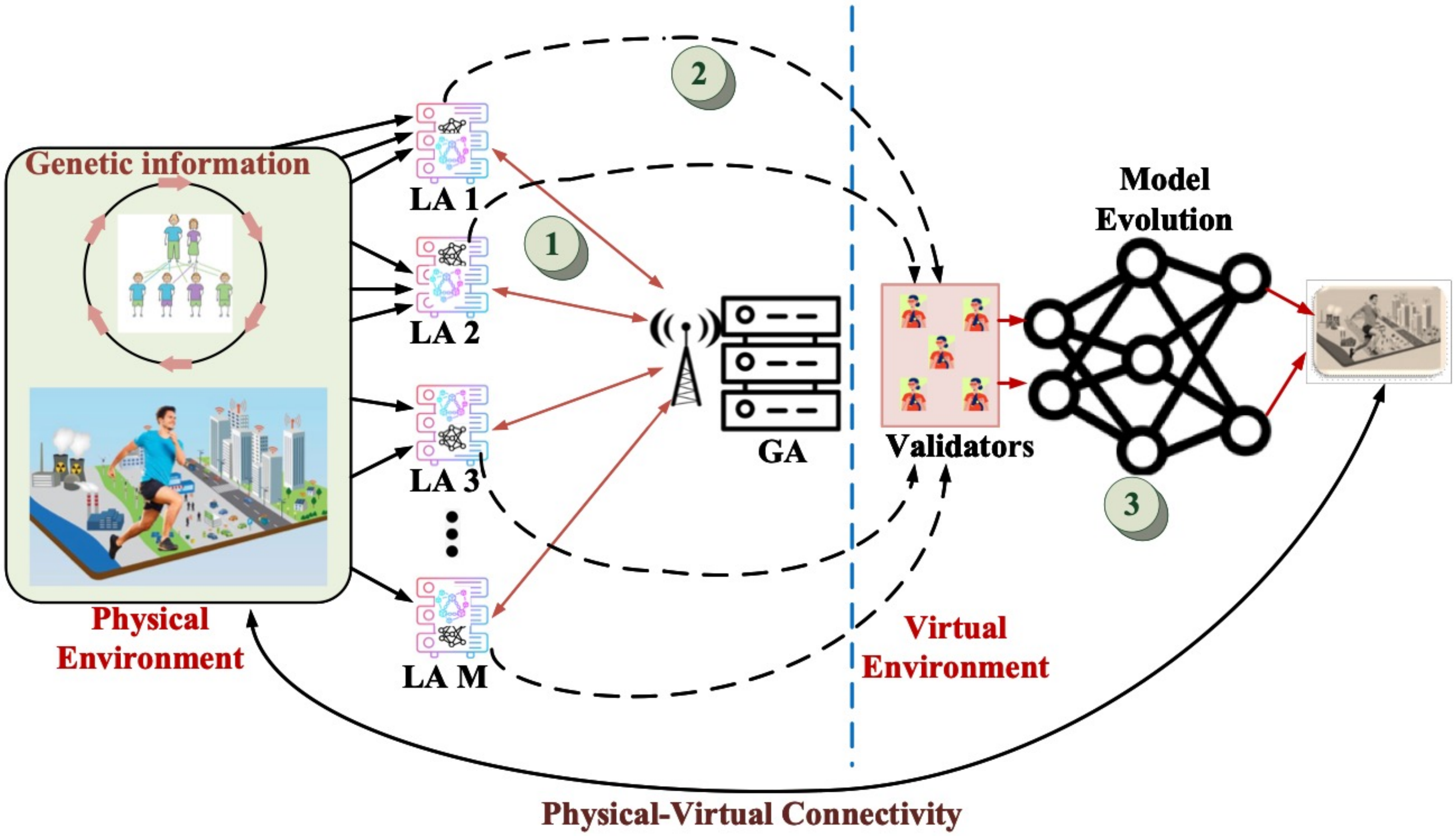
Differentially Private Federated Multi-Task Learning Framework for Enhancing Human-to-Virtual Connectivity in Human Digital Twin

