, DB2
OODBMS	Objectivity, O2,Versant, Ontos, Gemstone
ORDBMS	Unisql, Object Store
HDBMS	IBM IMS DB
NDBMS	CODASYL DB

접근하고 재구성할 수 있도록 정의한 테이블형의 데이터베이스인 관계형 데이터베이스이다. 분산 데이타베이스는 네트워크상의 여러 다른 지점에 분산되어 있거나 중복되어 있는 데이터베이스를 말하며, 객체지향 데이터베이스는 객체 클래스와 서브 클래스로 정의된 데이터가 서로 일치하는 데이터베이스이다.

DBMS(Database Management System)는 때로는 데이터베이스 관리 시스템이라고도 불리는데, 다수의 컴퓨터 사용자들이 데이터베이스 안에 데이터를 기록하거나 접근할 수 있도록 해주는 프로그램이다. DBMS는 사용자 요구사항들이나 다른 프로그램의 요구사항들을 관리함으로써, 사용자들이나 다른 프로그램들이 실제로 그 데이터가 저장매체의 어느 곳에 저장되어 있는지를 이해하지 않고서도, 다중 사용자환경의 그 누구라도 데이터를 이용할 수 있도록 해준다.

사용자 요구사항들을 처리함에 있어, DBMS는 데이터의 무결성 (이것은 데이터베이스가 계속해서 접근이 가능하며, 또한 의도한대로 조직화되어 있다는 사실을 확인해주는 것이다)과 오직 허가된 사용자들만이 데이터에 접근할 수 있게 하는 보안성을 보장한다. 가장 일반적인 형태의 DBMS가 관계형 데이터베이스 관리시스템, 즉 RDBMS 이다. RDBMS의 표준화된 사용자 및 프로그램 인터페이스를 SQL이라고 부른다. 좀더 새로운 종류의 DBMS로 OODBMS가 있다.

Ⅲ. DB 증거수집

OSP업체들을 대상으로 압수수색시, 업체에서 관리하는 데이터베이스 관리 시스템(이하 DBMS라 칭함)은 가입회원 정보는 물론 회원들이 업로드, 다운로드, 현금전환한 활동 정보들이 담겨있다. 때문에 가장 중요한 증거가 되며 이를 수집해야한다.

업체들마다 사용하는 DBMS가 다른 점을 고려하여 압수하는 자는 사전에 DBMS별 관련된 기술정보를 충분히 숙지할 필요가 있다.

3.1 증거수집대상

저작권법 위반, 불법복제물 및 음란물 유통 등의 혐의, 헤비업로더들의 활동 정보 등이 수사대상이라면 이와 관련된 정보를 반드시 수집해야 할 것이다.

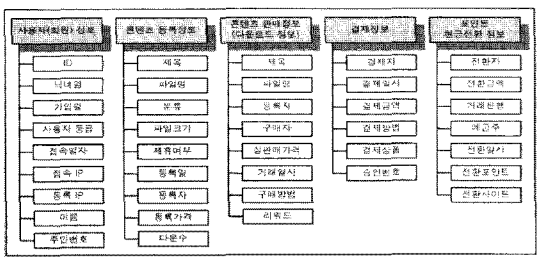
업로드되어 공유되고 있는 콘텐츠가 현재 정상적으로 보호를 받고 있는지, 보호요청되지 않은 콘텐츠들의 불법판매 등으로 취한 수익금, 불법복제물 및 음란물을 통해 이득을 취하는 업로더 목록과 각 업로더들의 신상정보 및 이득을 취한 금액 등 이와 관련된 정보가 담긴 데이터베이스가 대상이 된다.

필요한 데이터들이 한 데이터베이스에만 존재하는 것이 아니며, 한 테이블에만 존재하는 것이 아니기 때문에 가장 기본이 되는 데이터베이스 및 테이블을 기준으로 연관관계를 조사하여 연관된 데이터베이스 및 테이블을 찾아 관련된 데이터를 수집해야 한다.

