Tandem Outlier Detectors for Decentralized Data

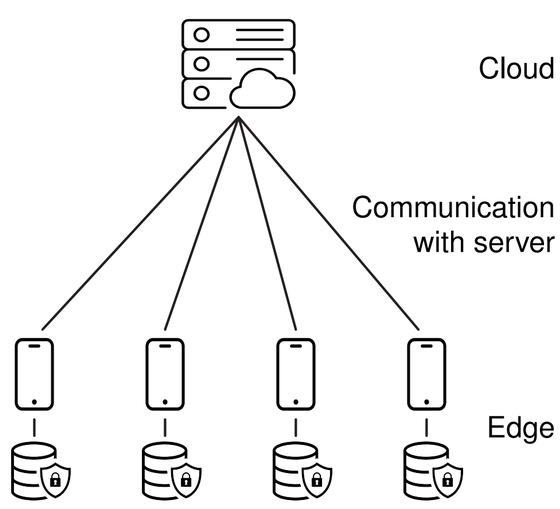
Marco Heyden, Jürgen Wilwer, Edouard Fouché, Vadim Arzamasov, Steffen Thoma, Sven Matthiesen, and Thomas Gwosch

Karlsruhe Institute of Technology

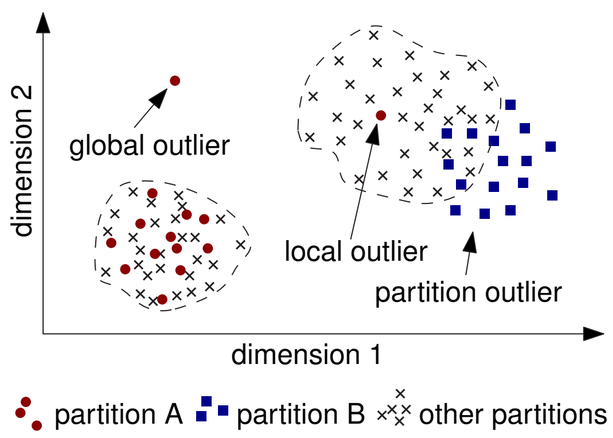
ENERGY STATUS DATA
COLLECT • ANALYSE • UTILISE

Scenario
Finding outliers in decentralized data

- Data creation is often decentralized
  - Smart meters, powertools, electric cars
- Each device holds part of the data generated in the network (a *partition*)
- **What outliers exist in decentralized data and how can we find them?**



Types of Outliers
Local, global, and partition outliers



Local and centralized outlier detection are insufficient

- Centralized:
  - Transfer data to central storage prior to outlier detection
  - Able to detect each outlier type
  - Problematic w.r.t. **privacy** and **efficiency**
- Local:
  - Identify outliers in a single partition
  - Data remains on the device → **efficient and privacy-preserving**
  - Only **limited data available** for outlier detection

Research Question

How can we **find local, global, and partition outliers in decentralized data without sharing raw data** with other devices?

Solution Sketch
Combining federated and local outlier detectors

1. Two outlier detectors L_i and F:
   - F: on single partition d_i
   - L_i: on entire network N with data $D = d_1 \cup d_2 \cup \ldots \cup d_{|N|}$ using Federated Learning
2. Outlier scores for device i:
   - $os_i^F = \{os_{i,1}^F, os_{i,2}^F, \ldots, os_{i,|d_i|}^F\}$
   - $os_i^L = \{os_{i,1}^L, os_{i,2}^L, \ldots, os_{i,|d_i|}^L\}$
3. Distinguish point outliers:
   - Global outlier if $os_{ij}^L > \varepsilon_i$ and $os_{ij}^F > \varepsilon_i$
   - Local outlier if $os_{ij}^L > \varepsilon_i$ and $os_{ij}^F \leq \varepsilon_i$
4. Identify partition outliers
   - Sort os_i^F, create bins of size b, compute averages
   - Set of average scores is os_i^*
   - Mann-Whitney-U test between os_i^* and $\bigcup_{k \in N \setminus i} os_k^*$

