Reliable Federated Learning for Mobile Networks

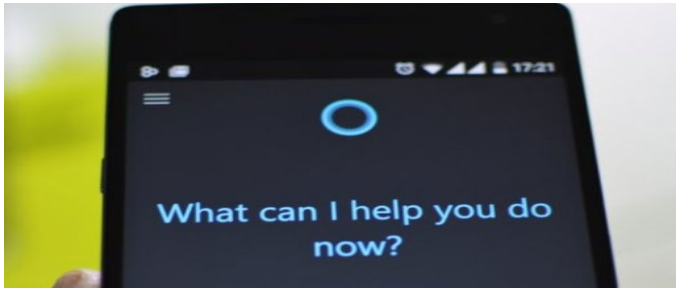
Dusit Niyato

Outline

- **Preliminary: Federated Learning Meets Blockchains**
  - Federated Learning
  - Security Challenges for Federated Learning
  - Solutions for the Security Challenges
- **Reputation for Reliable Federated Learning**
  - Reputation
  - Subjective Logic Model for Reputation Calculation
  - Blockchain
- **Numerical Results**
- **Summary**

Why we need federated learning?

Machine Learning for our life



AI-based voice input

HUAWEI P20 Pro with HUAWEI AIS

Other Phone

- It is increasingly popular to utilize machine learning technologies to dramatically enhance the performance of mobile applications
- Traditional machine learning require mobile devices to directly upload user data with sensitive private information to a central server for model training

- **Challenges:**
  - Large computation and storage overhead
  - Centralized management suffered from single point of failure and data manipulation
  - Hinder future development of machine learning

Can data is stored and processed at the edge?

Data storage and processing are moving on devices at the edge for:

- Improved latency
- Works offline
- Better battery life
- Privacy advantages

How about the learning?

E.g., on-device analytics inference for mobile keyboards and cameras.

A New Solution - Federated Learning

Google AI Blog

The latest news from Google AI

Federated Learning: Collaborative Machine Learning without Centralized Training Data

Thursday, April 6, 2017

Posted by Brendan McMahan and Daniel Ramage, Research Scientists

Standard machine learning approaches require centralizing the training data on one machine or in a datacenter. And Google has built one of the most secure and robust cloud infrastructures for processing this data to make our services better. Now for models trained from user interaction with mobile devices, we're introducing an additional approach: *Federated Learning*.

Federated Learning enables mobile phones to collaboratively learn a shared prediction model while keeping all the training data on device, decoupling the ability to do machine learning from the need to store the data in the cloud. This goes beyond the use of local models that make predictions on mobile devices (like the Mobile Vision API and On-Device Smart Reply) by bringing model *training* to the device as well.

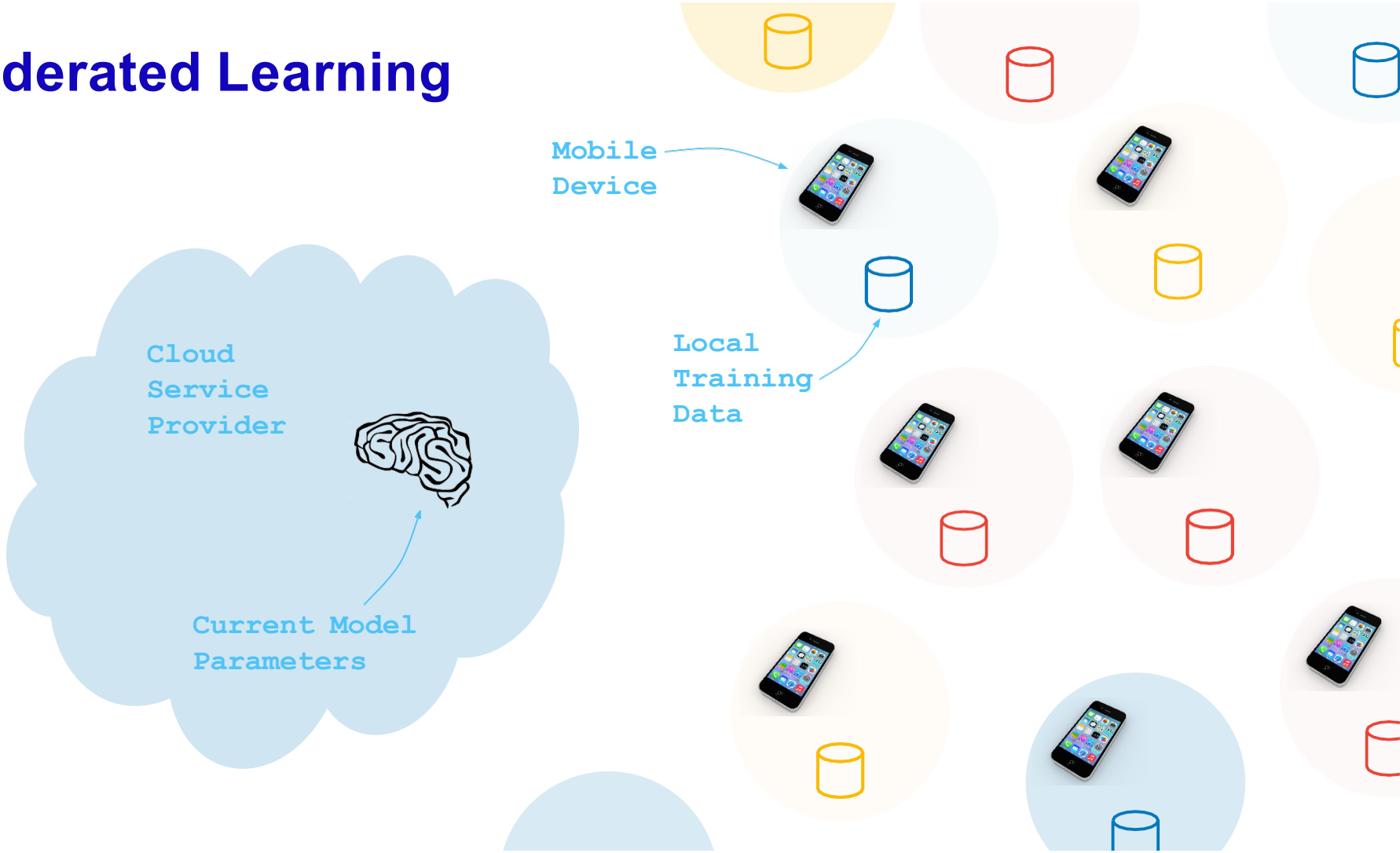
It works like this: your device downloads the current model, improves it by learning from data on your phone, and then summarizes the changes as a small focused update. Only this update to the model is sent to the cloud, using encrypted communication, where it is immediately averaged with other user updates to improve the shared model. All the training data remains on your device, and no individual updates are stored in the cloud.

What is federated learning?

Federated learning

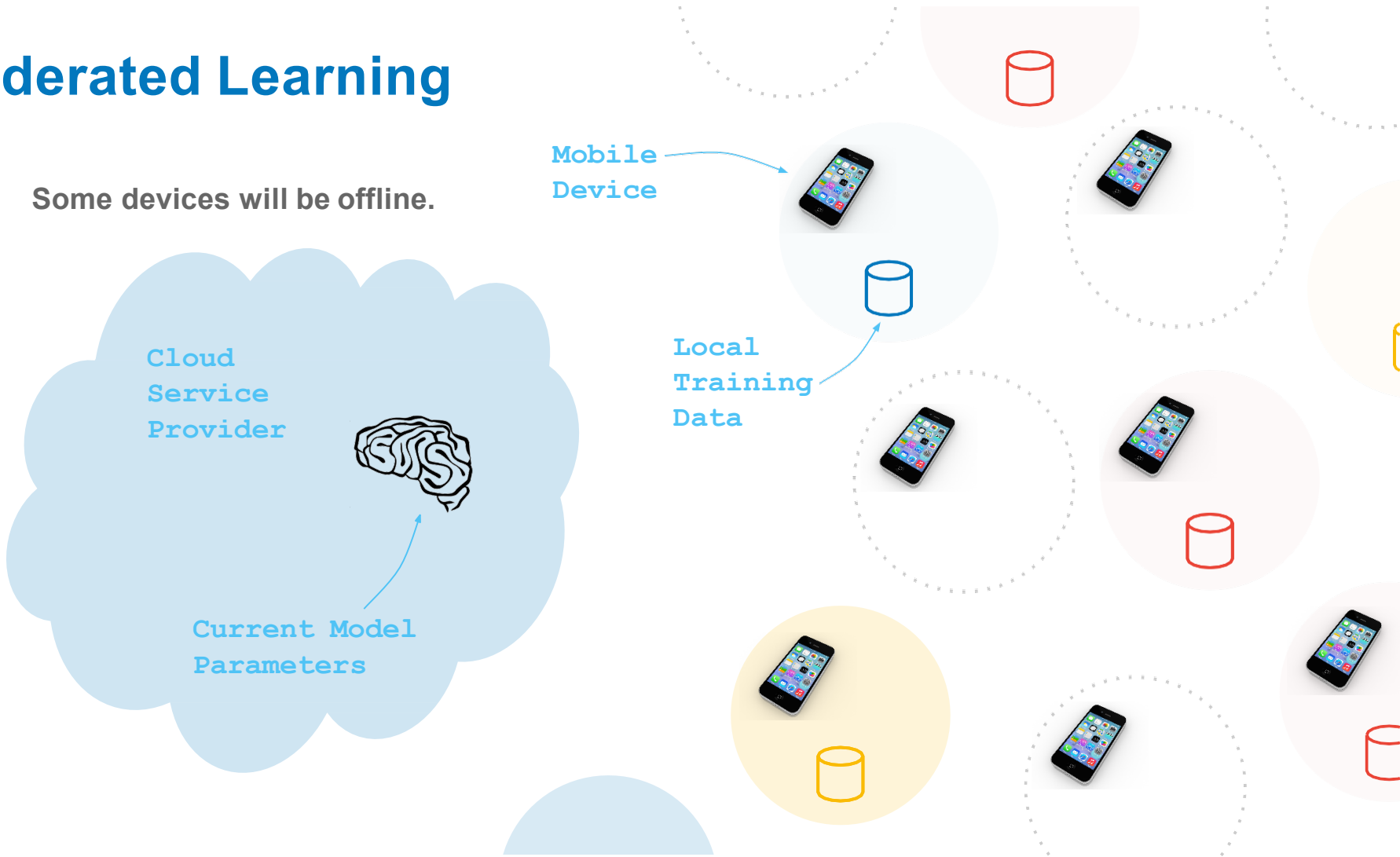
Enable machine learning engineers and scientists to work productively with decentralized data with privacy by default

