Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.24, NO.2, Apr. 2014 ISSN 2288-2715(Online)
http://dx.doi.org/10.13089/JKIISC.2014.24.2.431

페이스북 공개 정보를 이용한 사용자 출생지 추론*

최 대 선,[1†] 이 윤 호[2‡]
[1]한국전자통신연구원, [2]서울과학기술대학교

Inference of birthplaces of users with public information in FaceBook*

Daeseon Choi,[1†] Younho Lee[2‡]
[1]Electronics and Telecommunications Research Institute, [2]SeoulTech

요 약

본 논문에서는 페이스북 사용자들의 공개된 정보만으로, 그들의 출생지 정보를 추론할 수 있음을 보인다. 다양한 기계학습 알고리즘 및 노출 정보들의 조합을 통한 실험을 통해, 지지벡터기계 알고리즘 및 졸업고등학교소재지, 현 주소, 고등학교 졸업연도가 추론의 가장 최적의 성능을 나타냄을 발견하였고, 약 78%의 출생지 정보를 추론할 수 있었다. 출생지 정보는 패스워드 분실시 복구에 사용되는 질문에 자주 사용되고, 또한 주민등록번호의 일부를 이루는 중요한 정보이므로, 사용자들은 이러한 사실에 주의하여 페이스북을 사용하는 것이 필요하다.

ABSTRACT

This paper shows the users' birthplace information can be inferred with only the public information in FaceBook SNS. Through experiments with various machine learning algorithms and various parameters, we have found that SVM algorithm with the location of the highschool, the current address, and the graduate year of highschool performs best for the inference, as this can infer 78% of users' birthplaces correctly. Since the birthplace information is used for various security purpose such as questions for getting the forgotten password and a part of korean residence registration number, this is a non-trival security breach and users need be cautious about it.

Keywords: SNS, Privacy, Personal information, Data inference, Security

I. 서 론

스마트폰의 대중화는 사람들로 하여금 사회 관계망 서비스를 항시적으로 사용할 수 있는 계기를 마련하였다. 페이스북(Facebook), 트위터(Twitter) [1,2]와 같은 범 세계적인 사회 관계망 서비스는 많은 한국 사용자들이 존재한다.

이러한 사회 관계망 서비스의 사용은 사용자들의 사회 관계의 항시성 과 접촉 빈도를 높이는 장점을 낳았으나, 반대 급부로 자신의 개인 정보에 대한 노출을 야기시켰다. 그러나 사용자들은 자신이 공개한 정보의 가치와 위험도에 대해서는 인지를 못하고 있는 경향이 강하다.

본 연구에서는 페이스북 사용자들의 공개정보를 이용하면 사용자들의 출생지 정보(시(도) 수준)를 추론할 수 있음을 보인다. 구체적으로, 본 연구에서는 사용자가 일반적으로 노출시키는 정보들을 분석하였고, 그것들의 다양한 조합을 다양한 기계 학습 추론 알고리즘을 통하여 어떠한 정보가 출생지 정보를 추론할 때 유용한지 검증하였다. 검증 결과, 지지학습기계(Support Vector Machine:SVM) [3] 알고리즘을 이용하여, 졸업 고등학교 소재지, 현재 직장, 그리

접수일(2014년 1월 23일), 게재확정일(2014년 4월 2일)
* 이 연구는 서울과학기술대학교 교내 일반과제 연구비 지원으로 수행되었습니다.
† 주저자, sunchoi@etri.re.kr
‡ 교신저자, younholee@seoultech.ac.kr (Corresponding author)

고 고등학교 졸업 년도를 이용하여 학습 및 추론을 할 경우 가장 높은 추론 성능을 보인다는 것을 알아냈으며, 실험 결과 약 76.43%의 출생지를 성공적으로 추론할 수 있었다. 이 때 사용한 매개 변수는 svm_type=Nu_SVC, kernel type = RBF, C=1, nu=0.5, p=0.1, gamma= 0.0, degree = 3, coef0=0, shrinking=True, eps=0.001, normalization = True 이다.

