Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.25, NO.3, Jun. 2015 ISSN 2288-2715(Online)
http://dx.doi.org/10.13089/JKIISC.2015.25.3.627

MMORPG 게임 내 계정도용 탐지 모델에 관한 연구*

김 하 나,† 곽 병 일, 김 휘 강‡
고려대학교 정보보호대학원

A study on the identity theft detection model in MMORPGs*

Hana Kim,† Byung Il Kwak, Huy Kang Kim‡
Graduate School of Information Security, Korea University

요 약

온라인 게임시장의 성장으로 아이템 거래시장이 활성화됨에 따라 아이템 현금 거래시장은 1조 6,000억원 규모로 성장하였으며, 활성화된 시장으로 인해 아이템 및 게임머니의 현금화가 용이하게 되었다. 이러한 특성으로 악의적인 사용자들은 온라인 게임에서 계정도용을 통해 금전적인 가치가 높은 희귀 아이템 및 게임머니를 탈취하여 현금화하는 사례가 빈번히 발생하고 있는 실정이다. 본 연구에서는 MMORPG(Massive Multi-user Online Role Playing Game)내에서의 계정도용자들의 행위분석을 통한 계정도용 탐지모델을 제안한다. 계정도용의 경우 현금화 시킬 수 있는 아이템 및 게임머니를 탈취해야하기 때문에 게임 행동상에서 경제활동에 치중되어 있으며 아이템 생산, 아이템 판매, 게임머니 획득 이라는 특정 시퀀스를 가지고 있다. 이를 기반으로 계정도용 탐지모델을 제안하였으며, 본 논문의 탐지모델을 활용하여 분류한 결과 84%의 정확도를 보였다. 더불어 거래 네트워크 분석을 통해 계정도용 시 발생하는 거래특성에 대해 분석하였다.

ABSTRACT

As game item trading becomes more popular with the rapid growth of online game market, the market for trading game items by cash has increased up to KRW 1.6 trillion. Thanks to this active market, it has been easy to turn these items and game money into real money. As a result, some malicious users have often attempted to steal other players' rare and valuable game items by using their account. Therefore, this study proposes a detection model through analysis on these account thieves' behavior in the Massive Multiuser Online Role Playing Game(MMORPG). In case of online game identity theft, the thieves engage in economic activities only with a goal of stealing game items and game money. In this pattern are found particular sequences such as item production, item sales and acquisition of game money. Based on this pattern, this study proposes a detection model. This detection model-based classification revealed 86 percent of accuracy. In addition, trading patterns when online game identity was stolen were analyzed in this study.

Keywords: Identity theft, Online game security, User behavior analysis, MMORPG

I. 서 론

2009년부터 급증한 개인정보 유출사건으로 인해 많은 사용자들의 ID와 비밀번호가 유출되면서, 보이스 피싱, 대출사기, 계정도용 등의 2차 피해가 발생하였다. Fig.1.은 2009년부터 2014년까지 발생한 개

접수일(2015년 4월 20일), 수정일(2015년 6월 8일), 게재확정일(2015년 6월 8일)
* 본 연구는 2014년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. 한국연구재단에서 부여한 과제번호 : 2014R1A1A1006228)
† 주저자, hanada@korea.ac.kr
‡ 교신저자, cenda@korea.ac.kr (Corresponding author)

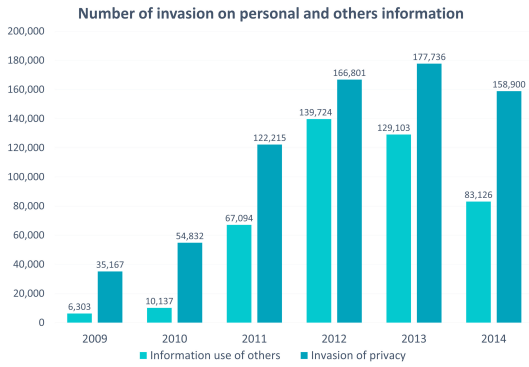

Fig. 1. Number of invasion on personal and others information

인정보 침해건수 및 타인정보 이용건수를 나타낸다[1]. 지난 2011년 SK커뮤니케이션즈에서 약 3,500만개의 개인정보가 유출되자 게임회사들은 계정도용을 예방하기 위해 사용자들에게 2차 보안 서비스를 이용할 것을 요구하였다[2]. 개인정보 유출사건이 계정도용으로 이어지는 이유로는 대다수의 사용자가 동일한 계정정보를 사용하는 것을 들 수 있다. 치안정책연구소의 2014년 조사에 따르면 34.6%의 사용자들이 3~4개의 ID를 사용하고, 44.1%의 사용자가 3~4개의 비밀번호를 사용한다고 나타났다[3].

2014년에 발생한 계정도용 사건으로는 CJ Mall, 네이버, 위메프 사건 등을 들 수 있다. CJ Mall은 49명의 고객의 240여만 원의 포인트가 무단으로 사용되었으며, 위메프의 경우 400여명의 고객의 10~40만 포인트(약 1천 100만원)가 무단 사용되었다[4][5]. 계정도용의 피해를 입은 사용자들은 개인정보 유출과 보상 문제로 기업 측에 불만을 가지게 되고, 기업 측은 대응 및 보상 문제로 금전적인 피해를 입게 된다. CJ Mall은 계정도용 사건에 대해 무대응으로 일관하다 고객들의 항의로 사과문을 게재하였다[5]. 이처럼 계정도용은 금전적인 피해뿐만 아니라 기업이미지에도 영향을 미치게 된다.