Federated Learning



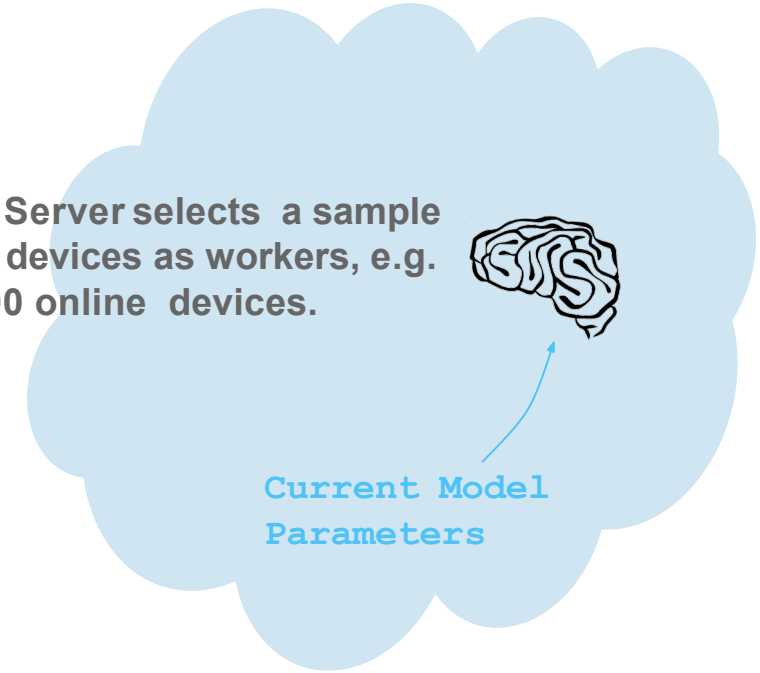
Federated Learning

Some devices will be offline.

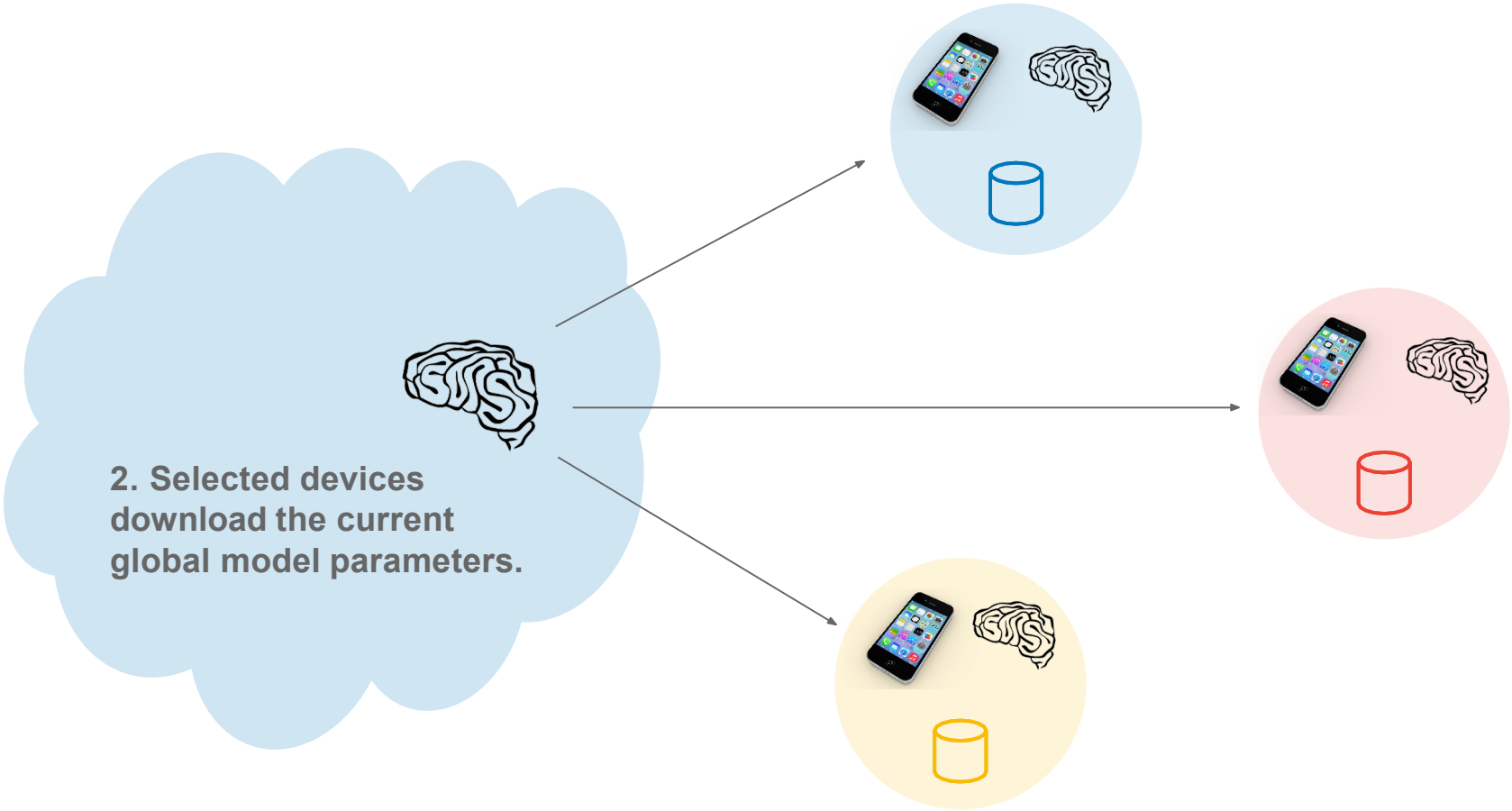


Federated Learning

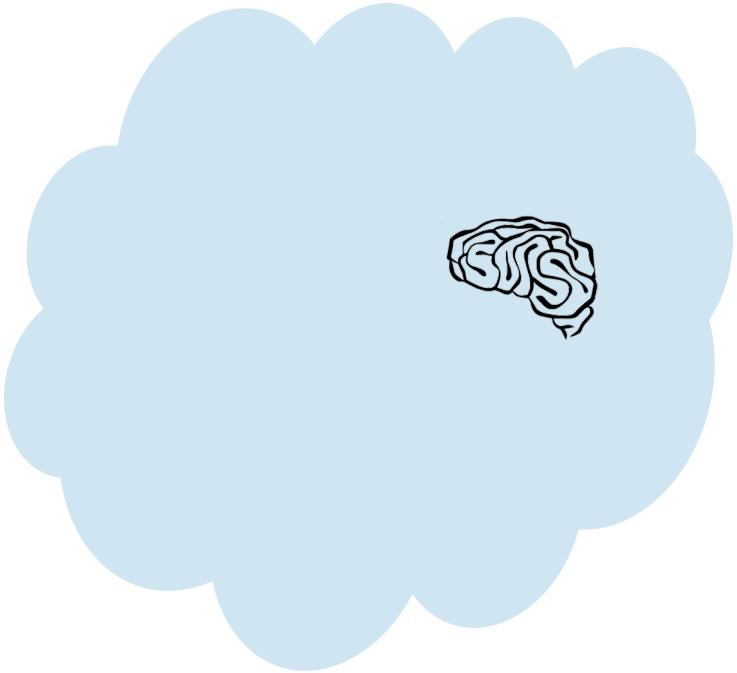
Some devices will be offline.