출생지 정보는 사용자 비밀번호 확인 시 자주 사용되는 질문 중의 하나이며, 주민등록번호 코드의 일부를 이루는 정보이다. 따라서 이러한 정보가 노출된다면, 사용자의 보안성에 간접적인 피해를 야기하므로, 추론에 사용된 정보를 선택적으로 숨기는 것이 자신의 출생지 정보의 노출을 회피하는 데 도움을 줄 것이라 생각한다.

참고로 사회 관계망 서비스의 개인정보 위험 및 추론에 관한 연구는 많이 있었으나 [4-11], 출생지 정보에 대한 내용은 없었다.

본 논문의 구성은 다음과 같다. 2절에서 출생지 정보 추론 과정 및 실험에 대해 설명하며, 3절에서 결론을 맺는다.

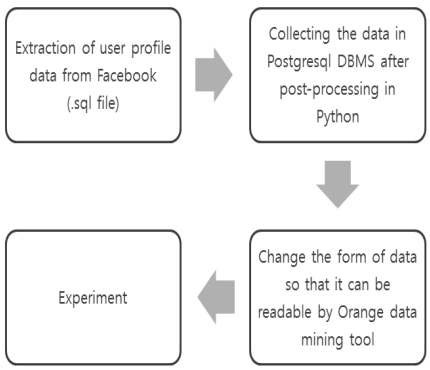

Fig.1. Experiment procedure

II. 출생지 추론

2.1 환경

본 항에서는 데이터 수집 방법과 구현 환경에 대해 기술한다. 데이터는 전자통신연구원에서 크롤링 도구를 이용하여 습득하였다. 우리는 이 중 약 3만여개의 데이터를 사용하였다. 또한 각 사용자들은 선정된 데이터 내에서 친구가 최소 1명 ~ 최대 10명까지 존재한다. 기술적인 이유로 10명 이상의 친구들은 포함시키지 못하였다. 각 데이터는 많은 필드가 존재하나, 우리는 다음과 같은 데이터만을 사용하였다. 전달받은 데이터는 postgreSQL DBMS [12]에서 가공을 수행하였다.

기본적인 실험 과정은 Fig. 1과 같다. 실험에서는 Orange [13] 에서 제공하는 Visual Programming 방법을 그대로 사용하여 실험을 수행하였다. Visual Programming 의 내용은 그림 2와 같다.

수행 시 사용된 데이터의 예는 다음과 같다. 다양한 항목이 존재하나, 우리는 그 중 일부를 선택하여 사용하였다. 추출한 항목은 ID, 성별, 국적, 혈액형, 친구 ID 리스트, 종교, 정치성향, 취미, 선호 스포츠 팀 연고지, 졸업 고등학교 소재지, 졸업 대학 소재지, 생일(년월일) 등이다. 물론 모든 사용자에 대해 모든 항목이 존재하지 않고, 존재하는 항목이 사용자마다 편차가 큰 특징이 있다.

2.2 지역 분류기 (local classifier) 를 이용한 추론 실험

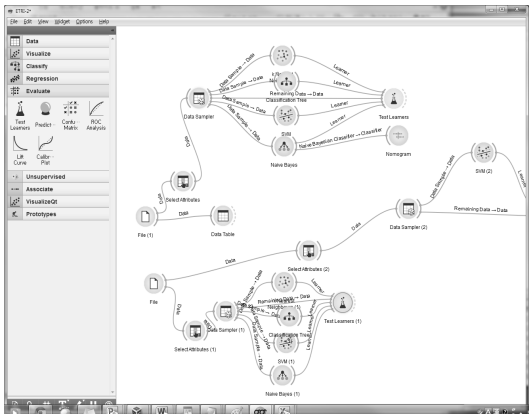

Fig.2. Implementation for the experiment by Visual programming in Orange tool.