지난 2월 발생한 아이핀 유출사건으로 부정 발급된 아이핀 75만 건 중 12만 건이 게임회사인 엔씨소프트, 엑스엘게임즈, 블리자드에서 회원가입 및 계정수정에 이용된 것으로 파악되었다[6]. 계정도용은 오래전부터 게임시장에서 빈번히 발생하였는데 대표적인 사건으로는 2006년 엔씨소프트사의 리니지, 넥슨사의 마비노기 사건을 들 수 있다[7]. 게임시장에서 계정도용이 발생하는 이유로는 활성화된 아이템 현금 거래 시장을 들 수 있다. 아이템 거래 시장은 약 1조 6,000억 원의 규모에 달하며, 아이템 중개 사이트인 아이템베이와 아이템 매니아의 연간 수수료 수익은 약 400억 원에 달한다[8]. 게임시장의 활성화로 게임 아이템 거래 시장 역시 활성화 되었으며, 유명한 게임 아이템의 경우 수십만 원에서 수천만 원에 거래가 이루어지고 있는 실정이다. 한국콘텐츠진흥원의 2014년 한일 게임이용자 조사보고서에 따르면 24.4%의 게임 사용자들이 게임 거래를 잘 알고, 직접 거래를 해본 경험이 있는 것으로 나타났다[9]. 게임이 취미생활을 넘어 자신의 개성을 드러내는 하나의 컨텐츠로 진화하면서, 게임 아이템에 현금을 투자하는 사용자들이 늘어나게 되었고, 이로 인해 아이템 거래 시장 내 아이템 및 게임머니 현금화가 용이해졌다.

악의적인 범죄자들은 이러한 특성을 악용하여 계정도용을 수행한 후 사용자들의 게임머니와 아이템을 탈취 한 뒤 현금화하여 금전적인 이득을 취한다. 온라인 게임에서 악의적인 범죄자들이 계정을 탈취하는 방법으로는 유출된 개인정보 사용, 사용자들 간의 계정정보 공유 습관 악용, 피싱 등을 들 수 있다. 다른 사이트와 동일한 ID, 비밀번호를 이용하게 될 경우 개인정보 유출로 인한 계정도용 피해가 발생할 수 있으며, 게임 내 사용자들 간의 계정 정보를 공유하는 것을 악의적인 범죄자들이 사회 공학적으로 접근하여 계정정보를 탈취할 수 있다. 마지막으로 게임 사용자들에게 게임사측에서 게임 사용자의 ID와 비밀번호를 요구하는 것처럼 피싱 메일들을 보내 계정정보를 탈취할 수 있다.

본 논문에서는 다양한 도메인 중 MMORPG에서 발생하는 계정도용에 대한 탐지 프레임워크를 제안하였으며, 거래 네트워크 분석을 통해 게임 봇과의 관련성을 제시하였다. 게임 봇(game bot)이란 자동으로 전투, 채집, 사냥 등을 수행하는 악성 행위 프로그램을 말한다. 게임 봇 사용은 정직한 게임 사용자들에게 상대적 박탈감을 주게 되어 게임에 대한 충성도를 감소시키고, 게임 사이클을 단축시키는 등의 문제점을 가지고 있다. 2장에서는 관련연구를 설명하고, 3장에서는 MMORPG에서 계정도용 탐지를 위한 전체적인 프레임워크를 설명하였다. 4장에서는 게임회사로부터 받은 로그 데이터를 사용하여 실험 및 평가하였고, 5장에서 결론과 앞으로의 연구에 대해 설명하였다.

II. 관련 연구

본 장에서는 온라인 상에서의 계정도용탐지에 관련

Table 1. Category of data in detecting and preventing from identification theft

Category	Description
Keyboard input data	Classification using typing rhythm[10]
Mobile device touch data	Classification using touch sensitivity finger movement[11]
Log data in online game	Classification using characters movement activity, login time [12][13][14][15]
Architecture & framework	Classification for processing of authorizing registration and laying out framework [16][17][18]

된 선행 연구들을 데이터 기반으로 분류하고 설명하였다. Table 1.은 계정도용 탐지 및 예방에 사용된 데이터를 기반으로 4가지로 분류한 것이다. 분류는 키보드 입력 데이터, 모바일 기기 터치 입력 데이터, 온라인 게임의 로그 데이터, 프레임워크 및 구조 제안으로 나누어진다.

키보드 입력 데이터는 계정도용 탐지에 키보드 타이밍 및 입력 데이터에 대한 시퀀스 분석 데이터를 사용한다. Gunetti Daniele 등[10]은 온라인 서비스 상에서 샘플 텍스트에 대한 타이핑 리듬 분석을 통계적 기법에 적용하여 사용자 인증 및 침입 탐지 방안을 제시하였다.

모바일 기기 터치 데이터는 터치 스크린에서의 사용자의 손가락 움직임, 터치 압력 강도, 회전 등 관련 분석 데이터를 사용한다. Hojin Seo 등[11]은 모바일 뱅킹에서 터치스크린 사용 시 손가락 입력 패턴 및 데이터 입력에 소요되는 시간과 관련된 입력 패턴을 데이터마이닝 기법에 적용하여 사용자 인증 방법을 제시하였다.

온라인 게임 로그 데이터는 계정도용 탐지에 게임 사용자의 접속시간, 접속한 IP address, MAC address, 캐릭터 이동경로, 캐릭터 행위 등의 분석 데이터를 사용한다. Kuan-Ta Chen 등[12]은 MMORPG에서 게임 이용 시간에 대한 패턴을 기반으로 KLD(Kullback-Leibler divergence) 기법에 적용하여 계정도용 탐지 방법을 제시하였다. Jehwan Oh 등[13]은 MMORPG에서 게임 내 캐릭터 경험치, 거래, 로그인 이용 시간 등의 데이터를 통계적 기법에 적용하여 계정도용 탐지 방법을 제시하였다. Hwa Jae Choi 등[14]은 계정도용자들의 유형을 속전속결형, 신중형, 대담무쌍형으로 분류하고 데이터마이닝 기법을 적용하여 계정도용 탐지 방법을 제시하였다. Jiyoung Woo 등[15]은 MMORPG에서 경험치, IP & MAC address, 접속 시간, 경고 횟수 등의 데이터를 데이터마이닝 기법에 적용하여 계정도용 방법을 제시하였다.

