소프트웨어 몽타주: 디지털 포렌식 수사를 위한 유사 소프트웨어 탐지 대상의 필터링

(Software Montage: Filtering of Detecting Target of Similar Software for Digital Forensic Investigation)

박 희 완 † 한 태 숙 ††
(Heewan Park) (Taisook Han)

요 약 소프트웨어 몽타주란 소프트웨어로부터 빠르게 추출 가능하고 내재된 특성을 함축하고 있는 정보를 의미한다. 잘 알려진 프로그램으로부터 몽타주를 작성하면 몽타주를 기반으로 유사 프로그램 탐지 대상을 필터링할 수 있다. 본 논문에서는 API 호출과 문자열 기반의 소프트웨어 몽타주를 제안한다. 제안된 몽타주를 평가하기 위해서 인스턴트 메신저 프로그램에 대한 유사 프로그램 탐지 대상의 필터링 실험을 하였다. 이 실험으로부터 제안된 몽타주가 잘 알려지지 않은 프로그램 탐지 대상을 필터링하는 포렌식 도구로 활용될 수 있다는 것을 확인하였다.

키워드 : 소프트웨어 몽타주, 유사 소프트웨어 필터링, 소프트웨어 포렌식, 디지털 포렌식

Abstract A software montage means information that can be extracted quickly from software and includes inherent characteristics. If a montage is made from well-known programs, we can filter candidates of similar programs among the group of programs based on the montage. In this paper, we suggest software montages based on two characteristics: API calls and strings. To evaluate the proposed montages, we performed experiments to filter candidates of some similar programs to instant messenger programs. From the experiments, we confirmed that the proposed montages can be used as a forensic tool that filters a group of similar programs even when their functions are not known in advance.

Key words : Software Montage, Similar Software Filtering, Software Forensics, Digital Forensics

1. 서 론

디지털 포렌식(Digital Forensics)이란 디지털 매체로부터 증거를 수집하고 분석하는 기술을 연구하는 학문이다[1]. EnCase를 비롯한 기존 포렌식 도구들은 대부분 분석 대상이 문서 파일이나 이메일, 이미지 파일과 같은 데이터에 국한된다는 문제가 있다[2].

소프트웨어 포렌식은 범죄와 연관될 수 있는 소프트웨어에 대한 포렌식을 의미한다. 대표적인 예로서 인스턴트 메신저는 최근 사이버 범죄의 도구로 빈번하게 사용되고 있기 때문에 이에 대한 포렌식 연구가 활발히 진행되고 있다[3].

특정 소프트웨어가 컴퓨터에 설치되었는지를 확인하기 위해서는 윈도우 레지스트리 정보나 참조 데이터 집합(Reference Data Set)을 활용할 수도 있다[4]. 대표적인 참조 데이터 집합인 NSRL(National Software Reference Library)은 알려진 파일의 해시값을 공개하여 디지털 증거 분석시 검색 시간을 줄이고 위변조 여부를 판단하는데 사용될 수 있다[5]. 그러나 NSRL에 등록되지 않은 최신 프로그램은 탐지할 수 없다는 문제가 있고, 한 비트만 달라져도 해시값이 바뀌기 때문에 프로그램의 유사성 여부를 판단하는 데에는 사용될 수 없다.

본 논문에서는 이미 잘 알려진 프로그램으로부터 공통적인 특징을 추출한 후, 알려지지 않았지만 유사한 기능을 가지고 있는 프로그램을 탐지하는 소프트웨어 몽타주 기법을 제안하고 효용성을 평가한다.

본 논문의 구성은 다음과 같다. 2장에서는 기존 연구에 대해서 설명한다. 3장에서는 소프트웨어 몽타주에 대해서 설명한다. 4장에서는 본 논문에서 제안하는 소프트웨어 몽타주 시스템의 구조를 설명한다. 5장에서는 실험 및 평가를 하고, 6장에서 결론을 맺는다.