Federated Learning

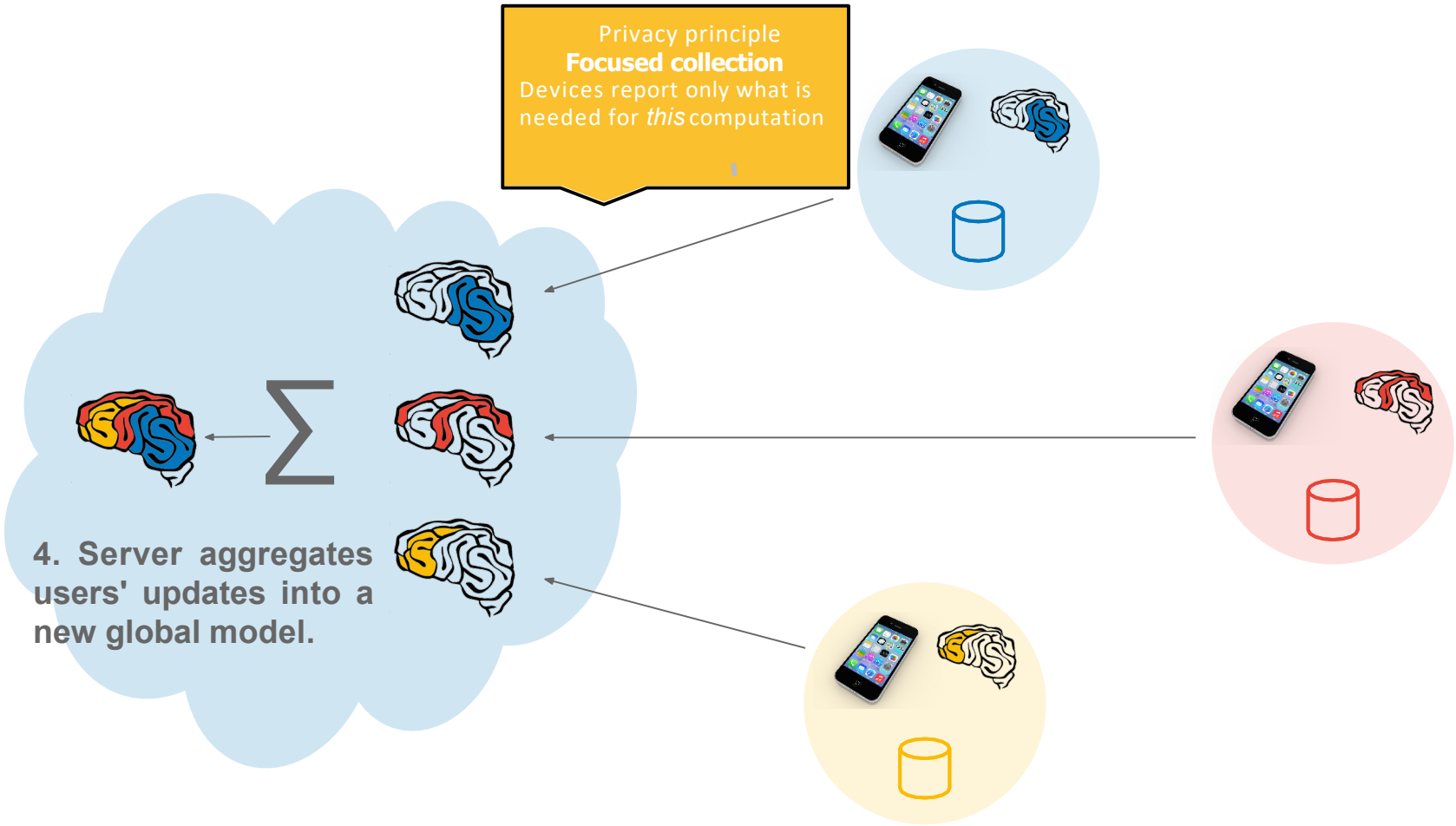


Federated Learning

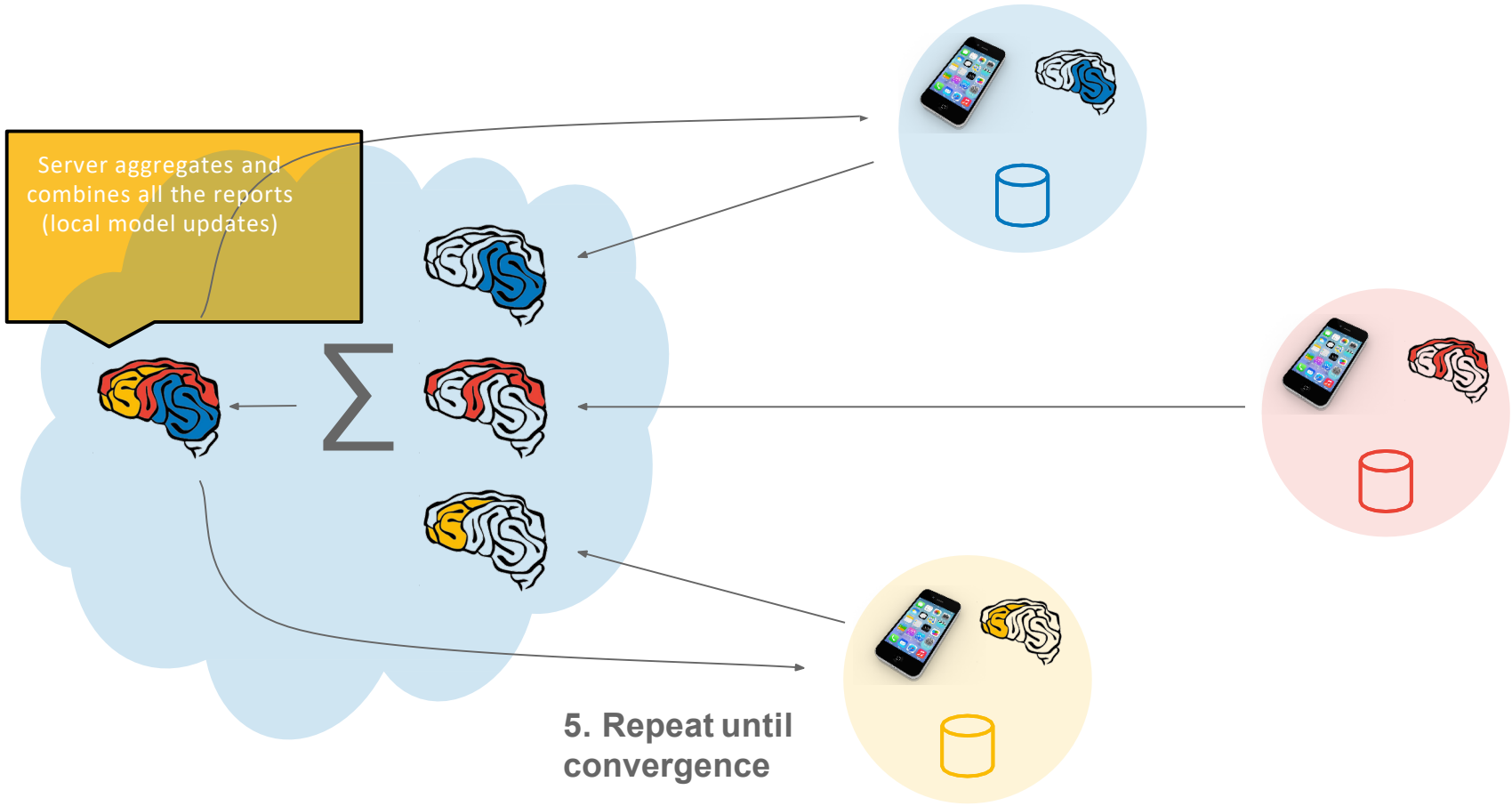


3. Devices compute a local model update using local training data

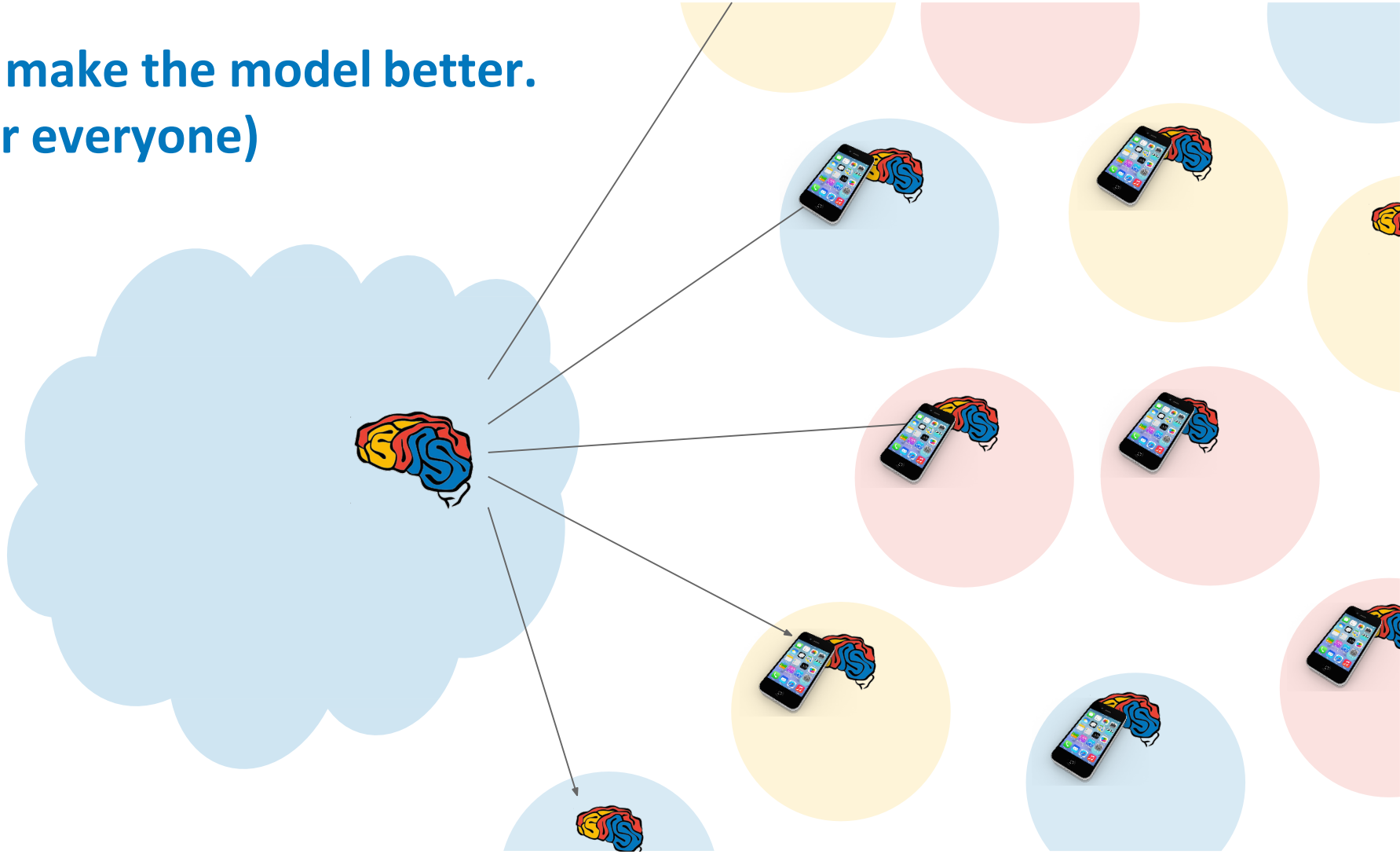
Federated Learning



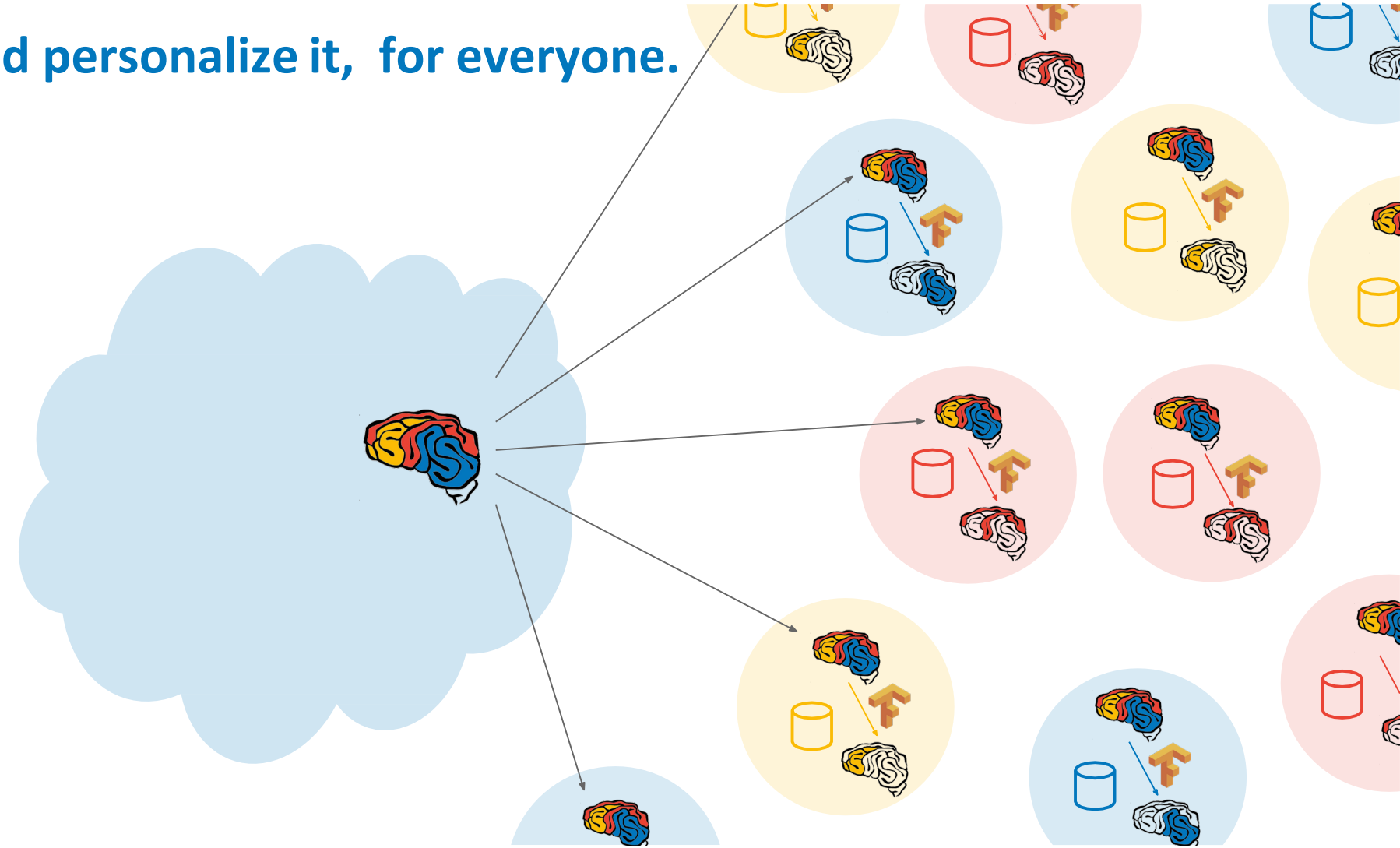
Federated Learning



To make the model better.
(for everyone)



And personalize it, for everyone.



Characteristics of federated learning
vs. traditional *distributed learning*

Data locality and distribution
- **decentralized, naturally arising (non-IID) partition**
- *system-controlled* (e.g. shuffled, balanced)

Data availability
- **limited availability, time-of-day variations**
- *almost all data nodes always available*

Addressability
- **data nodes are anonymous and interchangeable**
- *data nodes are addressable*

Node statefulness
- **stateless (generally no repeat computation)**
- *stateful*

Node reliability
- **unreliable (~10% failures)**
- *reliable*

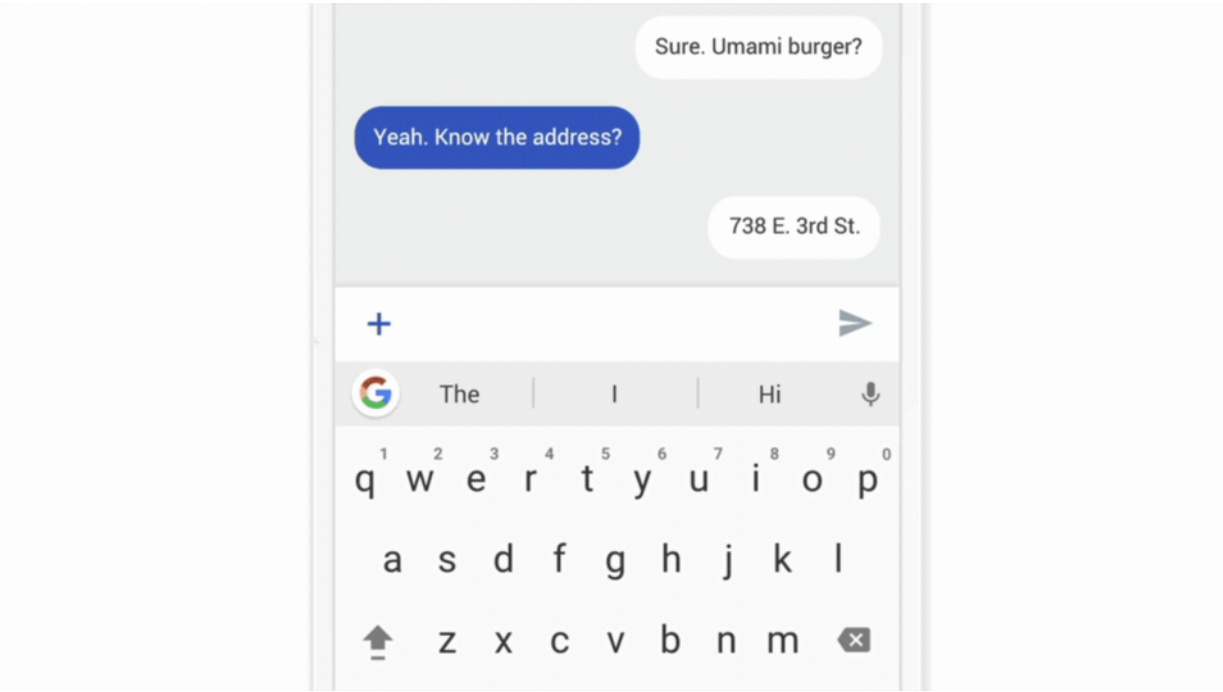
Wide-area communication pattern
- **hub-and-spoke topology**
- peer-to-peer topology (fully decentralized)
- *none (centralized to one datacenter)*

Distribution scale
- **massively parallel (1e9 data nodes)**
- *single datacenter*

Primary bottleneck
- **communication**
- *computation*

How can we use federated learning?

Federated learning for Keyboard



Gboard

