SecureNLP: A System for Multi-Party Privacy-Preserving Natural Language Processing

Qi Feng, Debiao He, *Member, IEEE*, Zhe Liu, *Senior Member, IEEE*, Huaqun Wang, and Kim-Kwang Raymond Choo, *Senior Member, IEEE*

***Abstract*—Natural language processing (NLP) allows a computer program to understand human language as it is spoken, and has been increasingly deployed in a growing number of applications, such as machine translation, sentiment analysis, and electronic voice assistant. While information obtained from different sources can enhance the accuracy of NLP models, there are also privacy implications in the collection of such massive data. Thus, in this paper, we design a privacy-preserving system SecureNLP, focusing on the instance of recurrent neural network (RNN)-based sequence-to-sequence with attention model for neural machine translation. Specifically, for non-linear functions such as sigmoid and tanh, we design two efficient distributed protocols using secure multi-party computation (MPC), which are used to carry out the respective tasks in the SecureNLP. We also prove the security of these two protocols (i.e., privacy-preserving long short-term memory network PrivLSTM, and privacy-preserving sequence to sequence transformation PrivSEQ2SEQ) in the semi-honest adversary model, in the sense that any honest-but-curious adversary cannot learn anything else from the messages they receive from other parties. The proposed system is implemented in C++ and Python, and the findings from the evaluation demonstrate the utility of the protocols in cross-domain NLP.**

***Index Terms*—Secure multi-party computation, natural language processing, seq2seq with attention, long short-term memory.**

I. Introduction

THE renewed interest in natural language processing (NLP) is partly due to recent trends in artificial intelligence (AI), and the potential of NLP in applications involving textual data. Generally, as shown in Fig. 1, NLP tools and techniques can facilitate the processing, analysis, and interpretation of significant volume of data, and in turn informs decision making and strategy formulation. For example, Tractica [1] estimated that the NLP market (including the NLP hardware, software, and services) is likely to worth $22.3 billion by the year 2025, and NLP software profits increase from $136 million in 2016 to $5.4 billion by the year 2025.

There are, however security and privacy concerns associated with NLP-enabled data analytics applications [2]–[5]. For example, as more data are collected from both our physical and cyber environments, there is a need to ensure private information about the users and their environments is securely communicated and stored, as well as the analysis of such data does not infringe on user privacy. The importance of data security and privacy is partly evidenced by the introduction of exacting privacy regulations, such as the General Data Protection Regulation (EU) 2016/679 (GDPR), whose clauses formally regulate the usage specification on user data, particularly relating to data security and privacy. Hence, in this paper we seek to preserve user privacy without affecting data utility in NLP services.

To achieve privacy preservation in NLP, a common approach is to utilize homomorphic encryption (HE), which allows (encrypted) data to undergo certain arithmetic operations (e.g, addition and multiplication) without the need to decrypt or reveal the underlying data. Hence in recent years, several HE-based building blocks have been proposed. Examples include SecureLR designed by Jiang *et al.* [6] and the computation toolkits designed in [7], [8].

Another approach to achieve privacy preservation is to use multi-party computation (MPC), where computations are performed on the secret inputs from various parties (such as several service providers, users and so on). The parties are not able to learn any information related with others' inputs, with the exception of what can be learned from the output. The output can be made available to all involved parties. As for the NLP tasks, for example, MPC can be leveraged when keywords from one article do not have sufficient scope points to perform tagging using NLP, but a combination of data from several filesets will make the NLP works well.

Manuscript received March 15, 2020; revised May 4, 2020; accepted May 19, 2020. Date of publication May 25, 2020; date of current version July 9, 2020. This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFC1604000, and in part by the National Natural Science Foundation of China under Grant 61932016, Grant 61972294, Grant 61802180, Grant 61941116, and Grant 61872192, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20180421, in part by the National Cryptography Development Fund under Grant MMJJ20180105, in part by the Fundamental Research Funds for the Central Universities under Grant NE2018106, and in part by the Major Science Research Project of Jiangsu Provincial Education Department under Grant 19KJA310010. The work of Kim-Kwang Raymond Choo was supported in part by the Cloud Technology Endowed Professorship and National Science Foundation CREST under Grant HRD-1736209. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Frederik Armknecht. *(Corresponding author: Debiao He.)*

Qi Feng and Debiao He are with the Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, School of Cyber Science and Engineering, Wuhan University, Wuhan 430072, China, and also with the State Key Laboratory of Cryptology, Beijing 100878, China (e-mail: fengqi.whu@whu.edn.cn; hedebiao@163.com).

Zhe Liu is with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China (e-mail: sduliuzhe@gmail.com).

Huaqun Wang is with the Jiangsu Key Laboratory of Big Data Security and Intelligent Processing, College of Computer, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: wanghuaqun@aliyun.com).

Kim-Kwang Raymond Choo is with the Department of Information Systems and Cyber Security, The University of Texas at San Antonio, San Antonio, TX 78249 USA (e-mail: raymond.choo@fulbrightmail.org).

Digital Object Identifier 10.1109/TIFS.2020.2997134