프레임워크 및 구조 제안은 계정도용을 방지하기 위해 사용자 인증 및 피싱 사이트 체크, 계정도용 법률 강화 등의 방법을 제안한다. Paul Madsen 등[16]은 서비스 상에서 2-factor 인증 및 2-channel 인증을 이용한 인증 구조를 통해 사용자 인증 방법을 제시하였다. Neil Chou 등[17]은 인터넷 상에서 URL, POST DATA, 링크, 이미지, 도메인 체크 및 평가를 통한 계정도용 방지 프레임워크를 제시하였다. Alexander Tsoutsanis[18]는 네덜란드, 영국, 미국에서의 SNS 계정도용에 대한 대응 조치를 설명하였고, 계정도용을 막기 위해 계정도용에 대한 법 강화와 지역 법원 방안을 제시하였다.

본 논문은 게임 로그 데이터를 분석하여 계정도용을 시간에 따른 유형으로 분류하고, 데이터마이닝을 적용하여 탐지방법을 제안하였다. 본 논문에서는 MMORPG 내의 계정도용자들의 행위분석을 통해 특정 시퀀스를 파악하였으며, 이를 기반으로 탐지모델을 제시하였다. 지금까지 시퀀스를 기반으로 계정도용을 탐지한 예는 없었으며, 계정도용 내 아이템 이동에 관련한 연구뿐 아니라 계정도용자들의 거래 네트워크 분석을 통하여 계정도용 거래 네트워크 특성에 관한 연구도 수행하였다.

III. 계정도용 탐지를 위한 방법론

제안하는 계정도용 탐지 방법론은 MMORPG에서 계정도용자가 수행한 행위를 기반으로 시간별 유형으로 구분하여 탐지하는 방법이다. 계정도용자들은 대부분 경제활동에 치중되어 있으며, 10분 이내에 게임머니 및 게임 머니 탈취가 이루어진다. 아이템 생성, 아이템 판매, 게임 머니 획득이라는 특정 시퀀스를 가지며, 계정도용 시간이 5분 이상인 경우 Fig.2.의 행위를 반복하는 것을 확인하였다.

특정 시퀀스가 반복되기 때문에 1분 미만, 5분 이

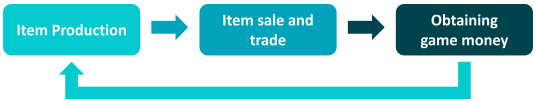

Fig. 2. The sequence of identity theft

하, 10분 이하, 10분 초과를 기준으로 유형을 구분하였으며, 3가지 데이터마이닝 학습알고리즘(Multilayer Perceptron, Logistic, Random Forest)을 사용하여 학습을 수행하였다. Multilayer Perceptron은 신경망 알고리즘으로 네트워크를 형성한 노드와 수정 가능한 링크에 가중치를 부여하고 반복적인 학습을 통해 모델을 생성하는 기법을 말한다. Logistic은 회귀분석 중 하나로 필요한 변수만을 골라내어 모형의 예측과 해석력을 높일 수 있는 기법이며, Random Forest는 앙상블 학습 방법으로 트레이닝 과정에서 다수의 결정트리 모델을 학습한 다음 여러 분류결과들을 종합하여 분류하는 기법이다.

제안하는 프레임워크는 Fig.3.과 같다. 첫 번째 단계에서 계정도용이 발생한 날의 모든 액션 로그들을 수집한 뒤 로그인-로그아웃 단위로 고유한 번호를 지정하였으며, 모든 분석은 로그인-로그아웃 단위로 진행하였다. 계정도용으로 신고 된 시간에 접근한 IP address들을 추출하여 제재 리스트를 생성하였고, 계정도용 당한 계정이 재제 IP address로 접근한 로그들만 추출하여 데이터 분석을 수행하였다. 먼저 계정도용자들의 행위를 분석하기 위해 계정도용 시 사용한 로그들을 Tabel 2.의 11 가지 행위로 분류하여 정

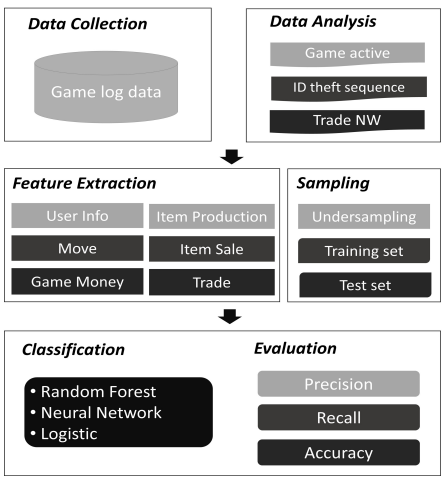

Fig. 3. Framework

상 사용자들과의 사용된 로그 비율을 비교하였으며, 계정도용 피해자들의 일주일간의 게임 머니 로그를 추출하여 계정도용일과 이전 6일의 지출비율을 비교하였다. 더불어 계정도용에 호출된 로그 비율을 기반으로 Tabel 3.과 같이 분류하여 시퀀스 분석을 수행하였다.

두 번째 단계에서는 시퀀스 분석결과를 기반으로 불필요한 분류항목과 로그들을 제외하고 feature를 선정하였으며, 선정된 feature는 Table 6.에 기술하였다. 계정도용 피해 사용자의 95%가 40레벨 이상(2010년 기준 최고레벨은 50)으로 고레벨이였으며, 88%의 사용자가 10분 이내로 게임을 실행하였다. 23개의 계정 중 1개의 계정을 제외하고는 습득한 경험치가 0인데 이는 게임 시 경제활동을 중점으로 수행하기 때문인 것으로 추측된다. 1개의 계정의 경우 4시간 27분 동안 게임을 실행하였으며, 계정도용 시 일반적으로 하지 않는 행위인 아이템 채집을 974번 수행하였다(다른 계정들은 아이템 채집을 수행하지 않음). 채집한 아이템은 요리를 만들거나, 무기를 만드는 아이템으로 값비싼 아이템은 아닌 것으로 파악되었다. 계정도용 시 탈취한 게임 아이템 또는 게임 머니를 전달하기 위하여 특정 위치로 이동하기 때문에 이동관련 로그들을 포함하였다. 아이템 생성 후 판매하여 게임 머니를 획득하고, 획득한 게임 머니를 탈취하기 때문에 게임 머니 관련 로그를 포함하였으며, 계정도용 시