Samuel D. Okegbile, Jun Cai, Hao Zheng, Jiayuan Chen, and Changyan Yi

https://ieeexplore.ieee.org/document/10234396

Differentially Private Federated Multi-Task Learning Framework



Federated Multi-Task Learning

Feature Classifier

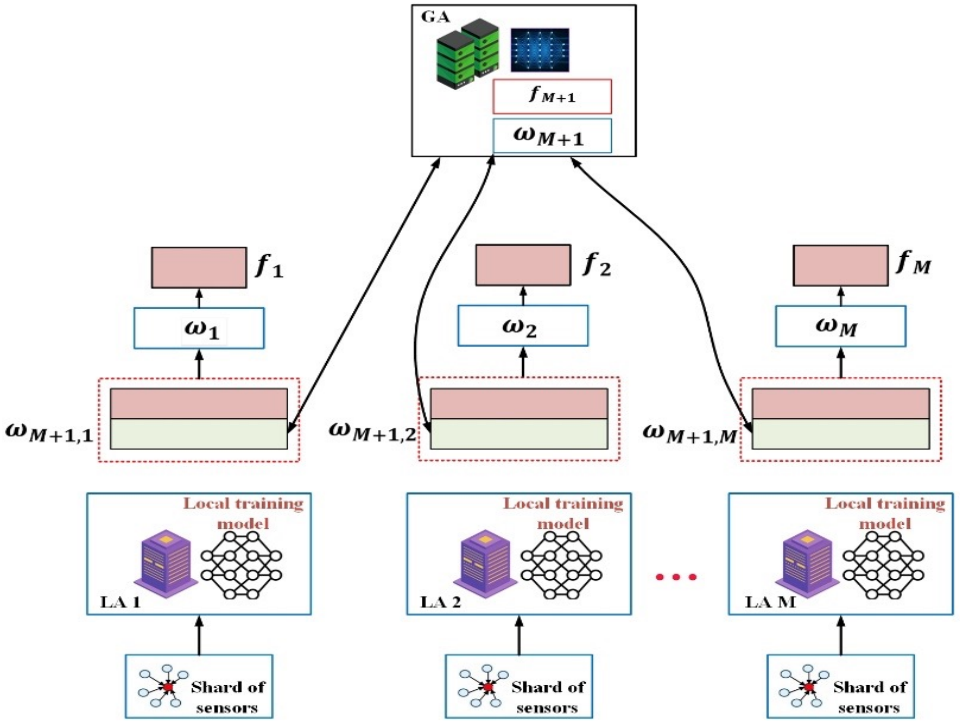
Domain classifier:

$$M_{MD}(i, GA) = \left|\left|\frac{1}{n_i}\sum_{k=1}^{n_i}\Phi(x_i^k) - \frac{1}{n_{GA}}\sum_{l=1}^{n_{GA}}\Phi(x_{M+1}^l)\right|\right|_{\mathcal{H}}^2$$

Logistic regression-based domain classifier:

$$\min_{\Phi(i,\mathrm{GA})}\left|\left|\mathrm{diag}(\mathcal{Q}_i)\left(x_i^k - x_{M+1}^l\Phi(i,\mathrm{GA})\right)\right|\right|_F^2 + \lambda\sum_{l=1}^{n_{GA}}\left|\left|\Phi_l(i,\mathrm{GA})\right|\right|_F^2$$

$$\min_{\Phi(\mathrm{GA},i)}\left|\left|x_i^k - x_{M+1}^l\Phi(\mathrm{GA},i)\right|\right|_F^2 + \lambda\sum_{k=1}^{n_i}\left|\left|\Phi_k(\mathrm{GA},i)\right|\right|_F^2$$



FL Learning

Local training: $$f_i(\omega_i, \omega_{M+1}) = \frac{1}{D_i}\sum_{j=1}^{D_i}\ell(x_{i,j}, y_{i,j}, \omega_i, \omega_{M+1})$$

Global training: $$f(\omega_{M+1}) = \sum_{i=1}^{M}\Omega_i f_i(\omega_{M+1,i}, \omega_{M+1})$$

Differential Privacy (DP)