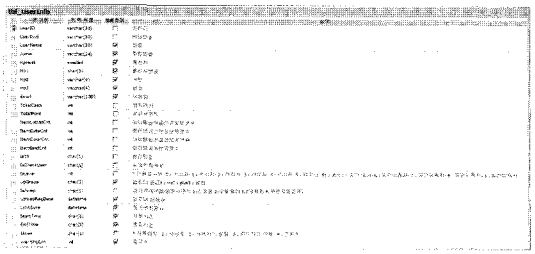
특히 헤비업로더들을 대상으로 데이터를 수집할 경우, 헤비업로더들을 특정지을 수 있는 정보가 반드시 필요하다. 예를 들어 이름, 주민등록번호, 전화번호, 집주소, 이메일주소, IP주소 등이 업로더를 특정지을 수 있는 정보이다.

3.2 DBMS 관련정보 수집

업체에서 사용하고 있는 DBMS에 관련된 정보를 가능한 한 많이 수집해야 한다. 운영되고 있는 데이터베이스의 서버 IP주소, DBMS의 종류, 데이터 저장 유형, 캐릭터셋, 데이터백업여부 등을 조사하여 현장에서 추출할 수 없는 자료를 전용분석공간에 별도로 환경을 구축하여 직접 추출할 수 있도록 필요한 정보들을 반드시 수집한다.



〔그림 1〕 DB 증거수집대상



〔그림 2〕 테이블 명세서

DBMS는 Windows, Linux 등 다양한 운영체제 환경하에서 구동되며, Microsoft사의 MS SQL Server, 공개형인 MySQL, IBM의 DB2 등 종류가 매우 다양하다.

현재 업체에서 운영하고 있는 서버의 타임스탬프, IP 주소, 운영체제 종류, 버전 등을 수집한다. DBMS의 경우에는 DBMS의 종류, 구축환경 등의 정보를 수집하고, 현재 생성되어 운영되고 있는 데이터베이스명, 그 안의 테이블명, 테이블개수, 테이블간의 연관관계 등의 정보를 수집한다. 업체마다 운영하는 테이블의 개수는 천차만별이지만 보통 수백개의 테이블을 생성하여 사용하고 있다. 이중에는 테스트용으로 만들어 진 것도 있지만 혐의를 입증할 수 있는 중요한 데이터들을 암시 및 저장하고 있기 때문에 매우 중요하다. 대부분의 업체에서는 통상적으로 관리목적하에 테이블의 구조 및 역할을 설명해놓는 테이블명세서를 작성해놓고 업무에 참고한다. 이런 자료를 수집한다면 추후 데이터를 추출하는데 많은 도움을 줄 뿐 아니라 확실한 혐의를 잡는데 도움이 된다.

각 테이블의 행수, 크기, 스키마 등도 같이 수집하는데, 이유는 수집한 증거에 대한 정보를 피의자와 같이 확인하는 절차를 거쳐야하고, 또한 추후에 분석을 위해 복구 및 임포트할 때 참고할 수 있기 때문이다.

3.3 로그 데이터 수집

로그 데이터를 수집하는 방법에는 DBMS 자체 백업기능을 이용한 수집, RAW형태의 파일 수집, 필요한 정보만 수집, 기존 백업파일 수집 등 크게 4가지 방법을 고려할 수 있는데, 각 방법에는 장단점이 존재한다.

우선 첫 번째 방법인 DBMS 자체 백업기능을 이용한 데이터 수집은 예를 들어 MS SQL Server는 자체 백업기능을 이용하면 쉽고 빠르게 백업파일을 생성하여 수집할 수 있고, 복구하기에도 쉽고 빠른 장점이 있다. 하지만 모든 DBMS가 자체백업기능을 제공하는 것이 아니며, 최적화되어 있지 않는 DBMS의 경우 속도가 매우 느릴 수 있다. 예를 들어 MySQL의 경우 mysqldump 명령으로 백업이 가능하지만 실제 생성되는 파일 내용을 들여다보면 데이터 생성 및 입력 등의 스크립트 형태의 텍스트파일이다. 이 형태는 속도가 매우 떨어져 신속한 증거 수집이 불가능하다.