Table 2. Game activity classification table

Category	Description
Battle	Hunting, Battle
Skill	Skill of the character
Friendship	Guild, Party, Quest
Trade	Log related to trade
Item production	Log related to item production
Item sale	NPC shop, User shop, Sales agency
Item etc.	Items that don't fall under production or sale categories
Game money	Game money
Movement	Teleport, Flight, Log related to movement
User information	IP address, Playtime
etc.	any other logs

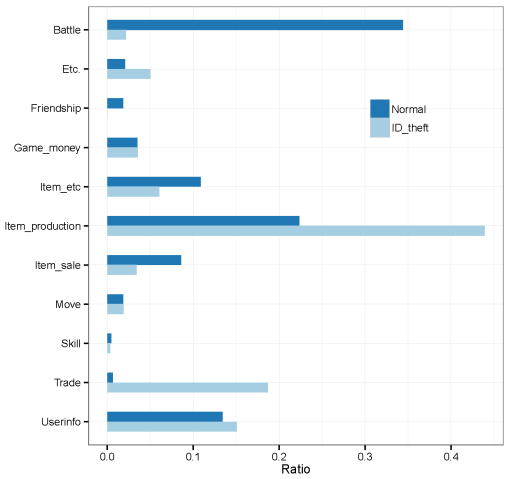

Fig. 4. Action log ratio of the normal user and account thieves

평균적으로 게임 머니량이 30% 이상 증가하고, 50% 이상 감소하였기 때문에 30% 이상 증가, 50% 이상 감소 횟수를 feature로 선택하였다. 더불어 로그인-로그아웃 동안 발생한 게임 머니 증가량과 감소량을 포함시켰다. 일반적으로 계정도용 시에는 아이템을 얻을 때와 아이템 채집을 수행하지 않으며, 아이템을 판매하기 위해 아이템 생성, 장비장착, 장비해제, 아이템 추출 등을 중점으로 수행한다. 로그인-로그아웃 동안의 아이템 생성, 장비장착, 장비해제, 아이템 추출 횟수를 feature로 선택하였다. 아이템 판매에서는 상인 NPC, 판매대행에 맡긴 아이템 수량이 전체 인벤토리(게임 내 아이템을 저장하는 창고)량의 90% 이상일 때의 횟수와 개인상점로그를 feature로 선택하였다. 마지막으로 거래횟수를 feature로 선택하였다.

세 번째 단계에서는 계정도용 사례 수가 정상 수보다 현저히 적기 때문에 정상 데이터들에 대해 Undersampling을 수행 후 학습 데이터를 구축하였다. 3가지 데이터마이닝 학습 방식(Multilayer Perceptron, Logistic, Random Forest)을 수행하였고, 10-folds Cross-validation 방식으로 훈련시켰다. 마지막으로 성능평가를 위하여 Precision(정밀도), Recall(재현율), Accuracy(정확도) 3가지 평가지표를 사용하였으며, 3가지 지표에 대한 정의는 식 (1)과 같다. Precision은 탐지된 결과들이 실제로 계정도용이 맞는가에 대한 비율이며, Recall은 실제 계정도용이 탐지모델을 통해 얼마나 탐지되는가를 나타내는 비율이다. Accuracy는 전체 사용자들에 대해 제안하는 모델이 얼마나 정확하게 예측했는가에 대한 지표이다.

$$\begin{aligned} \text{Precision} &= \frac{TP}{TP+FP} \\ Recall &= \frac{TP}{TP+FN} \\ Accuracy &= \frac{TP+TN}{total} \end{aligned} \quad (1)$$

TP: $True\ Positive$ TN: $True\ Native$
FP: $False\ Positive$ FN: $False\ Native$

IV. 실험 및 평가

4.1 계정도용자들의 행위 분석

2010년 6월 25일부터 2010년 7월 4일에 엔씨소프트사의 아이온에서 발생한 총 23건의 계정도용에 대하여 분석을 수행하였다.

4.1.1 계정도용자들의 행위분석 및 지출비율

계정도용자들의 게임 내 로그 비율을 파악하기 위해 다양한 로그들을 Table 2.에 나타낸 것처럼 분류하여 분석하였다. Fig.4.는 계정도용자들과 정상 사용자들의 게임 내 로그 비율로 계정도용자들의 경우 게임 내 친목(길드, 파티 등), 전투 등의 행위를 거의 수행하지 않고, 경제활동(아이템 생산, 거래)에만 치중 되어 있는 것을 확인할 수 있다.

계정도용 사용자들의 경우 게임 머니 및 아이템 탈취가 목적이기 때문에 원 사용자의 지출비율보다 높을 것으로 추측할 수 있다. 지출비율이란 한명의 사용자가 하루 동안 소비한 게임 머니 비율로 하루 동안 사용한 자산/(해당날 게임 시작시의 자산 + 하루 동안 사용한 자산) 으로 계산하였다. 계정도용 피해를 당한 23명의 일주일간 지출비율을 분석한 결과 일반적으로 20%~25% 이내의 지출비율이 발생하지만 계정도용 피해일의 경우 70% 정도의 지출비율이 발생하는 것을 Fig.5.에서 확인할 수 있다. Fig.5.에서 다이아몬드는 평균을 의미하고 직선은 중앙값, 박스는 지출비율을 보여준다.

행위분석과 지출비율 분석을 통해 계정도용 사용자들의 경우 사용자의 게임 머니와 아이템을 판매하는 경제활동을 수행한 후 게임 머니를 탈취하는 것을 확인할 수 있었다.

