비대면 업무에서 개인 정보 보호를 위한 화면 캡처 방법에 관한 연구

황교찬[1], 정효숙[2], 최규성[3]

[1]㈜워크스타일 대표, [2]㈜워크스타일 수석연구원, [3]㈜워크스타일 전임연구원

[1]eekdro@gmail.com, [2]hyosook.j@gmail.com, [3]tigermap7@gmail.com

A study on the screen capture method for privacy protection within untact (non-face-to-face)

Hwang Kyo Chan, Jeong Hyo Sook, Choi Gyu Seong

WORKSTYLE Co., Ltd.,

요 약

본 연구는 컴퓨터를 사용하는 사무직 근로자가 업무 중 침해받을 수 있는 개인 정보 보호 방법을 제시한다. 일반적으로, 회사에서 이용되는 보스웨어, 원격제어와 같은 프로그램들은 화면을 캡처해서 상대 또는 서버로 전송되는데 이때, 개인 정보 보호의 취약점이 발생하게 된다. 이에 본 논문에서는, 업무 중 사용하는 소프트웨어에 대한 상세한 정보를 수집할 수 있는 프로그램을 개발하여 화면 캡처 시 프라이버시 보호가 필요한 부분을 지정 및 선별하여 개인 정보의 유출을 방지하는 것을 목적으로 한다. 지정 및 선별 방법은 '키워드 기반의 정보 보호'와 '선별적 캡처' 방법으로 구분하여 구현하였다. '키워드 기반의 정보 보호'는 사전에 지정한 키워드를 중심으로 개인 정보가 포함된 캡처를 완전히 배제하는 방식이며, '선별적 캡처'는 사전에 지정한 프로그램을 캡처에서 제외하고 특정 프로그램의 크기만큼만 캡처하는 방식을 제안하고자 한다.

Ⅰ. 서 론

세계는 코로나-19에 의한 팬데믹으로 업무 형태가 빠르게 변화되고 있다. 업무 변화의 중심에는 비대면을 통한 업무 처리 과정이다. 이에 따라, 국내외 원격근무 솔루션 기업의 성장과 관련 원격프로그램 사용률이 높게 증가하고 있으며, 이러한 프로그램들은 실시간으로 컴퓨터 상태, 화면, 사용자의 음성, 영상 등을 공유하거나 수집하는 기능을 내포하고 있어 근무자의 의도와 다르게 기업 또는 개인 정보가 노출될 수 있다. 이런 문제 해결을 위해 기존 연구는 컴퓨터 화면을 캡처한 이미지를 광학 문자 인식 OCR(Optical Character Recognition)을 이용해 분석 후 정보 보호가 필요 내용을 선별하였다.[1],[2] 하지만, OCR을 통한 분석은 인식률 한계로 100%가 불가능하며, 이미지를 분석하거나 클라우드 처리는 컴퓨터의 부하와 높은 컴퓨팅 자원이 필요하다. OCR 분석은 회사가 사용자를 감시한다는 불쾌감과 관련 높은 자원 처리에 따른 업무에 영향을 주게 된다.

본 연구는 컴퓨터를 사용하는 사무직 근무자가 전산 업무 중 개인 정보를 침해받을 수 있는 프로그램 사용 비율을 확인하고 업무 중인 컴퓨터 화면을 캡처하려고 할 때의 정보 보호를 두 가지 방법으로 제시한다. 이를 위해서 화면 캡처 시 개인 정보 보호가 필요한 프로그램을 '선별적 캡처'와 '키워드 기반의 정보 보호'로 구현하였다. '선별적 캡처'는 사전에 지정한 프로그램을 캡처에서 제외하고 특정 프로그램의 크기만큼만 캡처하는 방법이다. '키워드 기반의 정보 보호'는 사전에 지정한 키워드가 캡처하려고 하는 프로그램의 타이틀에 포함되어 있다면 캡처에서 배제 시키는 방법이다.

본 연구의 실증을 위해 실험 대상자의 컴퓨터 사용 정보를 수집할 수 있는 프로그램을 개발하여 업무 중 사용하는 모든 프로그램에 대한 상세한 정보를 수집하여 연구를 진행했다.[3]

Ⅱ. 본 론

본 연구는 화면 캡처 시 발생할 수 있는 정보 보호를 위해 아래 4가지 해결 요구조건을 도출하였다.