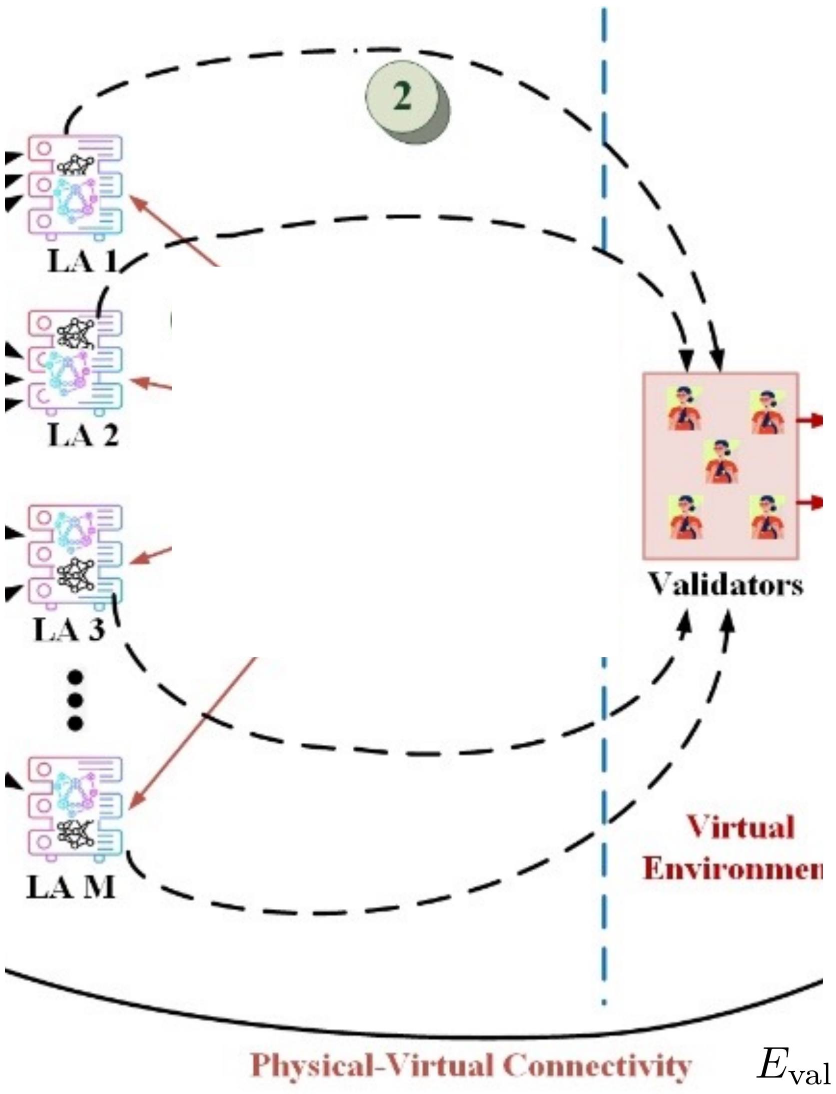
$(\epsilon, \delta) - \mathrm{DP}$

$$n_{M+1,i} \sim \mathcal{N}(0, \sigma^2 G_S^2)$$

Economical loss:

$$\psi = \frac{1}{\epsilon_{\max}}\sum_{i=1}^{M}\sum_{t=0}^{N_R-1}\epsilon_i v_i|\omega_{M+1,i}(t)|$$

Blockchain Based Validation Model



Proof of model quality (PoMQ)

$$R_{m_j}(i) = \begin{cases} 1 & \text{if } C_{\text{tot}}^{\text{off}}(i) \geq \theta_{\text{off}}, C_{\text{tot}}^{\text{cmp}}(i) \geq \theta_{\text{cmp}}, \text{and } C_{\text{tot}}^{\text{pvy}}(i) \geq \theta_{\text{pvy}} \\ 0 & \text{otherwise,} \end{cases}$$

Validation time cost

$$T_{\text{val}}(i) = \frac{|\omega_{f,i}|}{r_i} + \max_{m_j \in [V]} \left\{ \frac{c_v |R_{m_j}(i)|}{c_{m_i}} + \frac{|R_{m_j}(i)|}{r_v} \right\}$$

Transmission time Computation time Decision exchange time

Validation energy cost

$$E_{\text{val}} = \frac{N}{P_i h_{i,\text{val}}} \left[\exp\left(\frac{r_i}{B_0} - 1 \right) \right] + \sum_{m_j=1}^{V} \left\{ \frac{N}{P_j h_{j,k}} \left[\exp\left(\frac{r_v}{B_0} - 1 \right) \right] + \kappa_v c_0 |R_{m_j}(i)| c_{m_i}^2 \right\}$$