본 세부 절에서는 첫 번째 추론 과정 및 결과를 나타낸다. 본 추론 과정에서는 기존의 지역 분류기만을 이용하여 많은 사용자의 지역 정보를 추론하기로 한다. 추론의 아이디어는 다음과 같다. 우리가 생각한 추론 아이디어는 다음과 같다. 많은 경우 자신의 고등학교는 시(도) 범위 내에서 자신의 출생지와 동일한 고등학교를 나올 것이라는 가정에서 출발한다. 이와 유사하게 대학교 졸업 지역도 영향을 미칠 것이라 생

각되며, 마지막으로 현 주소지까지 포함한다면, 출생지를 유추하는데 좀더 확률을 높일 것으로 예상하였다.

Table 1. Performance measure

Peformance measure	Description
Accuracy	(# of correct inferences)/(# of total inference)
F measure	Harmonic average of Recall and Precision
Area under ROC curve	Used to measure the peformance in parameter-indepedent manner. Higher value means better peformance

Fig. 2에서 보면 알 수 있듯이 우리는 나이브 베이지안, SVM, kNN의 세가지 기계학습 알고리즘을 사용하였다. 85%의 데이터를 임의로 선정해 학습을 위해 사용하였으며, 나머지 15%를 테스트를 위해 사용하였다. 성능 평가를 위해 측정한 내용은 표 1에 표시되어 있다.

위와 같은 과정으로 수행한 결과는 아래의 Fig. 3에 나타난다. 이중 첫번째 그림 (a)는 다양한 데이터 필드를 이용하여 다양한 기계 학습 알고리즘을 이용하여 정확도(Accuracy)를 측정한 결과를 나타내고 있다. 그래프에서 보면 알 수 있듯이, SVM을 사용하면서, 동시에 고등학교 졸업 지역, 현 주소지, 그리고 대학교 졸업 지역을 이용하여 추론한 결과가 가장 높게 나타남을 알 수 있다. 마찬가지로 그 아래에 나타나 있는 그림 (b), (c)는 Area Under ROC 와 F measure를 이용해 측정한 결과를 나타내고 있다. 아래의 그래프들도 동일한 데이터 조합에서 동일한 SVM 추론 알고리즘 적용 시 가장 높은 성능을 나타냄을 알 수 있다. 결론적으로 SVM 알고리즘을 이용하여서 고등학교 졸업지역, 대학교 졸업 지역, 그리고 현 주소지를 이용하여 학습하였을 때, 가장 높은 확률로 출생 시(도) 를 추론함을 알 수 있었다.

III. 결 론

본 논문에서는 페이스북 사용자들이 노출한 공개된 정보를 이용하여 사용자들의 출생지 시(도) 정보를 단지 공개된 정보만으로 높은 확률로 추론할 수 있음을 보였다.

본 연구의 결과로써, 사람들이 자신의 개인정보 노출의 심각성을 이해하고, 이를 바탕으로 SNS에 정보를 노출 시에는 경각심을 갖고 주의하여, 온라인에서의 불법적인 행위로 인한 명예/재산상의 손해를 받지 않도록 하는 것에 기여할 것이라 생각된다.