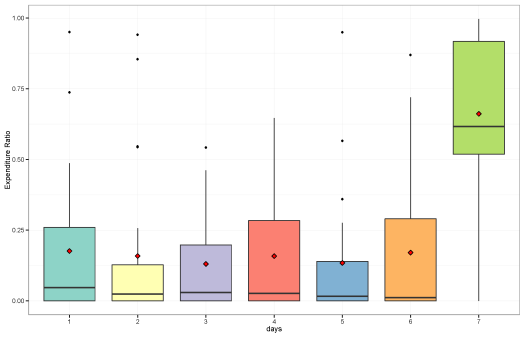

Fig.5. Expenditure Ratio

4.1.2 계정도용자들의 행위 시퀀스 분석

계정도용이 발생하였을 때 접속한 IP address를 추출하여 피해계정이 해당 IP address들로 접속한 동안의 모든 로그들을 추출하였다. 총 9개의 IP address에서 82번의 로그인-로그아웃이 발생하였으며, 59개의 로그가 8,332번 호출되었다. 호출된 횟수가 1번이거나 시스템 상 자동으로 기록되는 18개의 로그들을 제외하고, 계정도용자들의 시퀀스 분석을 수행하였다. Table3.는 호출된 로그들을 유사한 행위를 수행하는 그룹으로 분류하고, 이를 고유한 문자로 1:1 대응시킨 것이다.

시퀀스 분석을 위해 생물정보학에서 사용되는 유전자 서열정렬(Sequence Alignment)을 사용하였다. 서열정렬이란 유전자 등의 서열을 배열하고 유사도를 측정하여 생물체의 종을 구분하는데 사용하는 방식이다[19].

Table 3. Labeling table used for sequence analysis

Category	Character	Category	Character
Login	A	Item Production	G
Logout	B	Item purchased	H
User info	C	Item sale	I
Movement	D	Trade	J
Decrease the game money	E	etc.	K
Increase the game money	F		

82개의 로그인-로그아웃에 대한 특정 시퀀스는 발견되지 않아 게임 실행시간에 따라 분류하여 시퀀스 분석을 수행하였다. Table 4.는 총 82개의 계정도용의 게임 실행시간에 따른 비율로 88%가 10분 이내에 계정도용 피해를 입은 것을 확인할 수 있다.

Fig.6.는 유전자 서열정렬에서 일반적으로 사용되

Table 4. Comparison of the game time by identity theft

Type	Count	Rate(%)	Note
Less than 1 minute	32	39%	
Less than or equal to 5 minute	33	40%	
Less than or equal to 10 minute	7	9%	
More than 10 minute	10	12%	less than 1hour:8 more than 1hour:2

Table 5. 4 types of sequence

Type	Sequence
Less then 1 minute	AB, AGGGB
Less then or equal to 5 minute	AGGGGIIIIIIIIIEGGGGGGGEB
Less then or equal to 10 minute	AGGGGDDEIIIIIIIIJIEIIIIIEGGGGGEIIJIIIEB
More then 10 minute	AGGGGGDGGGGEIIIIIIIIEIIIIIIIIJIEIIIIIIIIIEGGGGGEIIIIIJGGGGGIIIIGGGIIEIIIIIJIIEIIIIIIEB

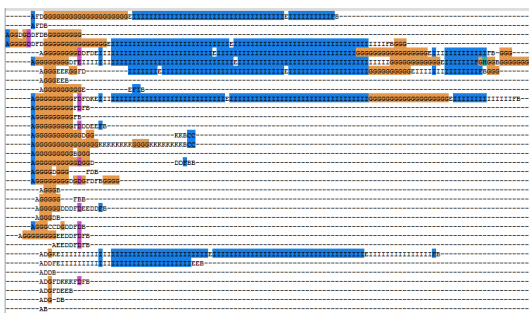

Fig. 6. sequence for type of less than or equal to 5 minute

는 ClustalX2 자동화 도구의 결과 화면이며, Table 5.는 시간분류별 추출된 시퀀스를 나타낸 것이다[20]. 1분 미만 유형의 경우 로그인(A)과 로그아웃(B)을 반복하는데 이는 계정도용자가 실 소유자(실제 계정 소유자)가 로그인을 하고 있는지 확인하기 위한 선행행위로 볼 수 있다. 10분 이하 유형의 경우 5분 이하 유형의 시퀀스가 반복됨을 확인할 수 있으며, 계정도용 시간이 10분을 초과한다 하더라도 5분 이하 유형의 패턴들이 반복적으로 진행됨을 확인할 수 있다. 즉, 로그인(A)후 판매하기 위해 아이템들을 생성(G)하고, 아이템 판매(I)나 거래(J)를 수행하며, 아이템 판매로 인해 게임 머니가 증가(F)하게 되지만 바로 게임 머니가 감소(E)하는 시퀀스가 반복적으로 나타났다.

시퀀스 분석 결과 계정도용자들의 경우 '아이템을 판매하기 위한 아이템 생성(생성, 수집, 강화 등) → 생성한 아이템 판매 및 거래 → 게임 머니 획득' 순으로 진행되는 것을 확인 할 수 있었다.

4.2 평가

시퀀스 분석을 통해 계정도용 시간이 길어지면, 특정 시퀀스를 반복하는 것을 확인할 수 있었다. 이를 기반으로 1분 미만, 5분 이하, 10분 이하, 10분 초과로 유형을 구분하여 평가를 수행하였다. 6월 25일부터 7월 4일까지의 전체로그를 로그인-로그아웃 동안으로 구분하였으며, 해당기간 동안 총 82번(Tabel.4 참고)의 계정도용 시도 및 수행이 일어났다. 계정도용 시도 및 수행으로 선택한 82번의 경우 23건의 계정도용이 신고된 시간에 접속한 IP address와 시간간격을 고려하여 계정도용이 일어난 것으로 간주하였다.

Undersampling을 통해 정상적인 사례를 추출하고 계정도용 사례와 함께 학습 데이터를 구축하였다. 데이터는 Table.6.에 나온 변수들로 구성되어 있으며, 결과는 1분 미만, 5분 이하, 10분 이하, 10분 초과, 정상 다섯가지로 분류하였다.

Random Forest가 가장 좋은 성능을 보였으며, Table 7.은 Multilayer Perceptron, Logistic, Random Forest를 통해 82개의 계정도용을 분류한 결과이다. Table 8.은 Random Forest의 유형별 탐지 성능 결과표로 1분 미만은 ID1, 5분 이하는 ID2, 10분 이하는 ID3, 10분 초과는 ID4를 의미한다.