Processing the text history on-device to suggest improvements to the next iteration of query suggestion model.

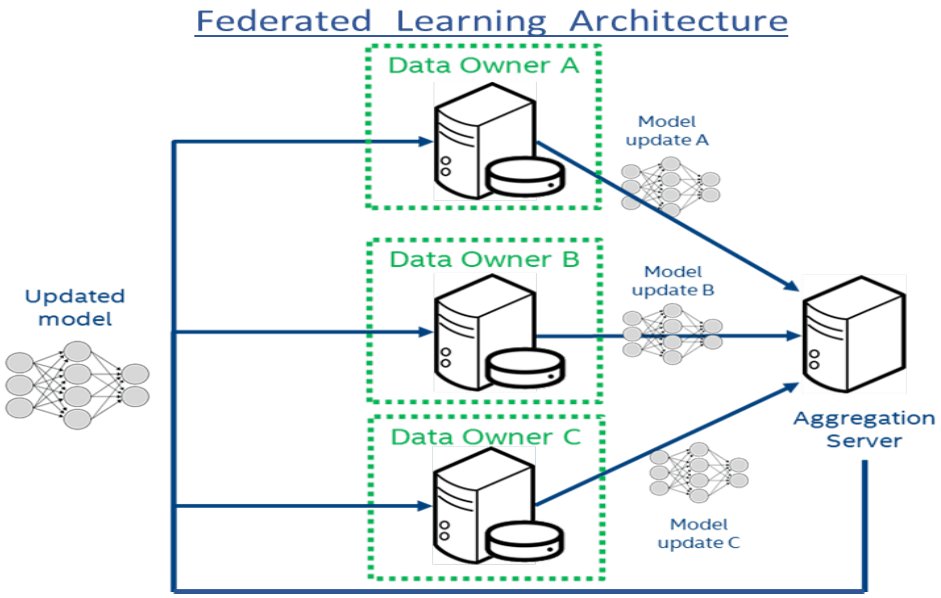
Federated Learning for Medical Imaging

While machine learning can benefit from this “big data” to generate state-of-the-art models, most healthcare data is hard to obtain due to legal, privacy, technical, and data-ownership challenges, especially among international institutions where HIPAA [1] and GDPR [2] concerns need to be addressed[3]

[1] Stanford Medicine. 2017 Health Trends Report: Harnessing the Power of Data in Health. Accessed online 8 FEB 2019.
[2] https://gdpr-info.eu/
[3] https://www.intel.ai/federated-learning-for-medical-imaging/#gs.lmrsj3

Intel AI for Medical Imaging

The encrypted model is sent to the individual institutions (Data Owners A-C) **which decrypt within a secure enclave in hardware** and then train on the local data. Only the model updates are shared with the central model aggregator. This provides protection to both the model and the data. The raw data never leaves the institutions, which not only adds privacy but also prevents large data transfers on the network.



Federated Learning Architecture using Intel hardware.

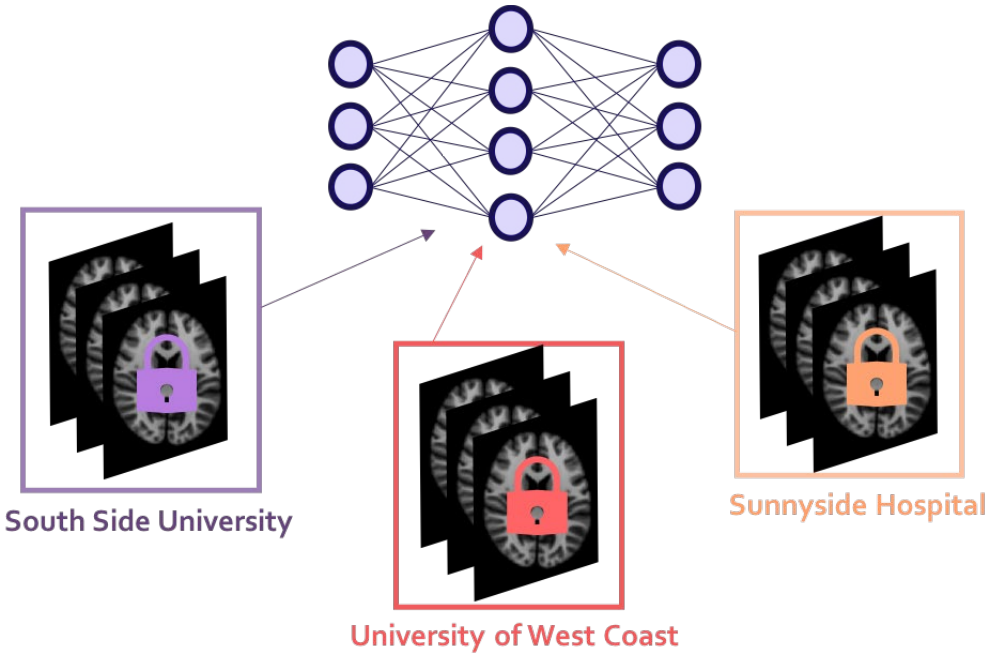
https://www.intel.ai/federated-learning-for-medical-imaging/#gs.lmrsj3

OpenMined for Medical Imaging

Federated Learning: allows us to train AI models on distributed datasets that you cannot directly access.