· 이 논문은 정부(교육과학기술부)의 재원으로 한국과학재단의 지원(No. 2010-0000-258)과 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음(NIPA-2010-C1090-1031-004)
· 이 논문은 제36회 추계학술발표회에서 '소프트웨어 몽타주: 디지털 포렌식 수사를 위한 유사 소프트웨어 탐지 기법'의 제목으로 발표된 논문을 확장한 것임

† 학생회원 : KAIST 전산학과
hwpark@pllab.kaist.ac.kr
†† 종신회원 : KAIST 전산학과 교수
han@cs.kaist.ac.kr

논문접수 : 2009년 12월 14일
심사완료 : 2010년 2월 10일

2. 관련 연구

유사한 파일을 찾기 위한 시도로서 파일의 블록 단위 해시를 이용하는 연구가 있었다[6]. 파일의 부분적인 변경에 대해서도 유사 블록의 개수에 의해서 적절한 유사도를 구할 수 있다. 그러나 수정된 부분 이후에는 해시 값이 모두 달라진다는 단점이 있다.

블록 해시 알고리즘의 단점을 보완하기 위한 방법으로 CTPH(Context Triggered Piecewise Hash) 알고리즘이 연구되었다[6]. 비교 대상 파일에서 Context라는 구분자를 식별하고 그것을 기준으로 블록에 대한 해시 값을 구하기 때문에 부분 변경에 대해서 블록 해시보다 더 적절한 유사도를 구할 수 있다. 그러나 두 알고리즘 모두 유사 파일 탐지에만 사용될 수 있고 유사 기능을 포함하는 프로그램을 탐지하는 데는 적합하지 않다.

유사한 소스 파일을 찾기 위한 노력은 MOSS(a Measure Of Software Similarity)와 같은 표절 검사 기법에서 다루어지고 있다[7]. 프로그램 소스를 토큰 단위로 분리하고 스트링 매칭을 통해서 유사도를 구하는 이 기법은 소스 파일에만 적용될 수 있다는 한계가 있다.

소프트웨어 버스마크는 소프트웨어가 가지고 있는 고유한 특성을 비교하는 방법이다[8]. 이 방법은 소스 코드가 없는 경우에도 바이너리에 직접 적용할 수 있는 방법이다. 하지만 소프트웨어 버스마크도 유사 소프트웨어를 탐지해주지는 못한다.

유사한 기능을 가진 소프트웨어를 분류하기 위해서 자동화된 소프트웨어 분류 기법이 제안되기도 했다[9]. 이 기법은 소프트웨어로부터 추출된 문자열을 기반으로 알려지지 않은 프로그램에 대한 자동 분류 테스트를 수행한다. 그러나 소프트웨어에서 추출 가능한 문자열만을 대상으로 했기 때문에 문자열이 거의 없는 프로그램이나 암호화된 문자열을 사용할 경우에는 분류가 될 수 없다는 한계가 있다.

3. 소프트웨어 몽타주

몽타주(Montage)라는 말은 조립한다(monter)라는 프랑스어에서 유래된 단어이다. 흔히 범죄수사에서 목격자의 진술을 토대로 범인의 모습과 비슷한 눈, 코, 입 등의 자료를 합성하여 범인의 모습과 유사하게 그린 얼굴 사진을 의미한다[10].

소프트웨어 몽타주는 일반적인 몽타주의 개념을 소프트웨어에 적용시킨 것이다. 즉, 소프트웨어의 기능 및 특성을 함축하고 있는 정보라고 정의한다. 예를 들어 이미지 처리에 관련된 소프트웨어가 있을 때 이 소프트웨어의 몽타주는 이미지 로딩, 확대 축소, 포맷 변환, 인쇄 등으로 요약될 수 있다. 이런 특징들은 주로 소프트웨어의 메뉴 구성 문구나 도움말 문구에서 추출할 수 있는데 이것을 소프트웨어의 몽타주라고 정의한다.