계정도용의 경우 정상 사용자가 계정도용자로 오해받을 경우 회사에 손해배상을 요구할 수 있기 때문에 Recall보다는 Precision이 더 중요하다. 평균 Precision의 값이 0.84이지만, ID2(5분 이하 유형)를 제외하고는 높은 값을 보이고 있다. 특히 1분 미만의 경우 계정도용자가 실 소유주의 접속 여부를 파악하기 위해 수행하는 선행 행위이므로 해당 계정도용 부분을 탐지한다면, 이후 발생할 수 있는 실제 게임 아이템 및 게임 머니 탈취를 예방할 수 있다.

정확도를 높이기 위해 향후 IP address, MAC address, 친구 네트워크와 거래 네트워크의 상관관계 등의 보조지표와 교차분석을 수행할 예정이며, 이는 계정도용 탐지 정확도를 향상 시킬 것으로 기대된다.

Table 6. Features

Category	Features
Userinfo	Level, Playtime, Experience point
Movement	Number of movements
Game money	Decrease the number of game money, Decrement of game moeney, Increase the number of game money, Increment of game money
Item production	Gain, Collection, Production, Item installation, Item uninstallation, Extraction, Number of extract
Item sale	NPC shop, User shop, sales agency
Trade	Trade, Moving between inventory

Table 7. Classification error table

	Multilayer Perceptron	Logistic	Random Forest
Precision	0.676	0.710	0.844
Recall	0.665	0.682	0.835
Accuracy	66%	68%	84%

Table 8. Classification error table for Random forest

Class	Precision	Recall
ID1	0.914	0.914
ID2	0.706	0.75
ID3	1	0.714
ID4	1	0.545

4.3 계정도용자들의 거래 네트워크 분석

일반적으로 게임 상에서 거래란 게임 사용자들 간에 아이템을 주고, 그에 상응하는 아이템이나 게임 머니를 주는 것을 말한다(양방향 거래, two-way trade). 계정도용자들의 거래 네트워크 분석결과 특정 위치에서 아이템을 주기만 하고 그에 상응하는 대가를 받지 않는 단방향 거래(one-way trade)가 이루어지는 것을 확인할 수 있었다. 일반 사용자들 역시 단방향 거래를 수행할 수 있기 때문에 실험기간동안의 단·양방향 거래 비율에 대해 분석을 수행하였다. 계정도용 시 발생한 단방향 거래 비율은 100%이며, 정상 사용자들의 단방향 거래는 81%, 양방향 거래는 19%로 나타났다.

계정도용 거래 네트워크를 분석하기 위해 제재 IP 리스트에 포함된 IP address로 거래를 수행한 그룹(이하 계정도용 그룹)과 수행하지 않은 그룹(이하 의심스러운 그룹)으로 분류하여 분석하였다. 피해계정들의 총 거래는 55건으로 계정도용 그룹이 43건, 의심스러운 그룹이 12건의 거래를 수행하였다. 55건 모두 단방향 거래였으며, Fig.8.은 계정도용 그룹의 거래 네트워크를 Fig.9.는 의심스러운 그룹의 거래 네트워크를 보여준다. 동그라미안의 숫자는 게임 사용자들을 나타내며, edge 위의 값은 거래 날짜와 거래를 수행한 장소(날짜_거래위치)를 나타낸다.

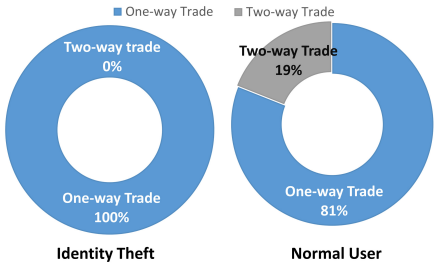

Fig. 7. Rate of one-way and two-way trade

Fig.8.(a), (b), (f)의 경우 오른쪽의 사용자가 왼쪽의 사용자에게 아이템 및 게임 머니(이하 게임 재화)를 준 것을 나타낸다. Fig.8.(c)의 경우 U8 사용자에게 5명의 사용자(U6, U7, U8, U9, U10)가 게임 재화를 주고, 최종적으로 U8 사용자가 팔각형의 U12 사용자에게 게임 재화를 준 것을 확인할 수 있다.

계정도용 거래의 특징은 한정된 장소에서 계정도용 피해자들이 다른 계정도용 피해자(이하 피해계정)에게 단방향 거래를 수행한다는 점이다. 계정도용 그룹은 총 4곳 A, B, C, D에서 거래를 수행하였는데 A와 C는 피해계정들이 다른 피해계정에게 게임 재화을 준 장소이며, B와 D는 다른 피해계정들로부터 게임 재화를 받은 피해계정이 다른 계정에게 최종적으로 거래를 수행한 장소이다. 거래가 수행된 장소가 4곳 이지만 게임 재화를 받는 장소와 최종적으로 탈취한 게임 재화를 주는 장소로 구분할 수 있다. 이는 일반 계정들의 단방향 거래의 장소의 수가 계정도용자들의 거래 장소의 수와 많은 차이를 나타내는 것을 알 수 있다. Fig.8.의 Group A는 피해계정들이 다른 피해계정에게 게임 재화를 준 장소가 A인 그룹이며, Fig.8.의 Group B는 거래 수행 장소가 C인 그룹이다. Fig.8.(c), (d)의 경우 여러 피해계정들로부터 아이템을 받은 피해계정이 마지막으로 다른 장소에서 다른 사용자에게 아이템을 주는 것으로 나타났다. 즉, 계정도용자들은 동시에 여러 계정들을 탈취한 뒤 한