Differential Privacy: allows us to make formal, mathematical guarantees around privacy preservation when publishing our results (either directly or through AI models).

Encrypted Computation: allows machine learning to be done on data while it remains encrypted.

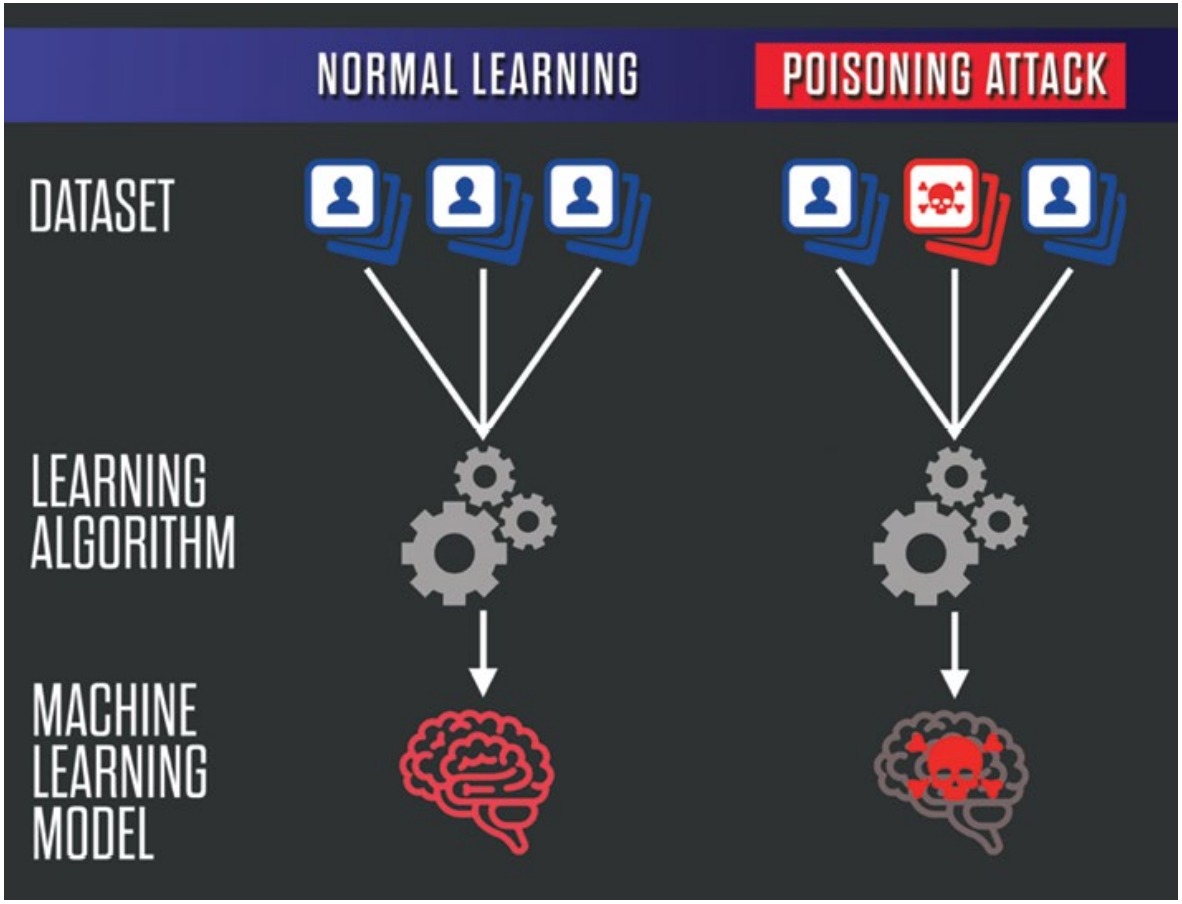


https://blog.openmined.org/federated-learning-differential-privacy-and-encrypted-computation-for-medical-imaging/

Security challenges for federated learning

Poisoning Attacks

- Malicious participants can corrupt the global model:
  - E.g. a retail store can corrupt a recommender system to its own advantage



Poisoning Attacks

- Malicious participants can corrupt the global model:
  - E.g. a retail store can corrupt a recommender system to its own advantage
- Types:
  1. Data poisoning attacks
     - Dirty-label samples to cause misclassification, e.g., in recommender system (Chen, 2017)
     - Sybil attacks: Improve data poisoning effectiveness by creating multiple malicious participants
     - Solution: In non-IID setting, sybil participants will contribute gradients that are more similar to each other than to honest participants. This allows them to be detected (Fung, 2018)

[Chen, 2017] X. Chen, C. Liu, B. Li, K. Lu, and D. Song, "Targeted backdoor attacks on deep learning systems using data poisoning," arXiv preprint arXiv:1712.05526, 2017.
[Fung,2018] C. Fung, C. J. Yoon, and I. Beschastnikh, "Mitigating sybils in federated learning poisoning," arXiv preprint arXiv:1808.04866, 2018

Poisoning Attacks

2. Model Poisoning Attacks

- Directly poison the global model by malicious local model updates, rather than manipulate data.

- More effective than data poisoning attacks
  - With just one attacker, the whole model can be poisoned (Bagdasaryan, 2018)

- Solutions (Bhagoji, 2018)
  - Check if model update from a participant can improve global model performance. If not, mark participant as a potential attacker
  - Check if model update varies a lot from other participants

- However, existing solutions are intractable with large cost.

[Bagdasaryan, 2018] E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, and V. Shmatikov, “How to backdoor federated learning,” arXiv preprint arXiv:1807.00459, 2018.
[Bhagoji, 2018] A. N. Bhagoji, S. Chakraborty, P. Mittal, and S. Calo, “Analyzing federated learning through an adversarial lens,” arXiv preprint arXiv:1811.12470, 2018.

Solutions for the Poisoning Attacks

- For the poisoning attacks, if a malicious data owner is selected to be a worker, the malicious worker may intentionally launch or collude with other workers to launch the attacks.

- Therefore, it is vitally important to design a **reliable worker selection** scheme for model training.

- Nevertheless, in federated learning, the following challenges for the worker selection need to be addressed.

- **No Reliable and Fair Metrics to Evaluate Workers**: A majority of federated learning systems randomly select mobile devices to be the workers through verifable random functions or resource conditions. However, the existing schemes cannot measure the trustworthiness level of workers to remove unreliable or untrusted workers.

Solutions for the Poisoning Attacks

- **No Effcient and Universal Worker Selection Schemes:** It is difficult to design an efficient and universal worker selection scheme for identifying high-quality data contributors and malicious worker candidates.

- **No Timely Monitoring Methods for Workers:** It is hard for the central aggregator to monitor the large-scale worker behaviors in real-time. As a result, a malicious or unreliable worker may be selected to be a worker again for a new federated learning task because of the lack of time-accumulated metrics to evaluate the worker's historical performance and the synchronous information of malicious and unreliable worker lists.

To address the above challenges, we introduce a reliable metric and design a reliable worker selection scheme for federated learning.

Our Work

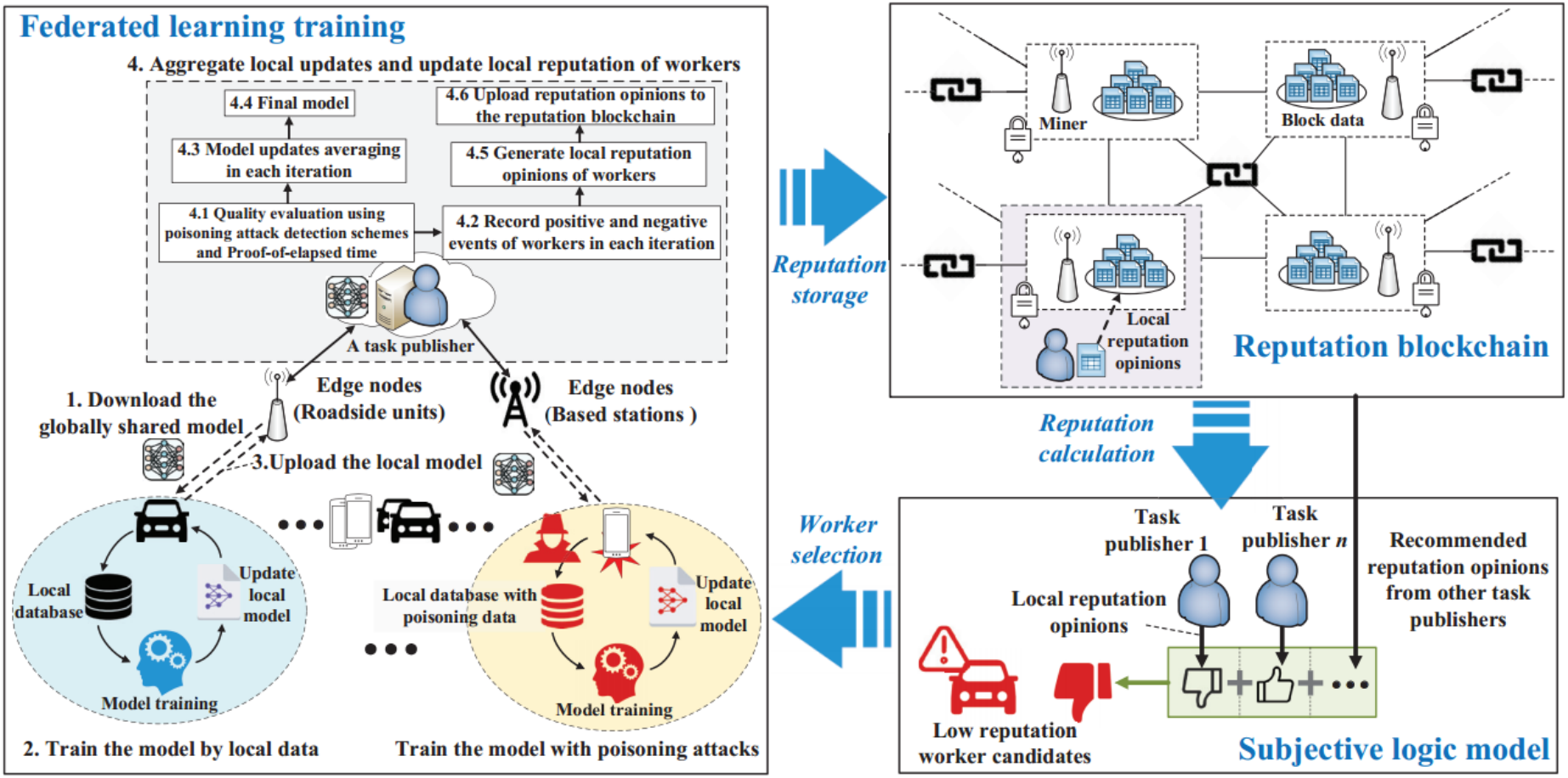
- **Motivation:** To defend against the poisoning attacks, it is vitally important to design a reliable worker selection scheme for model training. Reputation can be utilized as a fair metric that is a time-accumulated metric to indicate the rating of how reliable or trusted an entity is in certain activities according to its historical behaviors .

