PRIVACY PROTECTION IN LEARNING FAIR REPRESENTATIONS

YULU JIN AND LIFENG LAI UNIVERSITY OF CALIFORNIA, DAVIS, ECE DEPARTMENT

INTRODUCTION

As the number of IoT devices being introduced in the market has increased dramatically, inference as service (IAS) has been widely used in many sensitive environments to make decisions in the cloud [1]. In IAS, devices will send data to cloud and machine learning algorithms can be run on the cloud providers' infrastructure where training and deploying machine learning models are performed on cloud servers. However, two important issues, namely data privacy and fairness, need to be properly addressed.

Our goal is to address the fairness and privacy issues simultaneously in the IAS design based on our previous work [2]. Instead of sending data directly to the server, the user will pre-process the data through a transformation map. Then we analyze the trade-off among data utility, fairness representation and privacy protection, formulate an optimization problem, and design an iterative algorithm to find the optimal transformation map.

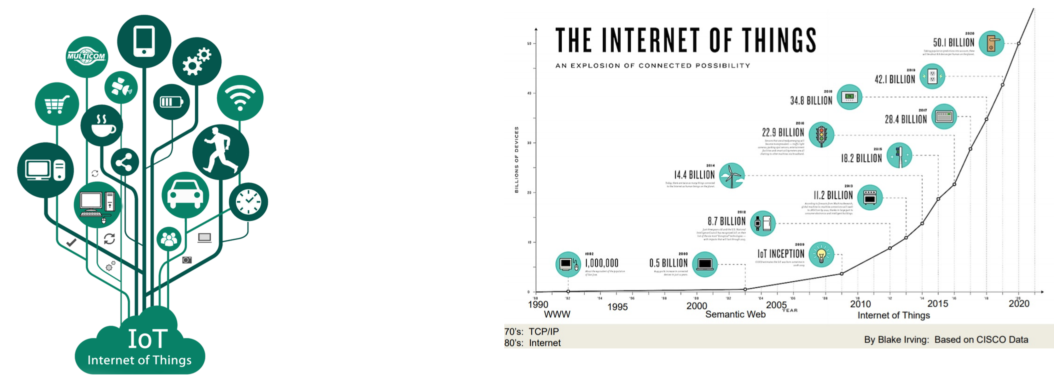

Figure 1: Internet of Things

ACKNOWLEDGEMENT

This work was supported in part by National Science Foundation under Grants CCF-1717943, CNS-1824553, CCF-1908258 and ECCS-2000415. Email:{yuljin,lflai}@ucdavis.edu.

REFERENCE

[1] A. Gujarati. Swayam: Distributed autoscaling to meet slas of machine learning inference services with resource efficiency. In *Proc. ACM Conference*, pages 109–120, Las Vegas, NV, Dec. 2017.

[2] Yulu Jin and Lifeng Lai. Privacy-accuracy trade-off of inference as service. In *Proc. International Conference on Acoustics, Speech and Signal Processing*, pages 2645–2649, Toronto, Canada, Jun. 2021.

PROBLEM FORMULATION

Consider an inference problem, in which one would like to infer the parameter $S \in \mathcal{S}$ of data $Y \in \mathcal{Y}$, where $\mathcal{Y}$ is a finite set. At the meantime, there is a sensitive attribute Z which contains sensitive information such as race, gender etc. Instead of sending Y directly to the server, we will learn a transformation map from Y to $U \in \mathcal{U}$, and send U to the server. The server will use U to conduct the inference task and the transformation mapping serves two purposes: fair presentation and privacy protection.

The optimization problem is

$$\max_{P_{U|Y}} \ \mathcal{F}[P_{U|Y}] \triangleq I(S;U) - \beta \mathbb{E}_{Y,U}\left[f\left(\frac{p(u|y)}{p(u)}\right)\right] - \alpha I(Z;U), \quad (1)$$

$$\text{s.t. } p(u|y) \geq \epsilon, \forall y, u, \sum_u p(u|y) = 1, \forall y \in \mathcal{Y},$$

where $d(y,u) = f(\frac{p(y)}{p(y|u)})$ and f is a continuous function defined on $(0, +\infty)$.

PROPOSED METHODS

As the objective function in (1) is a complicated non-convex function of $P_{U|Y}$, we first transform the maximization over single argument to an alternative maximization problem over multiple arguments. Then the Alternating Direction Method of Multipliers(ADMM) method is introduced to solve the sub-problems.