1) 기업과 개인별 프로그램 개별 등록이 가능해야 한다.
2) 화면 캡처는 특정 프로그램의 부분을 캡처해야 한다.
3) 문서 프로그램 이용 시 특정 파일명에 따른 캡처 제한이 필요하다.
4) 웹사이트는 업무와 개인의 공간이 혼재되어 있으므로 포괄적인 방법으로 개인과 기업의 정보 보호 기술이 모색되어야 한다.

1), 2) 해결을 위한 '선별적 캡처' 방법으로 프로그램을 구현하고자 했다. 1) 해결을 위해 특정 프로그램을 캡처에서 제외할 수 있게 사전에 지정하여 보관하고 캡처 시 해당 정보를 읽어와서 선별적 캡처를 했으며, 2) 해결을 위해 화면 전체가 아닌 화면상에 띄워져 있는 프로그램의 위치와 크기에 맞게 부분 캡처를 구현했다. 3), 4) 해결을 위해 '키워드 기반의 정보 보호' 방법으로 프로그램을 구현하고자 했다. Fig. 1를 보면 통상적으로 특정 파일을 읽어와 업무를 수행하는 프로그램들은 타이틀에 해당 파일명이 표기된다.

NO	Program	USE	Title
1	HWP 2018	Editing a document	A Study on selective data for personal information protection.hwp [c:\User\ForUS\Desktop\] - 한글
2	Microsoft Excel	Editing a document	Accounting management.xlsx - Excel
3	Visual Studio	Development tools	SimulatorAdmin
4	Chrome	Web browser	Naver

Fig. 1. Get information from the title of the program

이를 이용하여 사전에 지정한 키워드가 실행 중인 프로그램의 타이틀에 포함되어 있는지 확인하여 캡처에서 배제 시키시는 방법으로 구현했다.

이를 위한 프로그램은 웹(Web Application Server)기반으로 개발된 관리자 사이트와 실험 대상이 되는 근무자(Staff)의 PC에 설치되는 에이전트 프로그램(Agent program) 그리고 에이전트 프로그램과 통신을 통해 데이터베이스(Database)에 정보를 저장하고 캡처된 이미지를 저장소(Storage)에 저장하는 Api Server로 구성되어 있다.

Web Application Server를 통해 관리자는 근무자의 요청 또는 기업의 사용 프로그램을 고려하여 캡처 금지 키워드 및 캡처 금지 프로그램을 등록하게 된다

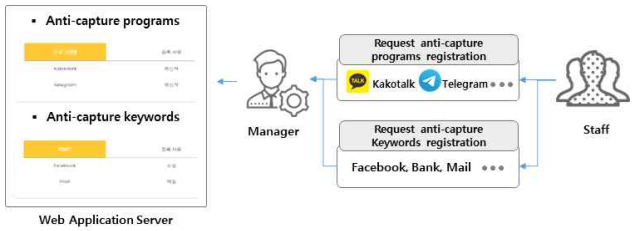

Fig. 2. Registration anti-capture keywords and programs

에이전트 프로그램은 근무자의 PC에서 주기적으로 화면을 캡처하고 실행되고 있는 프로그램의 정보를 Api Server로 정보를 전송하는 기능을 한다. Fig. 3은 에이전트가 PC 구동 후 처음 실행됐을 때 작동의 흐름을 나타낸다.