- **Ideas:**
  - To defend against unreliable model updates, **reputation is introduced as a reliable metric** to select trusted workers for reliable federated learning
  - A **multi-weight subjective logic model** is applied to design an efficient reputation calculation scheme according to both task publishers' interaction histories and recommended reputation opinions.
  - To achieve secure reputation management, the reputation is managed in a decentralized manner by employing the **consortium blockchain** deployed at edge nodes

Reputation-based worker selection scheme with consortium blockchain

J. Kang, Z. Xiong, D. Niyato, S. Xie, and J. Zhang, “Incentive Mechanism for Reliable Federated Learning: A Joint Optimization Approach to Combining Reputation and Contract Theory,” *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10700–714, Dec. 2019.

The Proposed Framework



- Reputation based worker selection: reputation calculation using multi-weight subjective logic model
- Blockchain for reputation management: Consortium blockchain for efficient management

Reputation Calculation using Subjective Logic Model

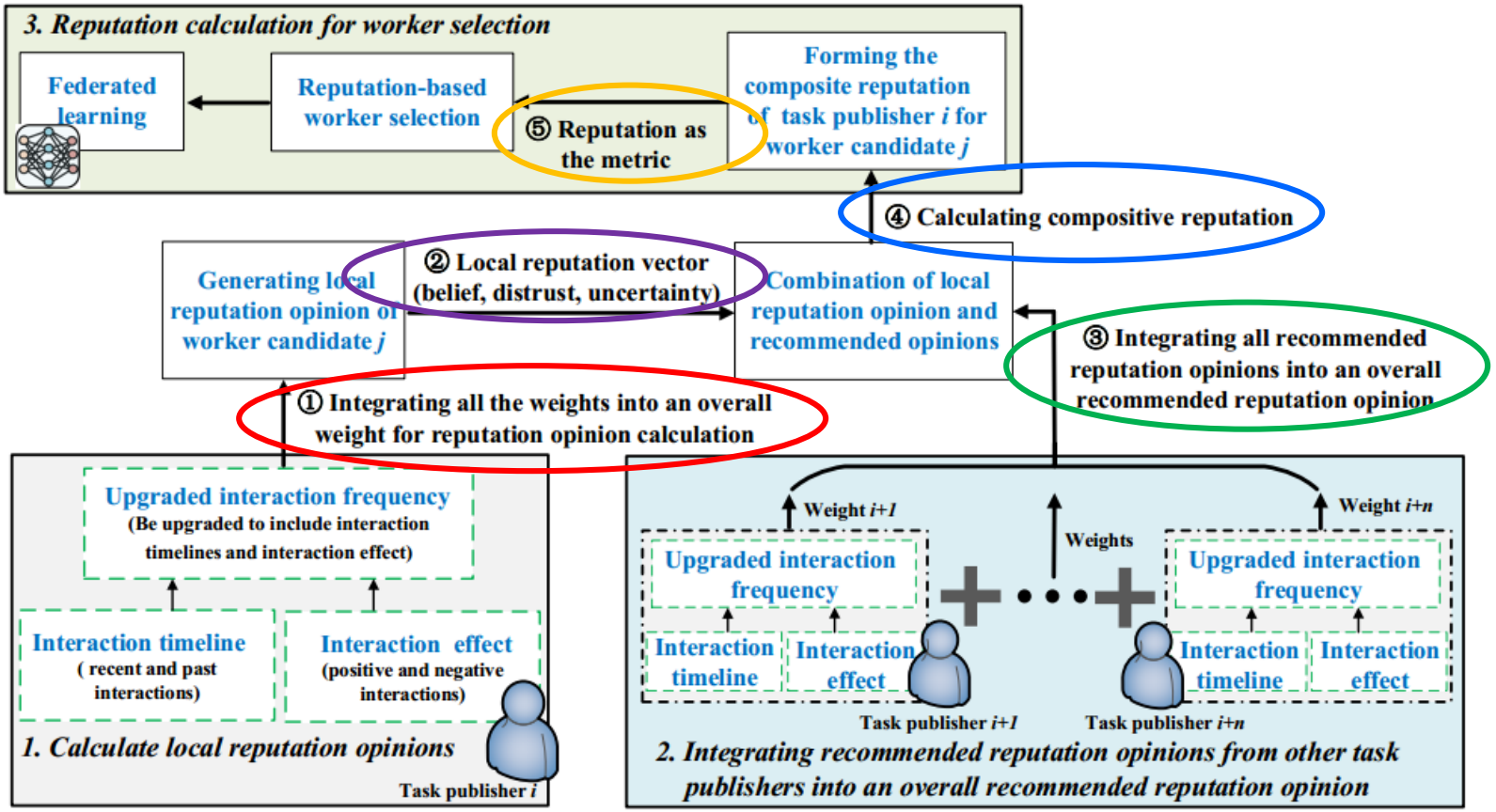
- **Reputation** is defined as the rating of an entity's trustworthiness by others based on its past behaviors, can be a metric to select worker.
- Reputation values are distributed, different, changing over time.
  - It is necessary to quantify miner candidates' reputation based on the interactions histories. The worker candidates with high-reputation should be the workers, who can ensure a secure and reliable federated learning.

Reputation Calculation using Subjective Logic Model

- **Subjective Logic** is utilized to formulate individual evaluation of reputation based on past interactions and recommended opinions.
- The subjective logic utilizes the term “opinion” to denote the representation of a subjective belief, and models positive, negative statements and uncertainty.

Multi-weight Subjective Logic Model

- Multi-weight subjective logic is an extension of subjective logic that takes different attributes of interaction events into consideration for more accurate and reliable reputation calculation.



An overview of reputation calculation

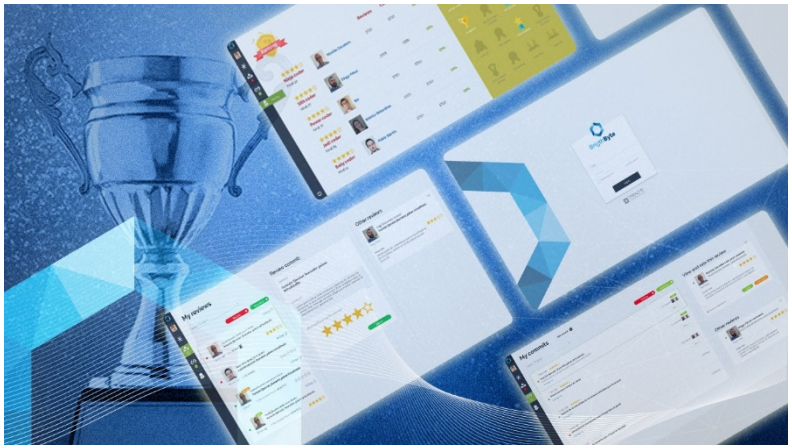
Multi-weight Subjective Logic Model

- Interaction Frequency:
  - represents the familiarity degree between a task publisher and a worker, which is expressed by the ratio of the number of times that the publisher interacts with the worker to the average number of times that the publisher interacts with other workers during a time window.
  - The **higher interaction frequency brings more prior knowledge about the worker** to the publisher, hence leading to a higher local reputation opinion for the worker.
- Interaction Timeline:
  - The trust level and the local reputation opinion of a worker for the same task publisher are changing over time. To evaluate the time effects on interactions, a time scale, for example, three days, is utilized to divide the interaction events into recent and past interactions.
  - **The recent interactions have a higher weight on the task publisher's reputation opinions。**
- Interaction Effects:
  - We classify the interaction events into positive and negative interactions. The negative interactions workers decrease the reputation of the workers, and vice versa.
  - **The positive interactions have a higher weight on the reputation opinion calculation.**

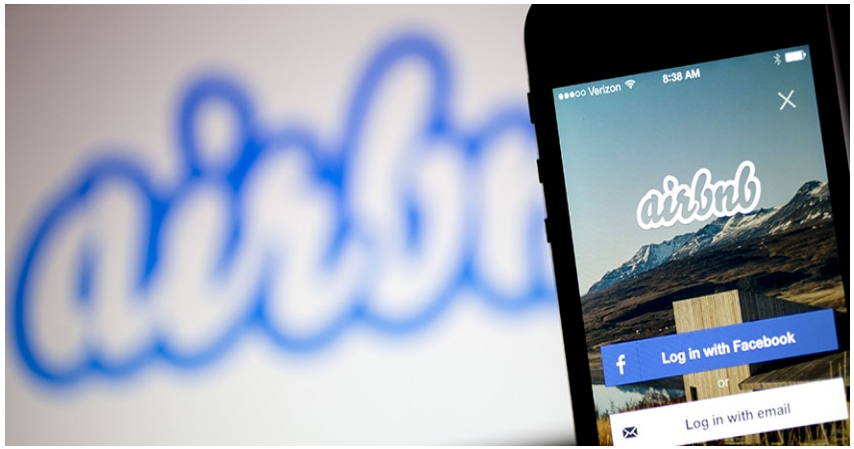
Taking the interaction timelines and interaction effects into consideration, the interaction frequency is upgraded to contain the above two weights.

Blockchain for Reputation Management

- Blockchain: provides a perfect way for distributed systems to record data（e.g., reputation records) that is designed to be transparent, permanent, auditable.