〔표 2〕 백업기능 및 RAW파일 수집시간 비교

수집방법	수집시간	비고
mysqldump	7분 20초	*4.56GB 기준
Raw파일 복사	6분 18초	

두 번째 방법은 DBMS의 RAW파일을 그대로 수집하는 방법이다. 이 방법은 특정 DBMS에서는 예외일 수 있지만 보통 DBMS에서는 전체 데이터를 가져오는 가장 빠른 방법이 될 수 있다. [표 2]는 MySQL의 mysqldump기능과 RAW파일을 복사 및 압축하는 시간을 비교실험한 결과다.

MS SQL Server의 경우 데이터를 .mdf, .ldf 2개의 파일로 관리한다. 본래 이 방법은 디스크 물리적 공간의 경로를 변경하고자 할 때 사용하는데 증거수집에 이용할 수 있다. 먼저 파일의 위치를 특정 명령을 통해 알아내고 그 파일들을 복사하는데, 이 파일들은 현재 DBMS에 묶여 구동되고 있어 개념상 이 파일들을 DBMS에서 분리시킨 후 복사해야 한다. 이 경우 서비스를 중지시켜야하는 단점이 있다.

- 'sp_helpdb 데이터베이스명' 명령으로 파일 위치 파악
- 'sp_detach_db 데이터베이스명' 명령으로 파일을 DBMS에서 분리
- .mdf, .ldf 파일에 대한 해쉬값 계산 및 복사

〔표 3〕 MS SQL Server 및 MySQL RAW파일 경로

DBMS	파일위치
MS SQL Server	C:\Program Files\Microsoft SQL Server\MSSQL.1\MSSQL\DATA\데이터베이스명.mdf C:\Program Files\Microsoft SQL Server\MSSQL.1\MSSQL\DATA\데이터베이스명.ldf
MySQL	Linux : /user/local/var /user/local/mysql/data Windows : C:\Documents and Settings\All Users\Application Data\MySQL\MySQL Server 5.x\data\데이터베이스명

Mysql의 경우 InnoDB, Myisam 2가지 형태로 데이터가 관리되는데, InnoDB의 경우 ibdata1, ibdata2의 파일에 데이터가 저장되고, 테이블 구조는 확장자가 .frm에 저장되는 구조다. Myisam의 경우는 .frm, .myd, .myi 3개의 파일로 구성되어 데이터들을 저장한다. 각 파일명은 데이터베이스내에 생성되어 있는 테이블명으로 되어 있다.

세 번째 방법으로는, 필요한 데이터만을 추출하는 것인데 특정 테이블내의 특정정보를 추출하는 SQL쿼리를 통해 필요한 데이터를 추출할 경우 해당 테이블의 행과 컬럼 수가 많고, 특수유형 OSP 특성상 특정시간대에 이용자들이 몰려 서버부하가 심해지면, 서비스가 다운될 뿐만 아니라 자료를 추출하는데도 매우 많은 시간이 소요된다. 이 방법은 피의자측에 요청하여 직접 필요한 데이터를 작성하도록 했을 경우 신빙성을 가지고 있는 반면에 지정했던 테이블 이외의 테이블에 예상치 못했던 혐의증거들이 담겨있음에도 불구하고 그것을 놓칠 수 밖에 없게 된다.

마지막으로, 업체는 서비스를 운영하면서 각종 로그기록을 남기는데, 고객지원을 위해 업체측에서도 중요하게 여기고 있어 주기별로 로그데이터백업을 대부분 수행한다. 주기적으로 정해진 시간에 백업을 수행토록 하는 스크립트 또는 백업도구의 옵션 등의 정보를 수집하고 백업된 데이터 존재여부를 파악하여 수집한다. 이 백업데이터가 중요한 이유는 압수수색을 시작하고 현재 구동되고 있는 데이터베이스의 데이터를 간단한 스크립트 또는 SQL쿼리문을 통해 삭제 또는 수정을 할 가능성이 크지만 백업된 데이터는 특정 포맷으로 저장 또는 압축되어 있어 변형이 쉽지 않기 때문이다.

