Doctoral School of Information and Biomedical Technologies Polish Academy of Sciences - TIB-PAN

Research domain: Informatyka Techniczna i Telekomunikacja

Topic: 1.4 Uczenie maszynowe – zagadnienia specjalne

Trustworthy Federated Learning in Cloud-based Intelligent Transport Systems

Supervisor; contact information

dr hab. Joanna Kołodziej (main supervisor);
tel. 601688140 oraz 22 38 08 303, joanna.kolodziej@nask.pl;
Naukowa i Akademicka Sieć Komputerowa – Państwowy Instytut Badawczy (NASK-PIB),
ul. Kolska 12, Warszawa
dr inż. Mateusz Krzysztoń (co-supervisor)
mateuszkr@nask.pl
Naukowa i Akademicka Sieć Komputerowa – Państwowy Instytut Badawczy (NASK-PIB),
ul. Kolska 12, Warszawa

Scope

The rapid development of research, technology and information tools for electronics, communication and control systems, sensor networks and the processing of enormous data sets has contributed to a real revolution in modern public and private transport management and the development of effective strategies for road infrastructure and management systems and the efficient use of existing infrastructure. The current Information and Communication Technology systems (ICT) have significantly improved the transfer of accurate traffic data, implementing control measures, and varying the level of uncertainty and randomness that characterised conventionally, manually managed transport networks. Successful implementation of Intelligent Transportation Systems (ITS) requires a good understanding of both local and global traffic and mobility models [1], and the impact of related possible traffic anomalies, such as the generation and propagation of shock waves, initiation of congestion, etc [2], [3], [4].

Federated (deep) learning (FL) has become an emerging paradigm for collaborative learning in large-scale distributed systems with a massive number of networked clients, such as smartphones, connected vehicles or edge devices [5]. Compared to other distributed learning approaches, federated learning allows the clients to train models without sharing raw data, which achieves privacy-preserving machine learning in real application scenarios. This brings great opportunities for deploying AI approaches in future intelligent traffic systems by means of Lora WAN and other communication networks. FL follows an orchestration architecture where its learning algorithms are embedded and executed in users' devices while being able to communicate and collaboratively train a model without exchanging any training data.

This PhD thesis aims to delve into the FL concepts to propose appropriate solutions for security and trustworthy preserving traffic data mining. It will consider the main analysis tasks, such as trajectory clustering, data mobility pattern mining and detection of anomalies. The proposed model(s) will address the following research questions:

1. How to adapt the FL pipeline to traffic data?

2. Which analysis methods, preserving the confidentiality of traffic data and detecting the possible anomalies in such data, are best suited to the architecture, and how to optimize or adapt them?

3. How to automatically deploy an FL-based security preservation solution in a cloud computing infrastructure?

All developed models and algorithms will be implemented and demonstrated over realistic datasets obtained as a result of collaboration with WobCom company (https://www.wobcom.de/).

Requested skills:

- MSc degree in computer sciences telecommunication or similar discipline,
- Backgrounds in machine learning,
- Advanced practical knowledge of Python/Java
- Experience in at least one ML tool (e.g. Tensorflow, PyTorch, scikit-learn)
- Advanced Level in English (speaking and writing).

References

1. Xu, Y., Kong, Q. J., Lin, S., and Liu, Y.: ”Urban traffic flow prediction based on road network model”, in Proc. of the 9th IEEE International Conference on Networking, Sensing and Control (ICNSC), pp. 334–339, 2012.
2. Li, R., and Lu, H.: ”Combined neural network approach for short-term urban freeway traffic flow prediction”, In Advances in Neural Networks–ISNN, pp. 1017–1025, 2009.
3. Vanajakshi, L., Subramanian, S.C., and Sivanandan, R.: ”Travel Time Prediction under Heterogeneous Traffic Conditions Using Global Positioning System Data from Buses”, IET Intelligent Transport Systems, vol. 3(1), pp. 1–9, 2009.
4. Subutai, A., Lavin, A., Purdy, S., and Agha.: ”Unsupervised real–time anomaly detection for streaming data”, Neurocomputing, vol. 262, pp. 134 – 147, 2017.
5. Kairouz, Peter, H. Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Keith Bonawitz et al. "Advances and open problems in federated learning." arXiv preprint arXiv:1912.04977 (2019).

Warsaw, February, 2023