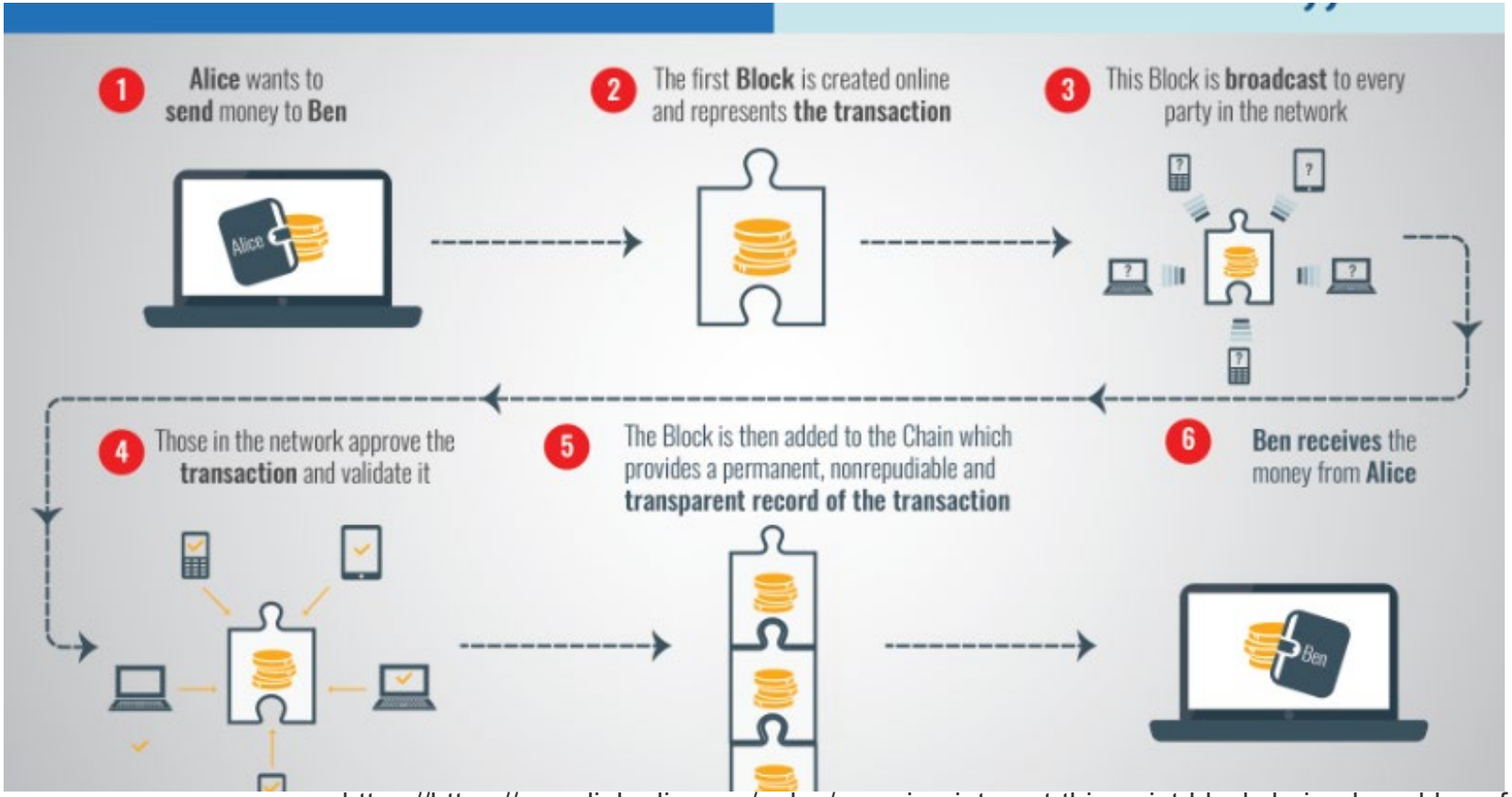


Tribalyte has created **BrightByte**, a reputation system powered by blockchain technology, which is aimed at generating a reputation for software developers, based on the quality of the written code



Airbnb founder envisions reputation blockchain empowering sharing economy. Firms could begin rating people in a shared blockchain in order to decide what services to provide and whether to

Blockchain Concept

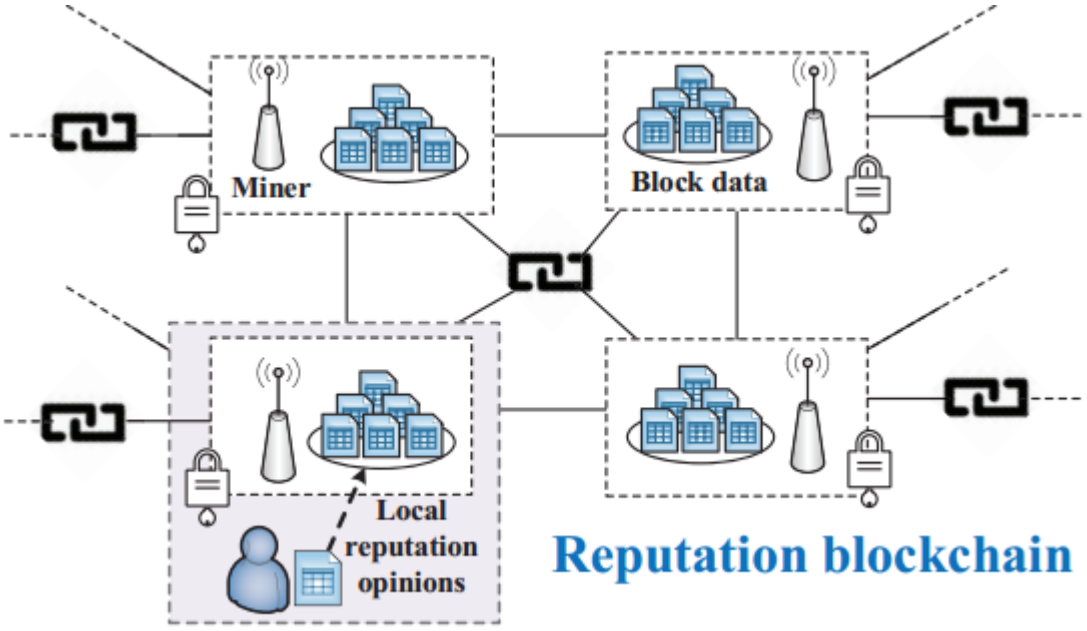


https://https://www.linkedin.com/pulse/securing-internet-things-iot-blockchain-ahmed-banafa/

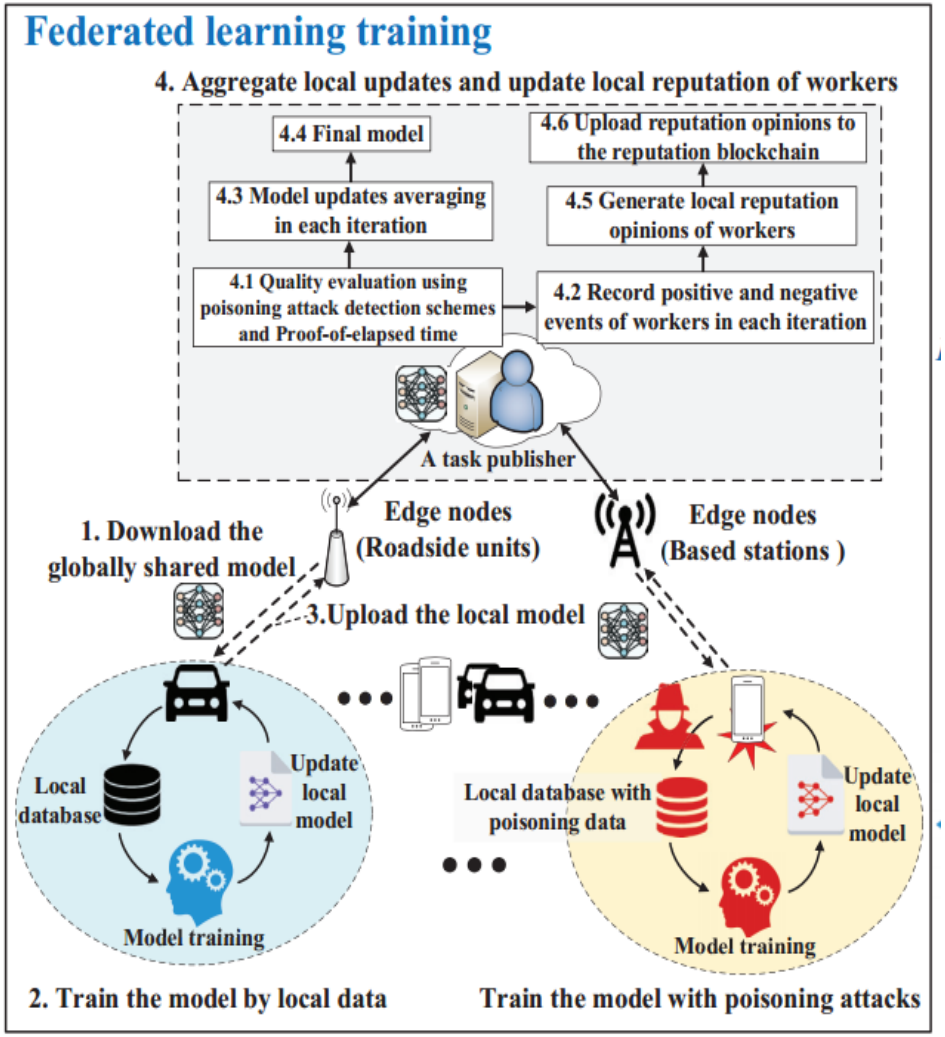
- **Blockchain: a distributed database that maintains a continuously growing list of ordered records called blocks. Anyone can create and complete smart contracts that are stored on the public ledger permanently.**

Consortium Blockchain for Reputation Management

- **Consortium blockchain**: blockchain with multiple authorized nodes to establish the distributed shared ledger with moderate cost
- Reputation values of the workers are securely managed and stored on the consortium blockchain consisting of the edge nodes (e.g., roadside units and base stations).



An Overview of the Proposed Scheme



Step 1: Task Publishment

- Federated learning tasks are first broadcast with specific data requirements (e.g., data sizes, types and time range).
- Mobile devices will send a joining request with identity and data resource information back to task publishers.