소프트웨어 몽타주는 대용량 저장 매체에 포함된 많은 실행 파일로부터 몽타주와 유사한 프로그램들을 1차적으로 필터링하기 위한 용도에 적합하다. 따라서 정확하고 엄밀한 분석보다는 빠른 추출 및 비교 속도가 중요하다. 몽타주를 이용한 검색은 양성 오류(False Positive)와 음성 오류(False Negative)가 생길 수 있다. 양성 오류란 두 프로그램이 서로 유사한 기능이 없지만 유사하다고 검색한 것을 의미하고, 음성 오류란 유사한 기능이 있는데도 검색되지 못한 것을 의미한다. 소프트웨어 몽타주는 양성 오류보다 음성 오류를 줄이는 것이 중요하다. 그 이유는 양성 오류 때문에 검색된 프로그램은 2차적으로 좀더 엄밀한 방법으로 검사하여 유사 여부를 추후에 확정할 수 있지만 음성 오류 때문에 검색되지 못한 프로그램은 추후 유사성을 판단하는 것이 불가능하기 때문이다.

4. 몽타주 기반 소프트웨어 검색 시스템

몽타주 기반 소프트웨어 검색 시스템의 구조는 그림 1과 같이 몽타주 생성 단계와 검색 단계로 나뉜다. 몽타주 생성 단계에서는 잘 알려진 프로그램 집합을 입력으로 받아서 공통적인 특징을 바탕으로 소프트웨어 몽타주를 생성한다. 몽타주 기반 검색 단계에서는 입력으로 대용량 저장 매체와 소프트웨어 몽타주를 받는다. 그리고 저장 매체에 포함된 수많은 프로그램 중에서 몽타주와 유사한 소프트웨어 집합을 유사도가 높은 순서대로 출력한다.

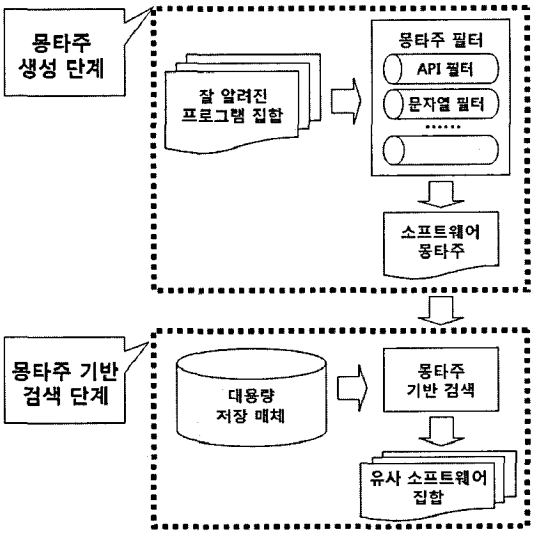

그림 1 몽타주 기반 소프트웨어 검색 시스템 구조도

5. 실험 및 평가

소프트웨어 몽타주의 효용성을 검증하기 위해서 메신저 프로그램에 대한 실험을 하였다. 시스템 구현은 Visual C++를 사용하였고, Intel Core2Quad CPU 2.83Ghz, RAM 4GB의 Windows XP 시스템에서 성능을 평가하였다. 몽타주 기반 조사 대상은 다양한 프로그램이 설치된 하드 디스크의 C 드라이브를 선택하였다. 드라이브 크기는 70GB이고 54.4GB의 용량의 파일이 저장되어 있다. 저장된 총 파일의 개수는 101,101개이고 검색 대상인 EXE와 DLL 파일의 개수는 8,212개이다.

몽타주 생성을 위해서 대표 메신저 프로그램 목록과 몽타주를 기반으로 검색하고자 하는 메신저 유사 프로그램 목록은 각각 표 1, 2와 같다. 일반적으로 잘 알려진 메신저 프로그램을 몽타주 작성에 사용했으며 비교적 잘 알려지지 않았거나 메신저와 유사 기능을 하는 프로그램을 검색 대상으로 선정하였다.