Synchronization Accuracy

Synchronization Accuracy = Synchronization gap + FML loss

Synchronization gap = the time since the last status update

Proposition: If C_{time} is i.i.d. exponential in steady-state, the density of S_{gap} at any time t can be obtained as

$$\mathcal{D}^{\mathrm{lcfs}}_{S_{\mathrm{gap}}}(t) = \frac{o_i[(\rho+2)(\rho-1)]t - \rho^2 + \rho + 3}{\rho^3 - 1}\exp(-\varrho_i t) + \frac{o_i(\rho+1)t + \rho(\rho+3) + 3}{\rho(\rho+1)+1}\exp(-\varrho_i[\rho+1]t) - \frac{\rho}{\rho-1}\exp(-o_i t)$$

Synchronization gap

$$S^{\mathrm{lcfs}}_{\mathrm{gap}} = \frac{o_i^4(2o_i + 7\varrho_i) + o_i^2\varrho_i^2(8o_i + 7\varrho_i) + \varrho_i^4(4o_i + \varrho_i)}{o_i\varrho_i(o_i+\varrho_i)^2(o_i^2 + \varrho_i o_i + \varrho_i^2)}$$

	Waiting on the queue	In service
Empty		
a arrives		a
b arrives	b	a
c arrives	c	a
a departs		c

Non-preemptive single-server last-come-first-serve (LCFS) queue with a buffer of size 2 and queue displacement policy

Connectivity Cost

Privacy cost

Long-term average connectivity cost:

$$C_{\text{conn}} = \lim_{\mathcal{U}\to\infty} \frac{1}{\mathcal{U}} \sum_{u=1}^{\mathcal{U}} \Big(\frac{1}{\epsilon_{\max}} \sum_{i=1}^{M} \sum_{t=0}^{N_R-1} \epsilon_i v_i |\omega^u_{M+1,i}(t)| + C^u_{\text{time}} + C^u_{\text{ene}} \Big)$$

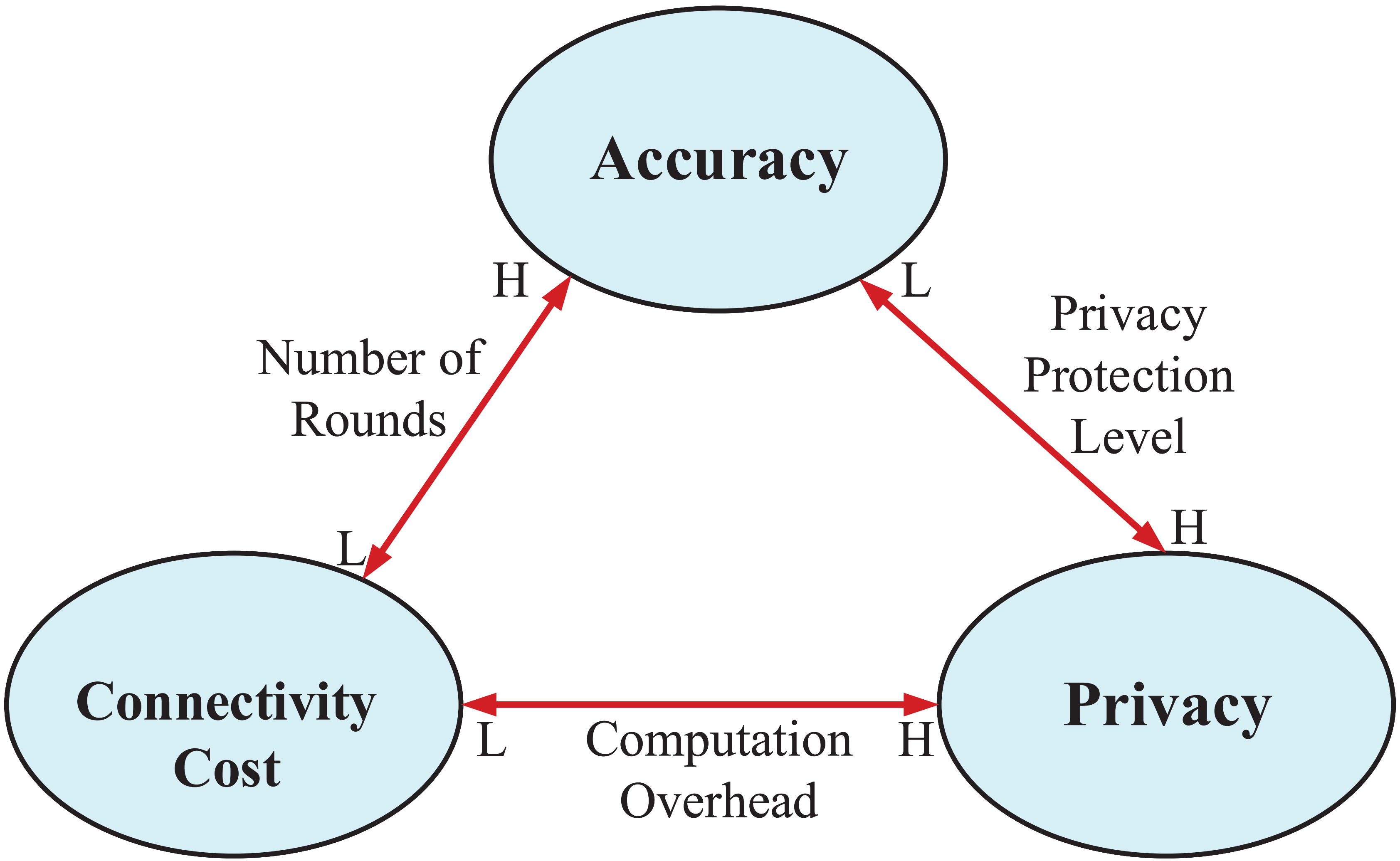
Overall time cost:

$$C_{\text{time}} = \sum_{t=0}^{N_R-1} \Big\{ \Big(\frac{c_{agg} \sum_{i=1}^{M} |\omega_{M+1,i}(t)|}{c_{GA}} \Big) + \max_{i\in[M]} \Big(\frac{|\omega_{M+1,i}|}{r_i} + \frac{c_r D_i(t)}{c_i} \Big) \Big\}$$
$$+ \Big(\frac{|\omega_{f,i}|}{r_i} + \max_{m_j\in[V]} \Big\{ \frac{c_v |R_{m_j}(i)|}{c_{m_i}} + \frac{|R_{m_j}(i)|}{r_v} \Big\} \Big)$$

Overall energy cost:

$$C_{\text{ene}} = \sum_{t=0}^{N_R-1} \Big\{ \kappa_{GA} c_0 \Big(\sum_{i=1}^{M} |\omega_{M+1,i}(t)| c^2_{GA} \Big) + \sum_{i=1}^{M} \Big(\kappa_i c_0 D_i(t) c_i^2 + t_i(t) \frac{N}{h_{i,\text{GA}}(t) P_i(t)} \Big[\exp\Big(\frac{r_i(t)}{B_0} - 1 \Big) \Big] \Big) \Big\}$$
$$+ \frac{N}{P_i h_{i,\text{val}}} \Big[\exp\Big(\frac{r_i}{B_0} - 1 \Big) \Big] + \sum_{m_j=1}^{V} \Big\{ \frac{N}{P_j h_{j,k}} \Big[\exp\Big(\frac{r_v}{B_0} - 1 \Big) \Big] + \kappa_v c_0 |R_{m_j}(i)| c^2_{m_i} \Big\}$$

Tradeoffs



Accuracy
H
L
Number of
Rounds
Privacy
Protection
Level
L
H
Connectivity
Cost
Privacy
L
Computation
Overhead
H

Optimization Formulation

$$\min_{o_i, V, M, \epsilon, N_R} \; [\Theta_1 (C_{\text{time}} + C_{\text{ene}} + f_i(\omega_i, \omega_{M+1})) - (1 - \Theta_1)\Theta_2 \psi_{\text{cost}}],$$

$$\text{s.t.} \quad o_i \in [0, 1],$$

$$\sum_{i \in [M]} i \leq M,$$

$$\sum_{j}^{V} m_j \leq V,$$

$$\epsilon_{\min} \leq \epsilon \leq \epsilon_{\max},$$

$$c_i^{\min} \leq c_i \leq c_i^{\max}, \forall i \in [M],$$

$$c_{m_j}^{\min} \leq c_{m_j} \leq c_{m_j}^{\max}, \forall m_j \in [V]$$

Overall time cost

Overall energy cost

Training accuracy

Privacy cost

o_i: status updating scheduling rate
V: validators
M: participating LAs
ε: privacy budget
N_R: FL rounds