3.4 DB 수집물 동일성 확인

디지털 포렌식에서 가장 중요한 것은 증거에 대한 무결성 및 동일성을 보장하는 것이다. 일반 디스크 포렌식의 경우엔 하드디스크와 같은 저장매체에 물리적으로 쓰기방지장치를 이용하는 등 데이터 무결성을 유지시킬 수 있다. 하지만 데이터베이스의 데이터는 하드웨어가 아니기 때문에 물리적인 조치를 취할 수가 없다.

이를 해결하기 위한 방법으로는 위에서 설명한 수집방법을 통해 얻은 파일들을 양측이 모두 소지하여야 하고 이 파일들의 해쉬값 증명서를 서로 교환하여 동일성을 인정하도록 하여 추후 관련문제가 발생되지 않도록 한다. 이 절차에서 반드시 피내사자 또는 제 3자(DB관리자)와 함께 데이터베이스에 접근하여 데이터 추출과정에 문제가 없음을 확인시켜준다.

Ⅳ. DB 증거분석 방안

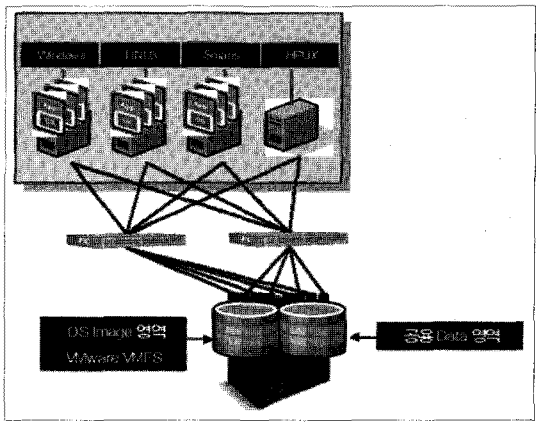
현장에서 수집한 데이터를 분석하기 위해서는 현장환경과 동일하거나 또는 유사하게 분석환경을 구축하여 정상적으로 데이터가 복구되도록 해야 한다. 그러기 위해서 현장에서 수집한 DBMS관련 정보(DBMS 종류, 캐릭터셋 등)을 이용하여 분석환경을 구축하여 분석자가 추출한 데이터가 증거효력을 잃지 않도록 한다.

4.1 분석환경 구축

수집해온 데이터를 분석하기 위해서는 복구 및 임포트를 해주어야 한다. 현장 환경과 유사하게 구축하기 위해서 수집당시 획득한 DB 정보를 참고한다. 이때 분석입장에서는 수집한 데이터 분석을 위한 분석환경을 구축하기 위해서는 다양한 환경이 가능한 유연성있는 시스템이 필요하다. Windows 환경의 MS SQL Server, MySQL, Oracle 등이 있을 수 있고, Linux 환경의 MySQL, Informix 등이 있을 수 있는데 그때마다 컴퓨터를 구해 구축할 수도 없는 노릇이다. 간단한 해결방안으로 제시할 수 있는 것이 하나의 서버에 여러 개의 가상환경을 구축할 수 있는 가상화 시스템이다.

4.2 테이블 분석

특수유형 OSP는 회원정보, 결제정보, 게시물정보, 포



〔그림 3〕 가상화 시스템 구상도

인트정보 등 수많은 데이터들을 관리해야 하기 때문에 이를 관리하기 위한 테이블 또한 수십개에서 수백개로 구성되어 있다. 많은 테이블들을 신속하고 정확하게 분석하여 어떤 데이터들을 담고 있는지를 파악하여 수사에 필요한 데이터들을 빠른 시일내에 추출할 수 있어야 한다. 그러기 위해서는 테이블명을 통한 추측과 실제 저장되어 있는 데이터들을 하나씩 확인하여야 한다.