표 1 몽타주 작성에 사용된 대표 메신저 프로그램

프로그램 이름	버전
네이트온[11]	4.0
MSN Live Messenger[12]	2009
버디버디 메신저[13]	7.0
야후 메신저[14]	9.0
Google Talk[15]	2009

표 2 몽타주 검색 대상인 메신저 유사 프로그램

프로그램 이름	버전
다음 메신저 터치[16]	5.5
세이클럽 타키[17]	2.9
드림위즈 지니[18]	7.0
미스리 메신저[19]	5
ICQ[20]	6.5
AOL 메신저[21]	6
스카이프[22]	3.1

5.1 API 기반 소프트웨어 몽타주 기법

API(Application Programming Interface) 기반 소프트웨어 몽타주는 비슷한 기능을 포함하는 소프트웨어들이 기능 구현을 위해서 유사한 Windows API 함수를 사용할 것이라는 전제에서 시작한다. 만일 메신저 프로그램으로부터 몽타주를 생성한다면 메신저 프로그램이 공통적으로 사용하는 네트워크 접속이나 메시지 송수신 관련 API 함수를 필수적으로 호출할 것임을 예측할 수 있다. 따라서 이러한 API 함수를 호출하는 프로그램을 검색해내는 것이 API 기반 소프트웨어 몽타주이다.

5.1.1 API 몽타주 생성

API 몽타주를 생성하기 위해서 대표 프로그램 집합으로부터 API 호출 정보를 추출해야만 한다. 이것을 위해서 Windows PE(Portable Executable) 파일의 임포트 테이블(Import Table) 정보를 사용할 수 있다.

5개의 대표 메신저 프로그램으로부터 실제로 생성된 API 몽타주 개수는 3,578개였고 추출에 소요된 시간은 0.8초로 측정되었다. 그 중에서 호출 빈도 별로 요약한 결과는 표 3과 같다. 5개의 프로그램으로부터 임포트 테이블만 검색해서 몽타주를 만들기 때문에 매우 빠른 속도로 몽타주 생성이 가능하다.

API 몽타주는 API 함수 이름과 그 함수 호출이 몇 개의 대표 프로그램에서 사용되었는지에 대한 정보를 포함한다. 대표 프로그램 집합에서 많이 사용된 API 함수일수록 그만큼의 가중치를 부여하여 유사도 계산을 한다.

표 3 API 기반 소프트웨어 몽타주

DLL이름:API 함수 이름	호출 빈도
WS2_32.DLL:WSAStartup	5
WS2_32.DLL:WSACleanup	5
⋮	
WS2_32.DLL:recv	4
WS2_32.DLL:htons	4
⋮	
USER32.DLL:SetClassLongW	1
USER32.DLL:DefDlgProcW	1

5.1.2 API 몽타주 기반 검색

생성된 소프트웨어 몽타주로부터 유사 소프트웨어를 검색하기 위해서는 검색 대상 매체에 저장되어 있는 모든 exe 파일과 dll 파일에 대해서 API 호출 정보를 추출하고 대표 프로그램으로부터 추출된 몽타주와 비교한다.

예를 들어 표 3과 같이 생성된 API 몽타주를 기반으로 유사 프로그램을 검색한다고 가정하자. 어떤 프로그램의 임포트 테이블에 WS2_32.DLL:recv 함수가 포함되어 있다면 몽타주에서 그 함수의 호출 빈도를 검색해서 해당 점수를 몽타주 유사도 점수에 추가한다. 즉, recv함수의 호출 빈도가 4이기 때문에 4점이 추가된다.

이런 방법으로 몽타주에 포함된 모든 API 함수를 대상으로 유사도 점수를 누적시켜서 최종 점수를 얻는다. 그리고 전체 파일에 대해서 점수가 가장 높은 순서대로 검색 결과 파일을 생성한다.

표 4는 API 몽타주 기반 검색 결과의 예이다. 전체 검색에 소요된 시간은 495.4초로 측정되었다. API 몽타주 생성에 사용된 대표 메신저 프로그램들이 상위에 랭크되어 있음을 확인할 수 있다. 이 결과로부터 검색하고자 하는 메신저 유사 프로그램이 얼마나 높은 순위에서 검색되었는지를 확인할 필요가 있다.

