Privacy-Protecting, yet Resilient Federated Learning

This invention balances privacy and security in federated learning. It is ideal for privacy-protected machine-learning applications such as diagnosing disease patterns from patient images. With the invention, several cooperating hospitals may for example train a common model without sharing sensitive or protected data and still be assured that only useful updates from the other participants contribute to the model.

IP RIGHTS
International application filed
2024

Excluding bad actors without compromising privacy

The technology can also be used to train driver assistance systems based on data from real customers, with smartphone applications like smart keyboards trained to suggest words based on a user's finger movement, or in any other context where participating clients correspond to personally identifiable individuals.

01 Leveraging the advantages of federated learning for additional privacy-critical applications

02 Uniting the requirements for privacy and security by group testing of overlapping client groups

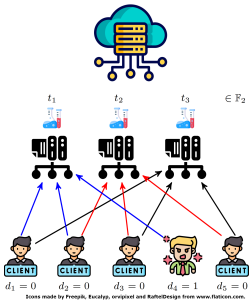
03 In uncertain cases, finding the optimal balance between keeping faulty data and excluding useful contributions

04 Experimentally tested on several benchmark data sets, with significant improvements observed

REFERENCES:

Patent appl. EP4446920A1

Xhemrishi et al. arXiv:2305.05506



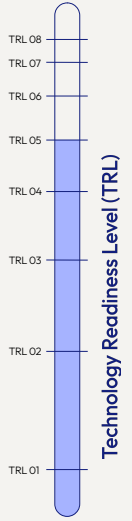
Securely aggregating client updates (lower row) into partially overlapping groups (second row) prevents the central server (top) from inferring individual client data, but still allows for estimating client performance from group performance, enabling the exclusion of malicious clients.

CHALLENGE

In federated learning, a central server shares a common model with several clients, which train the model with local data and provide their resulting model updates to the server. The server then updates and re-shares the model with the clients. To protect against model-inversion attacks and thus increase privacy, the clients may securely aggregate their updates, but this prevents the server from singling out malicious clients or faulty contributions.

INNOVATION

The patent-pending technology addresses the previous trade-off between privacy and security by securely pooling the clients into partially overlapping groups. Privacy is ensured by aggregating the updates within each group and revealing only the resulting group updates to the server in unobscured form. Yet, if the groups are composed accordingly, the server may still estimate the performance of individual clients from the performance of the respective groups in which they take part. Suspicious clients can then be excluded from future increments of the common model.



01 Basic principles observed · **02** Technology concept formulated · **03** Experimental proof of concept · **04** Technology validated in lab · **05** Technology validated in relevant environment · **06** Technology demonstrated in relevant environment · **07** System prototype demonstrated in operational environment · **08** System complete and qualified

BayPAT

Innovation from Research & Science – BayPAT markets inventions on behalf of the Bavarian universities and universities of applied sciences.

An invention of **Technical University of Munich**

Technische Universität München

As one of the largest technology transfer offices in Germany, BayPAT evaluates and commercializes innovations from 28 research institutions. With our extensive experience and expertise in patenting and licensing technologies and our comprehensive range of services, we are your partner of choice. For more information visit www.baypat.de

Bayerische Patentallianz GmbH
Dr. Sindre Haugland
shaugland@baypat.de
+49 (0)89 5480177-17
Prinzregentenstr. 52
80538 München