4.2.1 테이블명을 통한 데이터 분석

테이블명을 정하는 것은 개발자의 관점에 따라 달라지지만 일반적으로 자신뿐만 아니라 다른 개발자들도 쉽게 알아볼 수 있도록 하는 것이 관례이기 때문에 특정 패턴이 존재한다. 그래서 테이블명을 보고 어떤 데이터가 담겨있는지 어느 정도 파악이 가능하다. 하지만 이를 무시하고 철저히 자신만이 알아볼 수 있는 이름으로 생성한 경우에는 파악하기 쉽지 않다. 이러한 경우 테이블안의 컬럼명과 함께 데이터를 분석해야 한다.

또한, 저장된 값이 어떤 의미를 가지고 있는지도 파악해야 한다. 파일용량의 경우 byte, kilo byte, mega byte, giga byte, tera byte 등으로 구분되는데 저장된 데이터가 어떤 단위로 저장되어 있는지 확인해야 한다. 또, 플래그 형태의 값인 0 또는 1, Y 또는 N 등이 어떤 값이 참이고 거짓인지도 확인해야 한다.

4개의 업체 샘플을 비교분석한 결과 [표 4]와 같은 결과를 얻을 수 있었다. 회원들의 신상정보를 관리하는 테이블명은 사용자(user), 회원(member)를 뜻하는 단어를 사용하였고, 게시물을 관리하는 테이블명은 파일(file)이나 게시판(bbs)를 의미하는 단어를, 결제정보를 관리하는 테이블명은 지불(pay)을 의미하는 단어를, 포인트 정보를 관리하는 테이블은 포인트(point), 게임머니와 개념을 동일시한 현금(cash), 현금처럼 쓸 수 있는 포인트로 변환해주는 서비스인 포인트뱅킹을 의미하는 point banking으로 각 테이블을 식별할 수 있도록 해놓았다.

〔표 4〕 샘플별 테이블명명 비교

비교샘플	A	B	C	D
회원정보 테이블명	user_info_table	tbl_UserInfo	_member	user_db
게시물 관리 테이블명	club_file_table	tbl_Content	_webhard_file_bbs	file_db
결제정보 테이블명	mono_payment_tbl	tbl_AccountResult	_account	payment_db
포인트 및 현금전환 정보 테이블명	cash_info_table, point_banking_req	tbl_UserPayment	_point_log	point_log_200601

4.2.2 테이블 연관관계 분석

데이터베이스는 테이블들간의 연관관계가 존재한다. 테이블의 관계를 안다면 조사자가 필요로 하는 정보를 쉽게 분석할 수 있다. 사용자 정보를 담고 있는 테이블이 어떤 테이블과 연관되어져 있는가를 볼 수 있다면 관련 테이블을 쉽게 파악할 수 있게 된다.

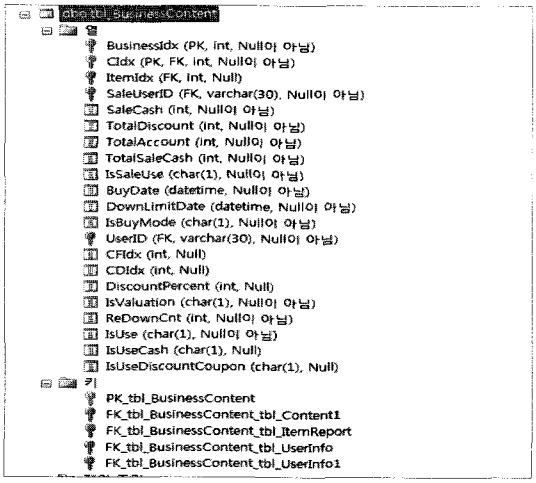
첫 번째, 하나의 테이블과 또 다른 하나의 테이블간의 데이터 관계가 있다. 사용자 정보 테이블과 결제 정보 테이블과의 관계는 한 사용자가 여러번 결제를 할 수 있기 때문에 일대다로 볼 수 있다. 게시물 목록 테이블과 게시물에 관한 정보가 담긴 테이블은 업로드된 파일은 유일하게 하나이므로 일대일 관계가 성립된다.