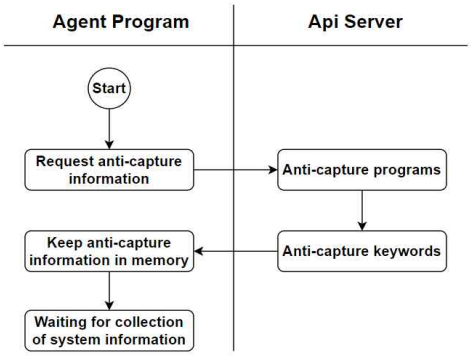

Fig. 3. Flow chart when the agent program is started on the PC

PC에 전원이 들어오고 윈도우(Window)가 구동되면 에이전트 프로그램이 시작되고 Api Server에게 정보 보호에 필요한 캡처 금지 정보를 요청하게(Request anti-capture information) 된다. Api Server는 Fig. 2에서 지정한 캡처 금지 프로그램(Anti-capture programs) 리스트와 캡처 금지 키워드(Anti-capture keywords) 리스트를 에이전트 프로그램에게 전달하게 된다. 에이전트 프로그램은 해당 데이터를 메모리에 보관하고(Keep anti-capture information in memory) 정보 획득 주기를(Waiting for collection of system information) 기다리게 된다. 정보 획득 주기는 사전에 관리자 사이트에서 지정하며 해당 주기마다 에이전트 프로그램은 근무자의 PC 화면 캡처 이미지 및 프로그램 사용 정보를 Api Server에게 전달하게 된다.

Fig. 4. Data collected by the agent program

에이전트 프로그램으로 선별적 캡처된 이미지와 이미지의 정보가 관리사 사이트에 등록된 모습을 Fig. 4와 같이 볼 수 있다. 데이터베이스(Database)와 저장소(Storage)에 정보가 올라가기 전에 PC의 에이전트 프로그램에서 선별적 캡처를 진행함으로써 캡처 이미지에 의해 발생할 수 있는 개인 또는 기업 정보의 유출을 원천적으로 보호할 수 있게 된다. 개인 정보 수집을 동의받은 9명 근로자가 근무 시간 동안 캡처된 이미지 개수를 정리한 것이 Table 1이다.

staff	Working time(h)	Capture Counter	Personal Messenger & keyword	anti-Capture Rate
A	17.3	230	26.3%	100%
B	19.7	178	43.2%	100%
C	24.3	250	31.1%	100%
D	19.9	234	15.3%	100%
E	22.3	202	34.9%	100%
F	17.1	132	25.3%	100%
G	17.8	123	31.9%	100%
H	13.6	118	6.4%	100%
I	22.0	278	8.7%	100%

Table 1. Analysis of the results of the anti-capture

캡처된 이미지 개수는 활동 시간에 비례하지 않는다. 같은 이미지의 경우 캡처를 저장하지 않고 캡처가 금지되면 역시 이미지를 저장하지 않기 때문이다. 캡처 금지 성공률은 캡처된 이미지 중 캡처 금지 키워드나 캡처 금지 프로그램에 포함되는 프로그램이 캡처 이미지에 포함되지 않는 비율이다. 100%가 나오는 이유는 본 연구에서 제시하는 선별적 캡처 방법을 통해 화면 캡처 전에 이미지 저장 여부를 선별하기 때문이다.

Ⅲ. 결론

기업에서 내부적인 자료의 보안 측면 이슈는 항상 있었다. 하지만 기업 보안 이슈와 함께 개인 정보 보호도 요구된다. CCTV 역시 자리 잡기 전까지는 보안과 사생활 보호로 많은 이슈가 있었다. 컴퓨터에 대한 모니터링 역시 기업의 보안과 개인 정보 문제는 향후 더욱 논쟁거리가 될 수밖에 없다.

본 연구는 중앙 관리 형태의 정보 보호 관리 정책을 기준으로 정보 보호 대상의 프로그램에 대한 사용률만을 측정하였다. 향후 연구 방향은 '사이트'나 '파일' 에 포함된 키워드 사용률도 추가로 측정하여 키워드에 속하는 개인 정보 침해를 방지하는 방안을 함께 모색할 것이다. 또한, 개인화된 정보 보호 캡처 방법에 관한 연구를 진행할 것이다.

ACKNOWLEDGMENT

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원(IITP)의 지원을 받아 수행된 연구입니다.(2022-0-00664). 관계부처에 감사드립니다.

참 고 문 헌

[1] Jong-Kyung Baek. (2020) A Personal Information Security System using Form Recognition and Optical Character Recognition in Electronic Documents. Journal of the Korea Academia-Industrial cooperation Society Vol. 21, No. 5 pp. 451-457

[2] Sukhyeon Kim. (2017) A Study on Detecting Personal Information from Image. Proceedings of the Korean Society of Computer Information Conference 25-1

[3] WorkStyle. (Url: http://ShareWBS.com)