표 4 API 몽타주 기반 검색 결과 Top 10위

순위	점수	검색된 프로그램
1	3286	c:\prog\Windows Live\Messenger\msnmsgr.exe
2	2447	c:\prog\Yahoo\Messenger\YahooMessenger.exe
3	1803	c:\prog\Google\Google Talk\googletalk.exe
4	1793	c:\prog\GPost\gplusphone\gmsgphone_cc.exe
5	1734	c:\prog\NATEON\BIN\NateOnMain.exe
6	1670	c:\prog\Messenger\msmsgs.exe
6	1670	c:\win\ServPack\i386\msmsgs.exe
6	1670	c:\win\ServPack\ServCache\i386\msmsgs.exe
9	1654	c:\prog\Common Files\Ahead\Lib\AdvrCntr2.dll
10	1649	c:\win\system32\wmp.dll

표 5 API 몽타주 기반 메신저 유사 프로그램 검색 결과

유사 프로그램 리스트	API 유사도 점수	검색 순위
스카이프 3.1	1580	16위
미스릭 메신저 5	1546	21위
ICQ 6.5	1427	33위
세이클럽 타키 2.9	1319	53위
AOL 메신저 6	1201	86위
드림위즈 지니 7.0	1014	214위
다음 메신저 터치 5.5	984	241위
평균	1295.9	94.9위

표 5는 API 검색 결과 메신저 유사 프로그램들의 API 유사도 점수와 몽타주 기반 검색 결과 순위이다. 7개의 유사 프로그램 중에서 5개가 100위 안에서 검색된 것을 확인할 수 있다. 그러나 드림위즈 지니와 다음 메신저 터치는 검색 순위가 200위권이었다. 그 이유를 조사해본 결과, 이러한 프로그램들은 자체 제작된 DLL을 이용해서 기능을 분산시키고 분산된 DLL에서 기능을 구현한 형태임을 확인할 수 있었다. 한 파일에 기능 구현을 집중시킨 프로그램과는 달리 이렇게 기능이 분리되어 구현된 프로그램은 유사도 점수가 상대적으로 낮아진다. 결과적으로 모든 유사 프로그램 리스트를 검색하기 위해서는 241위까지 검색해야만 한다. 전체 검색 대상인 exe와 dll 파일 개수가 8,212개이기 때문에 241/8212 ≈ 0.029, 즉 2.9% 정도로 검색 대상을 줄여줄 수 있다.

5.2 문자열 기반 소프트웨어 몽타주 기법

문자열 기반 소프트웨어 몽타주는 비슷한 소프트웨어들은 유사한 기능을 갖기 위해서 유사한 문자열을 프로그램 내에 포함하고 있을 것이라는 전제에서 시작한다. 즉, 소프트웨어에서 사용된 메뉴 문구나 입출력에 사용되는 메시지 문구 등의 문자열들을 추출한다면 유사 기능을 대표하는 문자열 집합을 만들 수 있다. 이러한 문자열 집합을 기반으로 유사 소프트웨어를 검색하는 방법이 문자열 기반 소프트웨어 몽타주이다.

5.2.1 문자열 몽타주 생성

문자열 몽타주를 생성하기 위해서는 대표 소프트웨어의 바이너리를 바이트 단위로 읽어서 연속된 아스키코드 정보를 추출해낸다. 그리고 사전 검색을 통해서 무의미한 단어를 걸러내고 몽타주에 포함시킬 단어를 최종적으로 선택한다. 이때 임포트 테이블에 포함된 API 함수 호출에 대한 정보도 아스키코드로 저장되어 있기 때문에 문자열 몽타주에 포함된다. 문자열 몽타주 생성은 프로그램 바이너리를 처음부터 끝까지 바이트 단위로 읽어서 스트링 토큰을 분리하고 사전 검색을 통해서 무의미한 단어를 배제시키는 작업이 필요하기 때문에 API 몽타주에 비하여 실행 성능은 떨어진다.