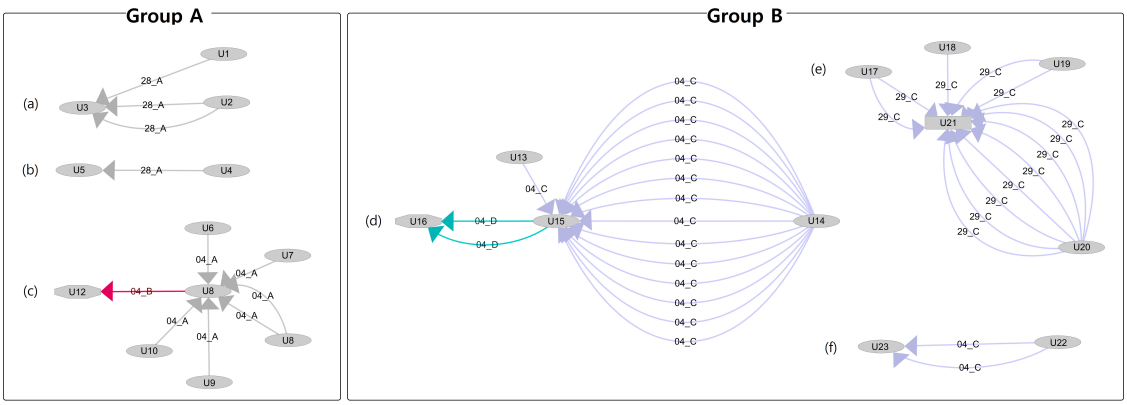

Fig. 8. Trade networks of identity theft group

개의 피해계정에 다른 피해계정들의 아이템을 모아 판매하는 것으로 추측할 수 있다. 게임 봇은 게임 재화를 수집하는 gold-farmer 그룹과 gold-farmer 그룹으로부터 자산을 받아 게임 머니화를 수행하는 merchant 그룹, 수집한 게임 머니만을 소유하는 banker 그룹으로 구분할 수 있다[21]. 한 계정에 게임 재화를 모아 다른 계정에서 최종적으로 주는 행위는 게임 봇의 3-depth 중 하나인 merchant 역할을 수행하고 있는 것으로 볼 수 있다.

Fig.9.의 거래를 살펴보면, 오른쪽에서 왼쪽에게 게임 아이템 및 게임 머니를 준 것을 볼 수 있다. Fig.9.(a)의 경우 왼쪽의 U2 사용자가 오른쪽의 U24 사용자에서 6월 28일에 E 장소에서 4번의 거래를 수행한 것을 보여준다. Fig.9.의 Group A는 계정도용 그룹과 거래 장소가 상이한 거래들이며, Group B는 계정도용 그룹과 거래장소가 1건이라도 동일한 거래들이다. 의심스러운 그룹의 경우 계정도용 시의 IP address로 거래를 수행하지 않은 그룹으로 Fig.9.에서 Group A의 경우 계정도용그룹과 장소 위치가 상이하기 때문에 실 소유주가 거래를 수행한 것으로 볼 수 있다. Fig.9.에서 Group B의 경우 실 소유자가 거래를 수행한 것을 확인하기 위해 IP address를 추적하였다. Fig.9.(e)와 (f)의 경우 일주일 동안에 접속한 IP address 중 하나로 접속하였기 때문에 계정도용과는 상관없는 것으로 추측할 수 있다. Fig.9.(g)의 경우 일주일동안 접속한 IP address가 아닌 IP address로 거래를 수행 하였지만, 계정도용 그룹이 최종적으로 거래를 수행한 D 에서 거래를 수행한 것으로 보아 이는 다른 계정도용자가 해당 계정으로 로그인하여 아이템을 탈취한 것으로 볼 수 있다. Fig.10.은 U21 사용자와 관계된 Fig.8.의 (e)와 Fig.9.의 (g)를 하나의 그림으로 나타낸 것이다. 즉, Fig.8.의 (c), (d)와 같이 탈취된 계정들로부터 게임재화들을 받고, 다른 사용

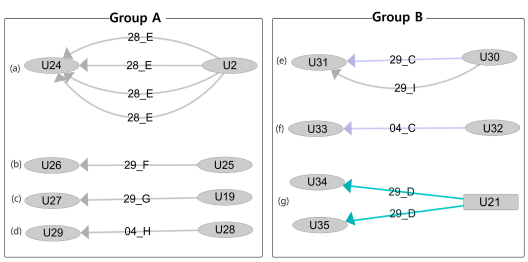

Fig.9. Trade network of suspicious group

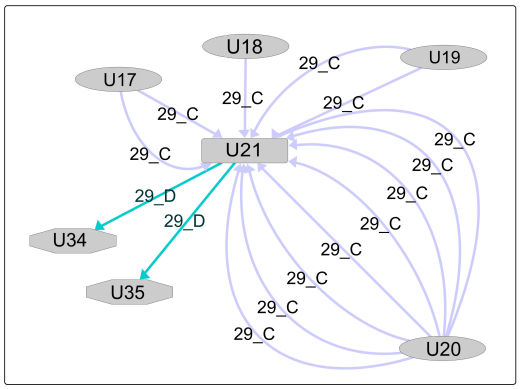

Fig. 10. Trading network of U18

자에게 게임재화를 전달하는 것을 볼 수 있다.

거래 네트워크 분석을 통해 계정도용자들이 탈취한 하나의 계정을 이용하여 특정 장소에서 아이템들을 모으고, 마지막에 탈취된 아이템들을 이동시키는 것을 확인할 수 있었다.

V. 결론 및 향후 계획

문화 콘텐츠 수출 중 과반수이상을 차지하는 게임 산업은 현금화가 용이하다는 특성으로 인해 계정도용에 노출되어 있으며, 증가하는 개인정보 유출로 인해 빈번하게 발생되고 있는 실정이다. 계정도용은 사용자들의 불만을 고조시키는 불법행위의 하나로 게임회사와 게임 사용자 모두에게 금전적인 피해를 남기며, 게임회사의 이미지에도 영향을 미치게 된다. 이로 인해 게임시장 내 계정도용 탐지가 요구되고 있다.