두 번째, 하나의 테이블과 다른 여러 테이블간의 관계가 있을 수 있다. 사용자 정보 테이블이 있다면 그 중 일반적으로 고유한 값으로 다루어지는 사용자의 고유번호나 고유ID가 다른 여러 테이블에서 사용될 수 있다. 사용자 정보 테이블의 사용자ID 값이 사용자의 현재 보유 포인트 정보가 담긴 테이블에서 사용될 수 있으며, 또한 기타 관련 정보가 담긴 테이블에서 사용되어진다. 회원이 탈퇴할 때 사용자 정보 테이블에서 삭제될 경우 관련 정보들도 같이 삭제되므로 운영상 유용하다.

두 번째 방법은 기본키/외래키를 통해 분석이 가능한데 MySQL의 경우 'show create table 테이블명' 명령을 통해 Primary Key와 Foreign Key를 통해 연관관계를 알 수 있고, MS SQL Server에서도 같은 방법으로 'show create table 테이블명' 명령을 이용해 외래키 설정을 확인할 수 있다. 오히려 MS-SQL Server는 전용툴에서 [그림 3]과 같이 테이블에 설정된 외래키를 쉽게 확인해볼 수 있다.

4.3 범죄혐의 증거 분석

업체에서 벌인 범죄행위는 로그기록만을 남기는 데이터베이스에서도 추출이 가능하다. 범인들은 항상 혼



〔그림 4〕 MS SQL Server 연관관계 정보

적을 남기기 때문에 같이 압수한 운영소스코드가 존재한다면, 소스코드를 분석하여 소스코드단에서 데이터베이스에 기록을 남기는 방법과 다르게 기록을 남기는 부분을 집중적으로 분석한다,

4.3.1 헤비업로더 분석

헤비업로더란 단순히 파일 공유 차원이 아닌 전문적으로 컨텐츠를 대량으로 업로드하여 이익을 취하기 위한 목적의 사용자들이다. 이들은 업로드를 권장하는 업체에 옮겨다니며 자신들이 업로드한 콘텐츠가 판매되면 받게 되는 보상포인트를 이용해 현금으로 받거나 특정 쇼핑몰에서 원하는 물품을 구매할 수 있다. 이뿐만 아니라 '해운대' 유출사건에서 엿볼 수 있듯이 개봉전의 영화가 유출되었을 경우 누가 최초 유표자인지 밝혀낼 수도 있다. 이러한 범죄행위를 밝혀내기 위해 업로더별 상세정보를 추출해야한다.

그 중에는 헤비업로더의 신상정보, 거래은행정보, 업로드한 콘텐츠 내역과 판매내역 등이 있고, 얼마나 보상을 받아 이득을 취하였는지에 대한 정보가 있다.

4.3.2 게시물 분석

사실상 특수유형 OSP에서 공유되고 있는 콘텐츠는 모두 저작권자를 보유하고 있지만 현재는 저작권 보호요청되어 있는 콘텐츠만 보호를 받고 있는 실정이다.

보호요청되지 않은 콘텐츠의 판매를 통해 이득을 취하는 것, 보호요청된 콘텐츠도 업로더들의 다양한 필터링우회방법으로 보호받지 못하는 콘텐츠의 판매수익금들이 모두 불법으로 간주된다. 게시물 분석을 통해 헤비업로더들의 업로드한 콘텐츠 내역을 추출하고, 실제 존재여부, 판매여부 등을 알 수 있다.

4.3.3 현금전환 내역 분석

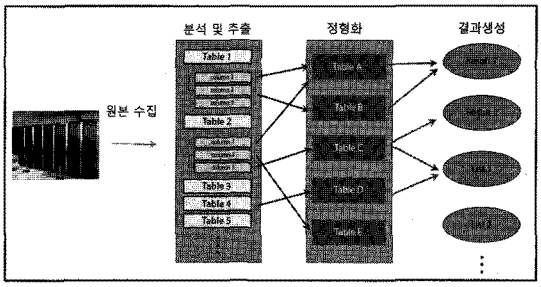
전문적인 업로더들은 자신들이 보유하고 있는 대량의 컨텐츠를 특수유형 OSP에서 공유함으로써 이득을 취하고 있다. 문제는 이들이 공유하는 컨텐츠 대부분이 저작권 침해가 우려되는 불법복제물이며 이뿐만 아니라 국내외의 음란물들을 무단 유통시키고 있다. 음란물의 경우 국내법상 저작권이 보호되지 않지만 정보통신망법에 의해 음란물유포죄에 해당되며 이들을 방관한 업체들 또한 방조혐의가 씌워지게 된다.