표 6 문자열 기반 소프트웨어 몽타주

문자열 이름	발생 빈도
accept	5
connect	5
⋮	
message	4
session	4
⋮	
forbidden	1
inactive	1

5개의 대표 메신저 프로그램으로부터 실제로 생성된 문자열 몽타주 개수는 1,768개였고, 생성에 소요된 시간은 1.64초로 측정되었다. 그 중에서 발생 빈도 별로 몇 개의 문자열들을 살펴보면 표 6과 같다.

문자열 몽타주는 해당 문자열이 몇 개의 대표 프로그램에서 사용되었는지에 대한 정보를 포함하고 있다. 많은 대표 프로그램에서 사용된 문자열일수록 그만큼의 가중치를 가지고 유사도 점수 계산에 사용된다.

5.2.2 문자열 몽타주 기반 검색

생성된 문자열 몽타주로부터 유사 소프트웨어를 검색하기 위해서는 검색 대상 매체에 저장되어 있는 모든 exe 파일과 dll 파일에 대해서 문자열 정보를 추출하고 대표 프로그램으로부터 정보와 비교한다. 대표 프로그램의 몽타주에 등록된 문자열과 같은 문자열을 포함하고 있는 프로그램은 발생 빈도에 해당하는 유사도 점수를 얻는다. 프로그램에 포함된 모든 문자열에 대해서 점수를 누적시키면 최종 점수를 얻을 수 있다. 그리고 점수가 가장 높은 순서대로 결과 파일을 생성한다.

표 7은 문자열 몽타주 기반 검색 결과의 예이다. 전체 검색에 소요된 시간은 871.4초로 측정되었다. 문자열 몽타주에서는 대표 프로그램 중에서 2개가 Top 10에 랭크되어 있음을 확인할 수 있다.

표 7 문자열 몽타주 기반 검색 결과 Top 10위

순위	점수	검색된 프로그램
1	2707	c:\prog\Skype\Phone\Skype.exe
2	2245	c:\nvidia\driver\SystemSoftware.exe
3	2162	c:\prog\Google Talk\googletalk.exe
4	2091	c:\win\system32\nvcpl.dll
5	2076	c:\prog\Nero\Nero 7\Core\nero.exe
6	2051	c:\prog\Rhino\ServU\ServUAdmin.exe
7	2025	c:\prog\Hnc\ImgFilters\GS\gsdll32.dll
8	1999	c:\prog\Acrobat 6.0\Acrobat\Acrobat.exe
9	1993	c:\prog\Yahoo\Messenger\YahooMessenger.exe
10	1966	c:\win\Microsoft.NET\v2.0.50727\mscorlib.dll

표 8 문자열 몽타주 기반 메신저 유사 프로그램 검색 결과

용의 프로그램 리스트	문자열 유사도 점수	검색 순위
스카이프 3.1	2707	1위
ICQ 6.5	1710	35위
AOL 메신저 6	1705	37위
세이클럽 타키 2.9	1648	45위
미스리 메신저 5	1630	49위
다음 메신저 터치 5.5	1469	87위
드림위즈 지니 7.0	1260	165위
평균	1732.7	59.9위

표 8은 문자열 유사도 점수와 검색 결과 순위이다. 7개의 유사 프로그램 중 6개가 100위 안에서 검색되었고 드림위즈 지니의 경우는 165위로 벗어났다. 영어 문자열을 기반으로 검색했기 때문에 한글 기반 메신저 프로그램들이 영어 메신저 프로그램보다 검색 순위가 떨어지는 것으로 조사되었다.

결과적으로 모든 유사 프로그램 리스트를 검색하기 위해서는 165위까지 검색해야만 한다. 전체 검색 대상인 exe와 dll 파일 개수가 8,212개이기 때문에 165/8212 = 0.020, 즉 2.0% 정도로 검색 대상을 줄여줄 수 있다.