본 논문에서는 계정도용을 시간별로 분류하여 탐지하는 프레임워크를 제시하였으며, 계정도용 거래 네트워크 분석을 수행하였다. 분석결과 계정도용자들은 아이템 생성, 아이템 판매, 게임 머니 획득이라는 특정 시퀀스를 가지며, 해당 시퀀스는 시간과 상관없이 반복하는 것을 파악하였다. 이를 기반으로 제시하는 탐지모델의 경우 84%의 정확도를 보였다. 더불어 계정도용이 실제로 발생하기 전의 선행행위로 분류한 '1분 미만' 유형에 대해 높은 탐지율을 보였으며, '1분 미만' 유형이 탐지되었을 경우 사용자에게 인증을 요구한다면 계정도용을 예방할 수 있을 것으로 기대된다.

계정도용 로그 분석결과 게임 봇과의 연관성들을 발견하였다. 첫 번째로 계정도용 피해자인 23명의 사용자 모두에게서 게임 봇을 사용한 흔적을 발견하였

다. 이러한 게임 봇 사용 흔적은 게임 봇이 게임 사용자들의 개인정보를 탈취할 수 있다는 가능성을 나타낸다. 두 번째로 계정도용 거래 네트워크 분석을 통해 탈취한 계정들 중 한 개의 계정을 merchant 로 사용하는 것을 파악하였다. 이는 게임 봇 그룹과 높은 연관성을 가지는 것으로 추측할 수 있다.

향후 기간을 추가하여 친구 & 거래 네트워크 분석한 후 확산모델에 대해 연구를 수행할 예정이며, 보조지표를 이용하여 정확도를 높일 예정이다.

References

[1] KISA, http://isis.kisa.or.kr/sub07/?pageId=070500#

[2] This is game, http://www.thisisgame.com/webzine//nboard/4/?outkey=zkcaebvn6rxg3m4ur1382qdqd9zmwy9fq8s438tf&page=312&n=25415, Jul. 2011

[3] Police Science Instituete, Police Science Institute Review no.39, Feb. 2014

[4] Boannews, http://www.boannews.com/media/view.asp?idx=42022&kind=1, Jul. 2014

[5] Sisaon, http://blog.sisaon.co.kr/130184176950, Jan. 2014

[6] IT Today, http://www.ittoday.co.kr/news/articleView.html?idxno=58787, Mar, 2015

[7] Hankyoreh, http://www.hani.co.kr/arti/economy/economy_general/105153.html, Feb, 2006

[8] Asia Economy Daily, http://www.asiae.co.kr/news/view.htm?idxno=2013030613360916782, Mar. 2013

[9] KOCCA, 2014 Korea, Japan Gamers Survey Report, May. 2014

[10] Daniele Gunetti and Claudia Picardi, "Keystroke analysis of free text," ACM Transactions on Information and System Security, vol. 8, no. 3, pp. 312-347, Aug. 2005

[11] Hojin Seo and Huy Kang Kim, "User input pattern-based authentication method to prevent mobile e-Financial incidents," Parallel and Distributed Processing with Applications Workshops, pp. 382-387, May. 2011

[12] Chen Kuan-Ta and Li-Wen Hong, "User identification based on game-play activity patterns," Proceedings of the 6th ACM SIGCOMM workshop on Network and system support for games, pp. 7-12, Sept. 2007

[13] Jehwan Oh, Zoheb Hassan Borbora, and Jaideep Srivastava, "Automatic detection of compromised accounts in mmorpgs," Social Informatics, pp. 222-227, Dec. 2012

[14] Hwa Jae Choi, Jiyoung Woo, and Huy Kang Kim, "Online game identity theft detection model based on hacker's behavior," Journal of Korea Society, Vol. 11 No.6. pp. 81-94, Dec. 2011

[15] Jiyoung Woo, Hwa Jae Choi, and Huy Kang Kim, "An automatic and proactive identity theft detection model in MMORPGs," Applied Mathematics & Information Sciences, Vol. 6, No. 1, pp. 291-302, Jan. 2012

[16] Paul Madsen, Yuzo Koga, and Kenji Takahashi, "Federated identity management for protecting users from ID theft," Proceedings of the 2005 workshop on Digital identity management, pp. 77-83, Nov. 2005

[17] Neil Chou, Robert Ledesma, Yuka Teraguchi, and John C. Mitchell, "Client-Side defense against web-based identity theft," 11th Annual network and Distributed System Security Symposium, Feb. 2004

[18] Alexander Tsoutsanis, "Tackling twitter and facebook fakes: ID theft in social media," World Communications Regulation Report, Vol. 7, No. 4, pp. 1-3, Mar. 2012

[19] Youngjoon Ki, Eunjin Kim, and Huy Kang Kim, "A novel approach to detect malware based on API call sequence analysis,"

International Journal of Distributed Sensor Networks, 501, 659101, Apr. 2015.
〔20〕 ClustalX2, ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/
〔21〕 Hyukmin Kwon, Kyungmoon Woo, Hyun-chul Kim, Chong-kwon Kim, and Huy Kang Kim. "Surgical strike: A novel approach to minimize collateral damage to game BOT detection," Proceedings of Annual Workshop on Network and Systems Support for Games, IEEE Press, pp. 1-2, Dec. 2013

〈저자소개〉

김 하 나 (Hana Kim) 학생회원
2013년 2월: 서울여자대학교 정보보호학과 졸업
2013년 3월~현재: 고려대학교 정보보호학과 석사과정
〈관심분야〉 온라인게임 보안, 데이터 마이닝

곽 병 일 (Byung Il Kwak) 정회원
2013년 2월: 세종대학교 컴퓨터공학과 졸업
2013년 9월~현재: 고려대학교 정보보호학과 석·박사통합과정
〈관심분야〉 온라인게임 보안, 데이터 마이닝, 네트워크 보안, IoT 보안

김 휘 강 (Huy Kang Kim) 종신회원
1998년 2월: KAIST 산업경영학과 학사
2000년 2월: KAIST 산업공학과 석사
2009년 2월: KAIST 산업및시스템공학과 박사
2004년 5월~2010년 2월: 엔씨소프트 정보보안실장, Technical Director
2010년 3월~2014년 12월: 고려대학교 정보보호대학원 조교수
2015년 1월~현재: 고려대학교 정보보호대학원 부교수
〈관심분야〉 온라인게임 보안, 네트워크 보안, 네트워크 포렌식