업체는 업로더들에게 보상금을 지급하기 위해 올앳, 띠앗 등의 포인트 적립회사와 연동하여 정책에 따라 일정 포인트 단위로 현금을 지급해주고 있다.

V. DB 분석 모델 연구

앞서 소개된 추출해야할 데이터(헤비업로더, 불법콘텐츠, 현금전환 기록 등)들은 수백개의 테이블과 수만라인의 소스코드를 분석함으로써 얻어낼 수 있는 결과이다. 이러한 방법은 매우 많은 시간을 소요하게 되어 있는데 이번 분석모델연구에서는 이러한 과정을 단순화시킴으로써 추후에 일부 자동화가 가능하도록 하는데 목적이 있다.

본 연구에서 제안하는 분석 모델은 현장에서 DB에 대한 증거분석시 분석 결과에 대해 객관적인 결과를 확보하여 도출 결과에 대한 신뢰성을 높이고자 한다. 다음



〔그림 5〕 분석 모델

그림은 분석 모델의 기본 구조를 보여준다.

분석 모델의 원리는 먼저 수집된 증거 데이터에 대한 정형화 작업이 필요하다. 이후 정형화 테이블을 활용하여 원하는 정보를 생성할 수 있다. 즉, 다양한 구조를 가진 OSP의 DB라 하더라도 정형화된 테이블에 데이터를 채울 수 있다면 그 결과는 객관적이고 신뢰성 높은 결과라 할 수 있다.

이 분석모델의 장점은 신속한 분석 결과를 도출할 수 있게 도와준다. 데이터베이스 내 테이블은 분석자가 직접 수동으로 구분지어야 하지만 수동단계를 거쳐 만들어낸 정형화된 테이블로 원하는 결과를 추출할 수 있는 자동화 시스템을 구축할 수 있다.

정형화된 테이블 안의 데이터들은 분석자에 의해 일정 기준으로 데이터가 일관성있게 구성되어 있어 스크립트나 분석 도구를 통해 원하는 정보를 추출할 수 있는데, 예를 들어 헤비업로더의 정보를 추출하려고 할 때 헤비업로더의 신상정보를 기본으로 신상정보에 보상포인트 정보(신상정보+보상포인트내역)를 추가하여 추출, 또는 헤비업로더의 결제내역(신상정보+보상포인트내역+결제내역)을 추가하여 추출하는 등 여러 결과를 만들어 낼 수 있게 된다.

Ⅵ. 결 론

특수유형 OSP의 대용량 데이터베이스의 수집과 분석과정에는 많은 어려움이 있다. 수집과정에서는 서비스의 피해를 최소화하여야 할 뿐 아니라 가능한 한 신속한 증거 데이터 수집이 이루어져야 한다. 방조 및 범죄수익금 분석을 위한 데이터를 추출하기 위해 현장에서 관련자료도 수집하여야 한다. 관련자료에는 웹소스코드, 매출 장부, 기타 부수적 파일 등이 필요하다. 분석과정에서는 현장과 유사한 환경을 구축하여 데이터를 임포트시켜야 하며, 임포트 과정에서 정상적으로 데이터가 저장되어야 한다. 적게는 수십개, 많게는 수백개의 테이블을 명명된 테이블명으로 구분지어야 하며 테이블 안의 컬럼 또한 명명된 컬럼명으로 구분짓고, 또한 각 값의 의미 또한 파악하여야 한다. 테이블의 연관관계를 분석하여 전체적인 구조를 파악한 후 최적화된 SQL 질의문 또한 작성해야 한다. 이런 문제점들을 본 논문에서 다룸으로써 신속한 증거 파일 수집 및 분석이 이루어질 수 있기를 기대한다.

