Task description - Master Thesis

Name, Vorname:	Bhatnagar, Prateek
Studiengang:	Distributed System Engineering (DSE)
Matrikelnummer:	4805056

Thema: ***SPML:*** *A Secure Privacy-preserving Machine Learning Framework using TEEs and Differential Privacy*

Machine learning techniques are widely adopted to build AI systems which are used in practice or in our daily life. For example, nowadays, we can pay at supermarkets using our faces; we can automatically translate the signs on the roads of a place for the first time we arrive by only using camera of our smartphone; and in hospitals, doctors can detect diseases at earlier stages than ever before. However, these machine learning systems rely and are shaped by users data which is increasingly private and sensitive. For example, the online service providers collect their customer's data to train recommender systems the predict future interest of the customers. Healthcare data from patients in hospitals can be collected to train diagnostic models. The banking transactions in the banks can be combined with merchant data and bank account information to train the models for fraud-detection engines. Thus, privacy and security have become paramount especially when these online-services deployed in an untrusted environment, e.g., a public cloud. Indeed, we have seen increasing attention of regulators on issues regarding the way in which personal data is handled and processed - the EU's General Data Protection Regulation a case in point. Thus, confidentiality and integrity of the data processing in clouds are becoming more important, not least because of increased demands for accountability regarding service providers, and the potential serious legal consequences (and fines) for the data mishandling, mismanagement and leakage, and more generally, for failing to implement the appropriate security measures. Service providers must ensure that data is always protected, i.e., at rest, during transmission, and computation.

A promising approach to helping resolve these security challenges is to make use of Trusted Execution Environments (TEEs), such as Intel Software Guard Extensions (SGX). Intel SGX protects the confidentiality and integrity of application code and data even against privileged attackers with root access and physical access. In general, Intel SGX provides an isolated secure memory area called enclaves, where the code and data can be executed safely. These security guarantees are solely provided by the CPU, thus even if the system software is compromised, the attacker can never access the enclave's content. This approach supports data analytics at processor speeds while ensuring the security guarantee for both computation and sensitive data. While promising at first glance, a machine learning system using TEEs, e.g., Intel SGX, still cannot ensure the privacy for data processed inside enclaves since an advisory can incurs several malicious queries to the data stored in enclaves to leak the private information from the data. To overcome this issue, in this thesis, we leverage the advantages of TEEs technologies and differential privacy mechanisms to build a secure and privacy-preserving machine learning system. The TEEs technologies provide strong confidentiality and integrity for the input data of the machine learning system while the differential privacy mechanism protect the privacy of the data against malicious queries. In the proposed system, we show the way to simultaneously unlock all of this power of machine learning while still respecting and protecting data privacy.