대표 프로그램 5개에 대한 결과만을 본다면 API 몽타주가 더 좋은 결과를 보였고, 7개의 대상 프로그램에 대한 결과는 문자열 몽타주가 더 좋은 결과를 보였다. 그 이유는 프로그램마다 API 몽타주가 더 적합한 경우도 있고, 문자열 몽타주에 더 적합한 경우도 있기 때문이다. 따라서 제한된 실험 결과만으로는 어느 몽타주가 더 우수하다고 판단할 수 없다.

6. 결 론

본 논문에서는 소프트웨어 포렌식 목적으로 사용될 수 있는 소프트웨어 몽타주 개념을 새롭게 제안하고, 인스턴트 메신저와 유사한 프로그램을 탐지하는 실험을 하였다. 그 결과 소프트웨어 몽타주는 대용량 저장매체를 대상으로 빠르게 유사 프로그램을 검색하기 위한 도구로 사용될 수 있다는 것을 확인하였다.

향후 과제로는 탐지 결과에 대해서 정확률(Precision)과 재현율(Recall)을 이용하여 수치적인 평가를 할 예정이다. 그리고 API 몽타주와 문자열 몽타주의 장점을 조합한 형태의 몽타주도 고려하고 있으며 몽타주로 사용될 수 있는 새로운 특성을 찾아내는 노력도 계속할 계획이다. 또한 후속 연구로서, 필터링된 검색 대상 중에서, 실제 유사 프로그램을 탐지하는 방법에 대한 구체적인 연구가 필요하다.

참 고 문 헌

[1] K. Lim, J. Park, S. Lee, "Trends and challenges of current digital forensics," *Journal of Sec. Eng.*, vol.5, no.4, pp.47-59, Nov. 2008. (in Korean)

[2] EnCase, "Complete data collection solution," http://www.guidancesoftware.com.

[3] S. Hong, J. Bang, K. Lim, S. Lee, "Instant messenger analysis in digital forensic viewpoint," *Proc. of the Info. Sec. & Crypt*, vol.18, no.1, pp.450-453, 2008. (in Korean)

[4] K. Kim, S. Park, "Trends of current software reference data set," *Journal of Korea Inst. of Info. Sec. & Crypt.*, vol.18, no.1, pp.70-77, Feb. 2008. (in Korean)

[5] NSRL, "National Software Reference Library," http://www.nsrl.nist.gov.

[6] K. Seo, K. Lim, S. Lee, "Detecting similar files for digital forensic investigation," *Journal of Sec. Eng.*, vol.7, no.2, pp.182-190, Apr. 2009.(in Korean)

[7] MOSS, "A System for Detecting Software Plagiarism," http://theory.stanford.edu/~aiken/moss/.

[8] H. Tamada, M. Nakamura, A. Monden, K. Matsumoto, "Java birthmark -detecting the software theft-," *IEICE Trans. on Info. & Syst*, vol.E88-D, no.9, pp.2148-2158. Sept. 2005.

[9] W. Cho, H. Park, T. Han, "Fast and automatic classification of software," *Proc. of the KIISE*, vol.35, no.2, pp.59-60, Oct. 2008. (in Korean)

[10] C. Choi, S. Lee, "Computing Similarity between Montages and Facial Photos," *Proc. of the KIISE*, vol.33, no.2, pp.453-458, Oct. 2006. (in Korean)

[11] Nateon Messenger, http://nateon.nate.com.

[12] MSN Live Messenger, http://download.live.com.

[13] BuddyBuddy, http://messenger.buddybuddy.co.kr.

[14] Yahoo Messenger, http://messenger.yahoo.com.

[15] Google Talk, http://www.google.com/talk.

[16] Daum Messenger, http://messenger.daum.net.

[17] Sayclub Messenger, http://tachy.sayclub.com.

[18] Dreamwiz, http://www.dreamwiz.com/mgn.

[19] Misslee Messenger, http://www.misslee.net.

[20] ICQ Messenger, http://www.icq.com.

[21] AOL Messenger, http://www.aim.com.

[22] Skype, http://www.skype.com.