Results
We could identify mistakes in operation and unusual user behavior

- **Partition outlier:** User performed the task very well

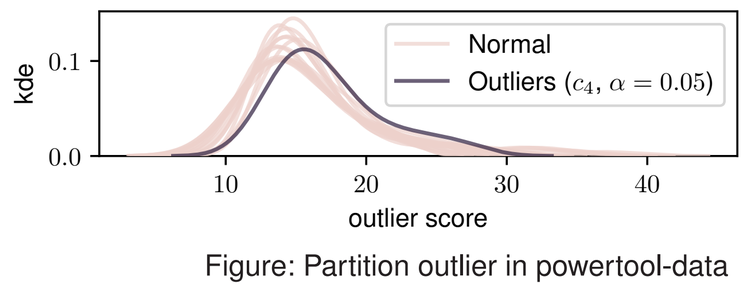


Figure: Partition outlier in powertool-data

- **Local outliers:** Minor deviation from regular usage
- **Global outlier:** User slipped off the screw-head

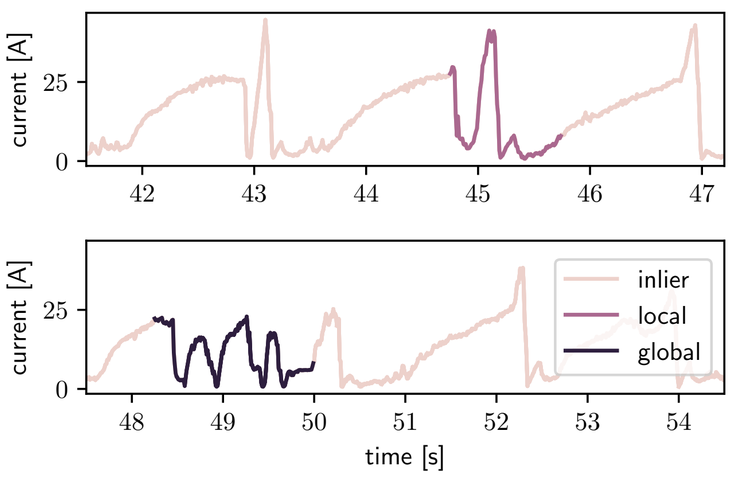


Figure: Local and global outlier in powertool-data

User Study
Connected Powertools

Protocol

- 15 participants
- 20 screwing tasks; 20 drilling tasks
- Prespecified goals (e.g., target depth)
- Acceleration, sound, magnetic field, gyroscope data

Outlier detectors

- F and L_i: single-layer Autoencoders
- Reconstruction error = Outlier score
- Trained on features extracted from sensor data in sliding windows of 25 ms

Related Work

- Server-coordinated [4] and peer-to-peer [6] Federated Learning
- Federated Learning for outlier detection [8, 9, 7, 5]
- Other definition of “local”: outlier occurs multiple times in a small geographical region [3, 2, 10]
- Wireless Sensor Networks [1, 2]

Data and Code

https://github.com/heymarco/TandemOutlierDetection

References

[1] S. Bharti, K. K. Pattanaik, and A. Pandey. “Contextual outlier detection for wireless sensor networks”. In: *J. Ambient Intell. Humaniz. Comput.* (2020).

[2] H. H. W. J. Bosman, G. Iacca, et al. “Spatial anomaly detection in sensor networks using neighborhood information”. In: *Inf. Fusion* (2017).

[3] N. Giatrakos et al. “TACO: tunable approximate computation of outliers in wireless sensor networks”. In: *SIGMOD*. 2010.

[4] B. McMahan et al. “Communication-Efficient Learning of Deep Networks from Decentralized Data”. In: *AISTATS*. 2017.

[5] T. D. Nguyen, S. Marchal, et al. “DÏoT: A Federated Self-learning Anomaly Detection System for IoT”. In: *ICDCS*. 2019.

[6] Abhijit Guha Roy et al. “BrainTorrent: A Peer-to-Peer Environment for Decentralized Federated Learning”. In: *CoRR* abs/1905.06731 (2019).

[7] R. A. Sater and A. B. Hamza. “A Federated Learning Approach to Anomaly Detection in Smart Buildings”. In: *CoRR* (2020).

[8] J. Schneible and A. Lu. “Anomaly detection on the edge”. In: *MILCOM*. 2017.

[9] S. Singh, S. Bhardwaj, et al. “Anomaly Detection Using Federated Learning”. In: *ICAIA*. 2021.

[10] X. Yu, H. Lu, et al. “An adaptive method based on contextual anomaly detection in Internet of Things through wireless sensor networks”. In: *IJDSN* (2020).