Privacy-concerns in The Era of AI and Social Network Sites

Xuan-Son Vu, Addi Ait-Mlouk, Erik Elmroth, Lili Jiang

[1]Department of Computing Science, Umeå University, Sweden;
[1]{sonvx, addia, elmroth, lili.jiang}@cs.umu.se

Abstract. Social Network Sites (SNS) (e.g., Facebook, Youtube), have been playing a great role in our lives. On one hand, they help to connect people in the way that would otherwise never possible before. Many recent breakthroughs in AI such as facial recognition were achieved thanks to the amount of available data in the Internet via SNS. However, on the other hand, they can have major impacts on people life, good or bad. Due to privacy concerns, many people have tried to avoid SNS [2] to protect their privacy. Similar to the security issue of the Internet protocol, Machine Learning - which is the core of AI, was not designed with privacy in mind. For instance, Support Vector Machines (SVMs) try to solve a quadratic optimization problem by deciding which instances of training dataset are support vectors. This means that the data of people involved in the training process will be also published within the SVM models. Recently, Fredrikson et al. [1] used hill-climbing algorithm on the output probabilities of a computer-vision classifier to reveal individual faces from the training data. Because of all these issues, privacy guarantees must apply to the worst-case outliers and thus will also destroy data utilities. From all above reasons, in my PhD project, we study on (1) how to protect privacy when learning predictive models and how to have a good trade-off between data utilities and privacy, to avoid privacy breaches such as Cambridge Analytical in the future.

1 Introduction

In the information age, online society on social network and the real social network are highly related. Many researchers have been studying user activities on Social Network Sites (SNS) (e.g., Facebook, Twitter) to understand user behaviour *. Recent thorough review of Waheed et al. [8] on 116 primary studies showed that *"activities performed on SNS are either associated with user behavior or reflect personality characteristics"*. This means that, understanding user activities on SNS is a key challenge to better support people for a sustainable society. A society that can assure its citizens equality, freedom and a healthy standard of living by understanding emotional changes and social behaviour of its citizens to better support them with privacy-guarantee (e.g., to avoid depression).

*There are ~3.2M results from Google Scholar for the keyword "social network and user behavior".

2 Methodologies

Our current research ranging from protecting privacy for data-sharing [7], predictive models [4], and to data analysis [6, 5, 3] including:

- Personality-Based Knowledge Extraction for Privacy-preserving Data Analysis [6].
- Self-adaptive Privacy Concern Detection for User-generated Contents [4].
- Generic Multilayer Network Data Analysis with the Fusion of Content and Structure [5].
- dpUGC: Learn Differentially Private Representation for User Generated Contents [7].
- Graph-based Interactive Data Federation System for Heterogeneous Data Retrieval and Analytics [3].

References

1. Fredrikson, M., Jha, S., Ristenpart, T.: Model inversion attacks that exploit confidence information and basic countermeasures. In: Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security. pp. 1322–1333. CCS '15 (2015)
2. Stieger, S., Burger, C., Bohn, M., Voracek, M.: Who commits virtual identity suicide? differences in privacy concerns, internet addiction, and personality between facebook users and quitters. Cyberpsychology, Behavior, and Social Networking **16**(9), 629–634 (2013). https://doi.org/10.1089/cyber.2012.0323, https://doi.org/10.1089/cyber.2012.0323, pMID: 23374170
3. Vu, X.S., Ait-Mlouk, A., Elmroth, E., Jiang, L.: Graph-based interactive data federation system for heterogeneous data retrieval and analytics. In: Demo Track, In: Proceedings of the The Web Conference 2019 (to appear). TheWebConf '19 - formerly WWW, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland (2019)
4. Vu, X.S., Jiang, L.: Self-adaptive privacy concern detection for user-generated content. In: Proceedings of the 19th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing), Vol. Volume 1: Long papers, p., March 2018 (2018)
5. Vu, X.S., Jiang, L.: Generic multilayer network data analysis with the fusion of content and structure. In: Proceedings of the 20th International Conference on Computational Linguistics and Intelligent Text Processing, April, 2019 (2019)
6. Vu, X.S., Jiang, L., Brändström, A., Elmroth, E.: Personality-based knowledge extraction for privacy-preserving data analysis. In: Proceedings of the Knowledge Capture Conference. pp. 45:1–45:4. K-CAP 2017, ACM, New York, NY, USA (2017). https://doi.org/10.1145/3148011.3154479, http://doi.acm.org/10.1145/3148011.3154479
7. Vu, X.S., Tran, S.N., Jiang, L.: dpugc: Learn differentially private representationfor user generated contents. In: Proceedings of the 20th International Conference on Computational Linguistics and Intelligent Text Processing, April, 2019 (2019)
8. Waheed, H., Anjum, M., Rehman, M., Khawaja, A.: Investigation of user behavior on social networking sites. PLOS ONE **12**(2), 1–19 (02 2017). https://doi.org/10.1371/journal.pone.0169693, https://doi.org/10.1371/journal.pone.0169693