An Overview of the Proposed Scheme

- **Step 2: Worker Selection**
  - The task publisher validates the identity and data resource information of the requesters, then the legal requesters can be the worker candidates.
  - The task publisher starts to select its workers from the worker candidates according to their reputation values calculated by the subjective logic model.
  - The worker candidates with reputation values above a threshold can be selected as the workers. The reputation values of the worker candidates are calculated and stored on an open-access consortium blockchain named reputation blockchain.

  More details about blockchain will be given later.

An Overview of the Proposed Scheme

- **Step 3: Reputation Calculation:**
  - The task publisher utilizes the subjective logic model to generate local reputation opinions for the worker candidates based on interaction histories.
  - The subjective logic model takes three weights about the past interactions into consideration to form the local opinions for each worker candidate. By combining the local reputation opinions with recommended reputation opinions, the task publishers generate a composite reputation as the final reputation for each worker candidate.
  - The recommended reputation opinions can be downloaded from the reputation blockchain and obtained from the latest block data.

  More details about reputation calculation are depicted later.

An Overview of the Proposed Scheme

- **Step 4: Federated Learning:**
  - The task publisher sends an initial SGD model to the selected workers. The workers collaboratively train the global model by using their own local data.
  - The workers generate local model updates and the corresponding local computation time and upload this information to the task publisher.
  - The local computation time is used to verify the reliability and authenticity of local model updates by employing the proof of elapsed time method based on Intel's SGX technology.
  - Poisoning attack detection schemes are performed to identify the poisoning attacks and unreliable workers, e.g., Reject on Negative Influence (RONI) scheme [4] for Independent and Identically Distributed (IID) scenarios and the FoolsGold scheme [5] for non-IID scenarios.

[4] M. Shayan *et al.*, "Biscotti: A Ledger for Private and Secure Peer-to-Peer Machine Learning," 2018; available: https://arxiv.org/abs/1811. 09904.
[5] C. Fung *et al.*, "Mitigating Sybils in Federated Learning Poisoning," 2018; available: https://arxiv.org/abs/1808.04866

An Overview of the Proposed Scheme

- With the help of these schemes, the task publisher removes malicious updates from poisoning attacks and unreliable local model updates from the lazy or untrusted workers.
- Then, the task publisher generates a new global model and sends the new global model to the selected workers for the next model iteration.
- The workers obtain rewards from the task publisher according to their data contribution and model training behaviors for the federated learning task.
- During the federated learning process, either the lazy and unreliable workers or the workers with poisoning attackers in each model iteration are recorded as a negative interaction by the task publisher.
- Finally, the task publisher generates local reputation opinions for the workers based on their performance in the federated learning task.

An Overview of the Proposed Scheme

- **Step 5: Reputation Updating:**
  - To achieve secure reputation management, the task publisher's interaction histories and local reputation opinions for the workers with digital signatures are recorded as "transactions" and uploaded to the pre-selected miners in the reputation blockchain.
  - These miners execute consensus algorithms, such as PBFT, and the reputation opinions and interaction histories are stored as a data block to be added into the reputation blockchain.
  - All task publishers can obtain the latest reputation opinions for a certain worker candidate from the reputation blockchain. Lastly, with the help of the reputation blockchain, the task publishers are able to select high-reputation workers for federated learning tasks.

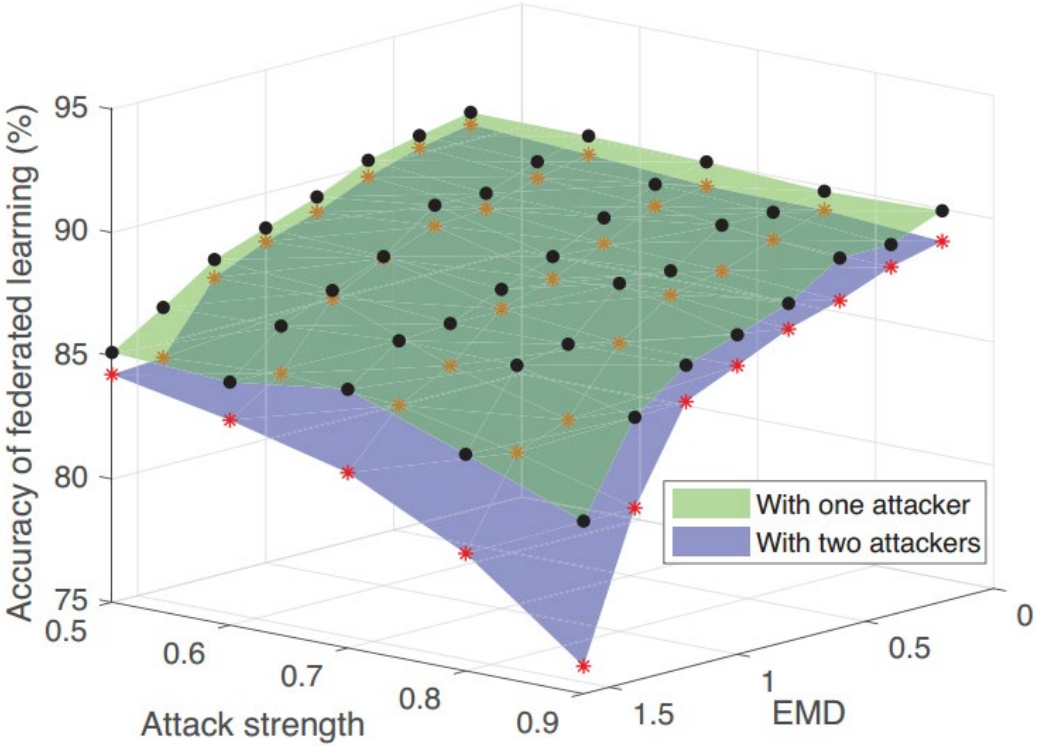
Performance Evaluation

Simulation Setting

- We perform simulation on a well known digit classification dataset named MNIST by using Tensorflow 1.12.0 for a digit classification.
- We establish the reputation blockchain system on the Hyperledger Fabric v1.4.0 and use PBFT algorithm with mild overhead and latency as the consensus algorithm.
- We consider ten workers in this federated learning task including two malicious workers who launch poisoning attacks, four unreliable workers with low-quality data, and four well-behaved workers.
- The Earth Mover's Distance (EMD) is used as a metric to measure training data quality of the unreliable workers.
- We compare the proposed MSL scheme with a Traditional Subjective Logic (TSL) scheme from [6], and an Aggregated Trust Value (ATV) scheme referred to in [7].

[6]J. Kang et al., "Incentive Mechanism for Reliable Federated Learning: A Joint Optimization Approach to Combining Reputation and Contract Theory," *IEEE Internet of Things J.*, vol. 6, no. 6, Dec. 2019, pp. 10700–714.
[7]Z. Yang et al., "Blockchain-Based Decentralized Trust Management in Vehicular Networks," *IEEE Internet of Things J.,* vol. 6, no. 2, April 2019, pp. 1495–1505.

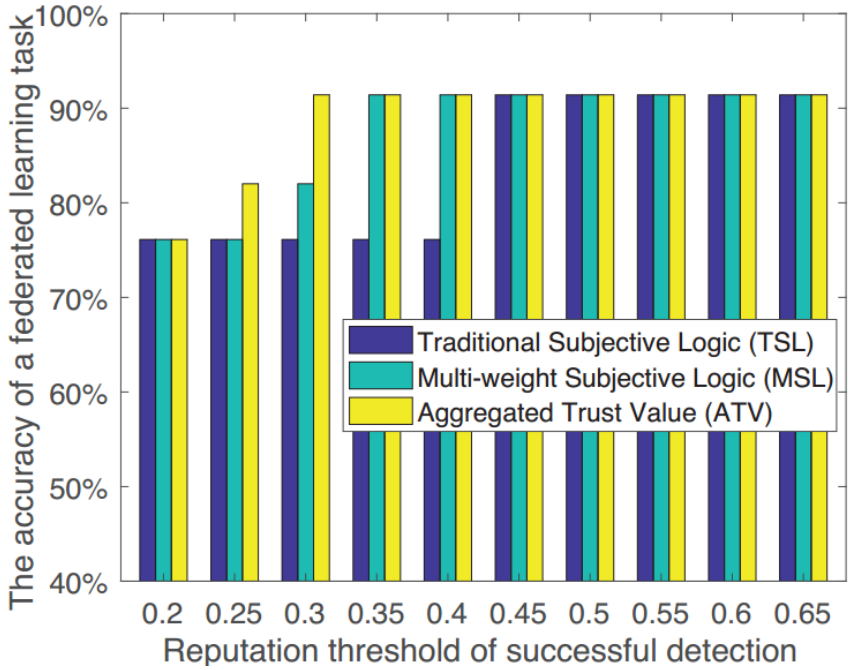
Accuracy of Federated Learning



The accuracy comparison with respect to attack strengths and data quality levels

- There are three factors that affect the learning accuracy: EMD, attacker number, and attack strength. An increase of any one of the above factors leads to a decrease of accuracy.
- The unreliable and untrusted workers with low-quality training data have negative impacts on the accuracy.

Performance of Proposed Reputation Scheme



The impact of reputation thresholds of successful detection on the federated learning accuracy

- Reputation threshold of successful detection: A metric that only the reputation of malicious workers below the threshold can be detected successfully.
- The proposed MSL scheme can achieve a more accurate and fair reputation calculation, thereby leading to a more reliable worker selection in federated learning

Future Work

- Economics and pricing issues, e.g., profit maximization of the FL task publisher and utility maximization of the FL workers
- Reputation management for other reliability issues, e.g., backdoor attack and free riding
- Federated transfer learning
- Consider more general computing paradigms such as Coded Distributed Computing (CDC)

Summary

- We addressed worker selection issues to ensure reliable federated learning in mobile networks.
- A reputation-based scheme was designed to select reliable and trusted workers. For efficient and secure reputation management, we calculated workers' reputation by using a multi-weight subjective logic model, and employed consortium blockchain to manage the reputation with tamper resistance and non-repudiation in a decentralized manner.
- Numerical results showed that our schemes can bring reliable federated learning to mobile networks.

More details can be found in the journal paper:

J. Kang, Z. Xiong, D. Niyato, S. Xie, and J. Zhang, "Incentive Mechanism for Reliable Federated Learning: A Joint Optimization Approach to Combining Reputation and Contract Theory," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10700–714, Dec. 2019.

Thank you!