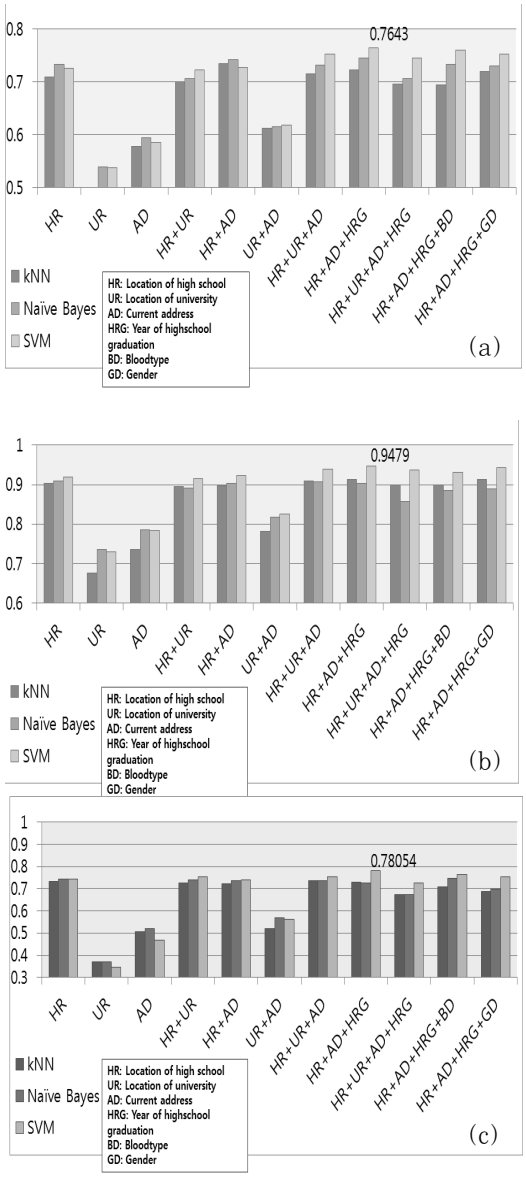

Fig.3. Experiment result: (a) Accuracy (b) F measure (c) Area under ROC

References

〔1〕 Facebook. https://www.facebook.com/
〔2〕 Twitter. https://twitter.com/
〔3〕 C. Cortes and V. N. Vapnik, "Support-Vector Networks," Machine Learning, vol. 20, no. 3, pp. 273-297, 1995.
〔4〕 S. Choi, et al, "A technique for analyzing personal information leakage threat in Bigdata," Journal of Korean Institute of Information Security & Cryptology 23(3), pp. 56-60, 2013.
〔5〕 K. Liu, E. Terzi, "A Framework for Computing the Privacy Scores of Users in Online Social Networks," ACM Transactions on Knowledge Discovery from Data 5(1) Article No. 6, Dec, 2010.
〔6〕 E. Zheleva and L. Getoor, "To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles," in Proc. 18th WWW, pp. 531-540, 2009.
〔7〕 C. Zhang, et al, "Privacy and security for online social networks: challenges and opportunities," IEEE Network, 24(4) pp. 13-18, Jul., 2010.
〔8〕 M. Hirose et al, "A Private Information Detector for Controlling Circulation of Private Information through Social Networks," In Proc. 7th ARES Conf, pp. 473-478, Aug., 2012,
〔9〕 Y. Yang et al, "Stalking online: on user privacy in social networks," In Proc. 2nd ACM Conf. CODASPY, pp. 37-48, 2012.
〔10〕 H. Mao, X. Shuai, and A. Kapadia, "Loose Tweets: an analysis of privacy leaks on twitter," In Proc. 10th ACM Workshop on Privacy in the electronic society, pp. 1-12, 2011.
〔11〕 R. Heatherly et al, "Preventing Private Information Inference Attacks on Social Networks," IEEE Transactions on Knowledge and Data Engineering, 25(8), pp.1849-1862, 2013.
〔12〕 PostgreSQL, an open-source based object-relational database system. http://www.postgresql.org
〔13〕 Orange: Open source data visualization and analysis for novice and experts. Data mining through visual programming or Python scripting. http://orange.biolab.si/

〈저자소개〉

최 대 선 (Daeseon Choi) 종신회원
1995년: 동국대학교 컴퓨터공학과 학사
1997년: 포항공과대학교 컴퓨터공학과 석사
2009년: KAIST 전산학과 박사
1997년~1999년: 현대전자/현대정보기술 연구소 선임
1999년 ~ 현재: 한국전자통신연구원 책임연구원/인증기술연구실장
〈관심분야〉 인증, 개인정보보호, 빅데이터 분석

이 윤 호 (Younho Lee) 종신회원
2000년/2002년/2006년: KAIST 전산학과 학사/석사/박사
2006년~2007년 KAIST 전자정보연구소 박사후연구원
2007년~2009년 GeorgiaTech Information Security Center Postdoc
2009년~2013년 영남대학교 정보통신공학과 조교수
2013년~현재 서울과학기술대학교 글로벌융합산업공학과 ITM 전공 조교수
〈관심분야〉 응용 암호, 데이터 보안, 개인정보보호