HomE: Enabling Homomorphic Encryption of Deep Neural Network Models and Datasets in Production

Description

Deep learning (DL) is widely used to solve classification problems previously unchallenged, such as face recognition, and presents clear use cases for privacy requirements. Homomorphic encryption (HE) enables operations upon encrypted data, at the expense of vast data size increase. RAM sizes currently limit the use of HE on DL to severely reduced use cases. Recently emerged persistentmemory technology (PMEM) offers larger-than-ever RAM spaces, but its performance is far from that of customary DRAMtechnologies.

This project aims to spark a new class of system architectures for encrypted DL workloads, by eliminating or dramatically reducing data movements across memory/storage hierarchies and network, supported by PMEM technology, overcoming its current severe performance limitations. HomE intends to be a first-time enabler for the encrypted execution of large models that do not fit in DRAM footprints to execute local to accelerators, hundreds of DL models to run simultaneously, and largedatasets to be run at high resolution and accuracy. Targeting these ground-breaking goals, HomE enters into unexplored field resulting from the innovative convergence of several disciplines, where wide-ranging research is required in order to assess currentand future feasibility. Its main challenge is developing a methodology capable of breaking through the existing software and hardware limitations.

HomE proposes a holistic approach yielding highly impactful outcomes that include novel comprehensive performancecharacterisation, innovative optimisations upon current technology, and pioneering hardware proposals. HomE can spawn a paradigm shift that will revolutionise the convergence of the machine learning and cryptography disciplines, filling a knowledge gap and opening new horizons such as DL training on HE, which is currently even too demanding for DRAM. HomE, based on solid evidence, will unveil the great unknown of whether PMEM is a practical enabler for encrypted DL workloads

Barcelona Supercomputing Center - Centro Nacional de Supercomputación