참고문헌

[1] 김병준 "MySQL 데이터베이스 최적화," IT BRIDGE, 2006.

[2] Martin S Olivier "On Metadata Context in Database Forensics," ICSA Research Group, 2008.

[3] Kyriacos Pavlou and RichardT.Snodgrass "Forensic Analysis of Database Tampering," University of Arizona, 2008.

[4] "SQL Server Database Forensics," Black Hat USA 2007.

[5] Baron Schwartz, Peter Zaitsev, Vadim Tkachenko, Jeremy D.Zawodny, Arjen Lentz & Derek J. Balling "High Performance MySQL Second Edition," OREILLY, 2008.

[6] 박장규 "MySQL 5.1 Reference Guide : 처음부터 끝까지", 혜지원, 2009.

[7] 폴 드보이 저/김형훈 역 "MySQL : 한국어판, MySQL의 사용, 관리, 프로그래밍을 위한 완벽 가이드 (4판)," 지앤선, 2009

[8] 데이터베이스 사랑넷 "http://database.sarang.net/

[9] Microsoft Technet "http://technet.microsoft.com/ko-kr/library/ms175195.aspx"

[10] 마이크로소프트 개발자 네트워크 "http://msdn.microsoft.com/ko-kr/library/ms190969.aspx"

[11] 웹마스터 웹디자인 리더 "http://www.web-reader.co.kr/sql/sql-2.htm"

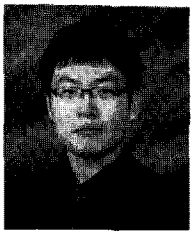
[12] Microsoft 고객지원 "http://support.microsoft.com/kb/601427/ko"

[13] IBM "http://publib.boulder.ibm.com/infocenter/tpfhelp/current/index.jsp?topic=/com.ibm.ztpf-ztpfdf.doc_put.cur/gtpm7/m7dirstruct.html"

[14] MySQL Korea "http://www.mysqlkorea.co.kr/"

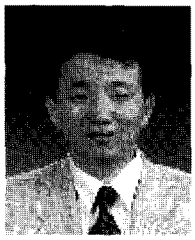
[15] KLDP Wiki "http://wiki.kldp.org/wiki.php/MySQL%B8%AE%C7%C3%B8%AE%C4%C9%C0%CC%BC%C7"

〈著者紹介〉



이 동 영 (Dong-Young Lee)
정회원
2009년 2월 : 한서대학교 컴퓨터통신공학과 졸업
2011년 2월 : 한서대학교 디지털포렌식학과 석사 졸업
2010년 2월 : 한국인터넷진흥원 인터넷침해대응센터 연구원
2010년 3월~현재 : 한국저작권위원회 디지털정보보호팀 주임
<관심분야> 디지털포렌식, 정보보호

전 완 근 (Wan-Keun Jeon)
종신회원
2005년 2월 : 한서대학교 정보보호공학과 박사 졸업
2000년 3월 ~ 2006년 10월 : 한국인터넷진흥원 선임연구원
2004년 1월 ~ 2006년 1월 : 서울중앙지검 첨단범죄수사부 파견
2006년 11월 ~ 현재 : 대검찰청 디지털포렌식센터 근무
2010년 1월 ~ 현재 : 한국저작권위원회 저작권포렌식 자문위원
2009년 ~ 현재 : UNODC VFAC 국제사이버범죄방지포럼 전문가
<관심분야> 디지털포렌식 2.0, 사이버범죄, 네트워크포렌식, 금융포렌식



김 홍 윤 (Hong-Yun Kim)
정회원
1982년 2월 : 인하대학교 전자계산학과 (이학사)
1984년 2월 : 인하대학교 전자계산학과 (이학석사)
1996년 2월 : 인하대학교 전자계산학과 (이학박사)
1995년 3월 ~ 현재 : 한서대학교 컴퓨터공학전공 교수
<관심분야> 센서 네트워크, 디지털 포렌식