Journal of Ambient Intelligence and Humanized Computing (2023) 14:553–566
https://doi.org/10.1007/s12652-021-03312-8

ORIGINAL RESEARCH

Privacy-preserving remote deep-learning-based inference under constrained client-side environment

Amine Boulemtafes[1,2] · **Abdelouahid Derhab**[3] · **Nassim Ait Ali Braham**[4] · **Yacine Challal**[5]

Received: 28 August 2020 / Accepted: 15 May 2021 / Published online: 28 June 2021

Abstract
Remote deep learning paradigm raises important privacy concerns related to clients sensitive data and deep learning models. However, dealing with such concerns may come at the expense of more client-side overhead, which does not fit applications relying on constrained environments. In this paper, we propose a privacy-preserving solution for deep-learning-based inference, which ensures effectiveness and privacy, while meeting efficiency requirements of constrained client-side environments. The solution adopts the non-colluding two-server architecture, which prevents accuracy loss as it avoids using approximation of activation functions, and copes with constrained client-side due to low overhead cost. The solution also ensures privacy by leveraging two reversible perturbation techniques in combination with paillier homomorphic encryption scheme. Client-side overhead evaluation compared to the conventional homomorphic encryption approach, achieves up to more than two thousands times improvement in terms of execution time, and up to more than thirty times improvement in terms of the transmitted data size.

Keywords Deep learning · Deep neural network · Privacy · Sensitive data · Constrained · Inference

1 Introduction

Deep learning is an advanced approach of machine learning, and is nowadays being used in different domains such as speech recognition, computer vision, and medical predictions. It allows classification, prediction, and discovery of new knowledge (Litjens et al. 2017; Kamilaris and Prenafeta-Boldú 2018; Aldweesh et al. 2020; Ferrag et al. 2020). The power of deep learning is particularly leveraged when it is run on powerful infrastructures such as clouds that provide high-power computation and massive storage. In fact, remote deep learning paradigm not only enables the access to proprietary deep models, but it also allows to better fit the increasing depth and computation requirements of powerful deep learning models. Furthermore, it was shown that remote inference can incur shorter execution time compared to the local inference in appropriate settings. More specifically, remote paradigm is particularity useful when the client-side environment, comprising the client devices and network, is resource-constrained. One good example of such an environment is the remote health monitoring, based on sensors and mobile devices that are relying on cellular connectivity. The adoption of remote deep learning paradigm is therefore desirable in such cases, allowing to support powerful analysis (Baryalai et al. 2016; Boulemtafes et al. 2020; Xu et al. 2020).

However, the external infrastructure might feed the deep learning model with sensitive data that are transmitted from client devices. This raises a number of privacy concerns,

✉ Amine Boulemtafes
aboulemtafes@cerist.dz

Abdelouahid Derhab
abderhab@ksu.edu.sa

Nassim Ait Ali Braham
en_ait_ali_braham@esi.dz

Yacine Challal
y_challal@esi.dz

[1] Division Sécurité Informatique, Centre de Recherche sur l'Information Scientifique et Technique, Algiers, Algeria

[2] Département Informatique, Faculté des Sciences exactes, Université de Bejaia, 06000 Bejaia, Algeria

[3] Center of Excellence in Information Assurance, King Saud University, Riyadh, Saudi Arabia

[4] CNRS, LAMSADE, Université Paris-Dauphine, PSL Research University, 75016 Paris, France

[5] Laboratoire de Méthodes de Conception des Systèmes, Ecole Nationale Supérieure d'Informatique, Algiers, Algeria