DRL Solution

MDP Formulation

$$(\mathcal{S}^{(t)}, \mathcal{A}^{(t)}, \mathcal{R}^{(t)})$$

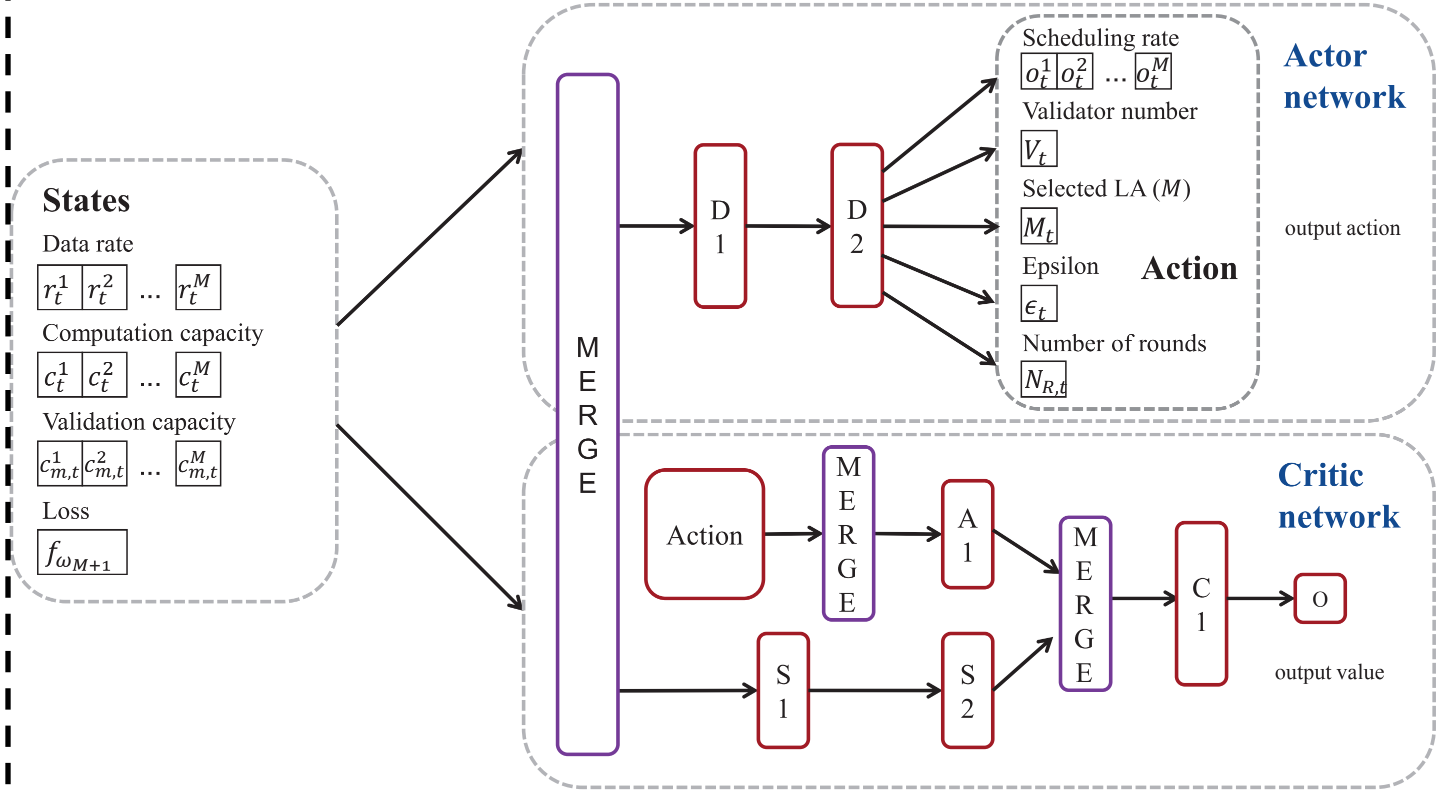
$$\mathcal{S}^{(t)} = \{r(t), c(t), c_m(t), f(\omega_{M+1}(t))\}$$

$$\mathcal{A}^{(t)} = \{o(t), V(t), M(t), \epsilon(t), N_R(t)\}$$

$$\mathcal{R}^{(t)} = -O(t)$$

$$\min \mathbb{E}\Big[\sum_{t=0}^{N_R-1} \gamma \mathcal{R}(\mathcal{S}^{(t)}, \mathcal{A}^{(t)})\Big]$$

DDPG Solution



Simulation Results

FeDAvg: conventional FL

DPFedAvg: FeDAvg+DP

VDPFML: DPFML+PoS