The objective function in (1) can be rewritten as

$$\mathcal{F}[P_{U|Y}, P_U, P_{Z|U}, P_{S|U}] = I(S;Y) + \beta \mathbb{E}_{Y,U}[d(y,u)] - \sum_{u,y} p(y)p(u|y) D_{KL}[p(s|y) \parallel p(s|u)] - \alpha I(Z;U).$$

For consistency, we require

$$p(u) = \sum_y p(u|y)p(y), \forall u, \quad (2)$$

$$p(z|u) = \frac{\sum_y p(u|y)p(z,y)}{p(u)}, \quad (3)$$

$$p(s|u) = \frac{\sum_y p(u|y)p(s,y)}{p(u)}. \quad (4)$$

Lemma 1 *Suppose that $f(\cdot)$ is a strictly convex function. Then for given $P_U, P_{Z|U}, P_{S|U}$, $\mathcal{F}[P_{U|Y}, P_U, P_{Z|U}, P_{S|U}]$ is concave in each $P_{U|y_i}, \forall y_i \in \mathcal{Y}$. Similarly, for given $P_{U|Y}, P_{Z|U}, P_{S|U}$, $\mathcal{F}[P_{U|Y}, P_U, P_{Z|U}, P_{S|U}]$ is concave in P_U. For given $P_{U|Y}, P_U, P_{S|U}$, $\mathcal{F}[P_{U|Y}, P_U, P_{Z|U}, P_{S|U}]$ is concave in $P_{Z|U}$. For given $P_{U|Y}, P_U, P_{Z|U}$, $\mathcal{F}[P_{U|Y}, P_U, P_{Z|U}, P_{S|U}]$ is concave in $P_{S|U}$.*

Under this property, we convert the original optimization problem to

$$\max_{P_{S|U}} \max_{P_{Z|U}} \max_{P_U} \max_{P_{U|Y}} \mathcal{F}[P_{U|Y}, P_U, P_{Z|U}, P_{S|U}].$$

$$\text{s.t. } p(u|y) \geq \epsilon, \forall y, u, \sum_u p(u|y) = 1, \forall y,$$

$$p(u) > 0, \forall u, \sum_u p(u) = 1, (2),$$

$$p(z|u) \geq 0, \forall u, z, \sum_z p(z|u) = 1, \forall u, (3),$$

$$p(s|u) \geq 0, \forall u, s, \sum_s p(s|u) = 1, \forall u, (4).$$

Then we find the solution to (1) iteratively.

In the first step, given $P_{S|U}^{(j-1)}$ and $P_{Z|U}^{(j-1)}$, we apply ADMM to solve

$$\max_{P_{U|Y}} \max_{P_U} \ \mathcal{F}[P_{U|Y}, P_U | P_{S|U}^{(j-1)}, P_{Z|U}^{(j-1)}],$$

$$\text{s.t. } p(u|y) \geq \epsilon, \forall y, u, \sum_u p(u|y) = 1,$$

$$\forall y, p(u) > 0, \forall u, \sum_u p(u) = 1,$$

$$\delta(u) = p(u) - \sum_y p(u|y)p(y) = 0, \forall u.$$

In the second step, we obtain $P_{Z|U}^{(j)}$ by the consistency equation (3).

In the third step, obtain $P_{S|U}^{(j)}$ by solving

$$\max_{P_{S|U}} \ \mathcal{F}[P_{S|U} | P_{U|Y}^{(j)}, P_U^{(j)}, P_{Z|U}^{(j)}],$$

$$\text{s.t. } p(s|u) \geq 0, \forall u, s, \sum_s p(s|u) = 1, \forall u, (4),$$

which has a closed form solution that is the same as the consistency equation (4).

ALGORITHM

Algorithm 1 Design the optimal transformation map

Input:
Prior distribution P_S, P_Z and conditional distribution $P_{Y|S,Z}$.
Trade-off parameter α, β.
Converge parameter η, η_p.

Output:
A mapping $P_{U|Y}$ from $Y \in \mathcal{Y}$ to $U \in \mathcal{U}$.

Initialization:
Randomly initiate $P_{U|Y}$ and calculate $P_U, P_{Z|U}, P_{S|U}$ by (3), (4) and (5).

$j = 1$.
while $\left\| P_{S|U}^{(j)} - P_{S|U}^{(j-1)} \right\|_F > \eta$ **do**
  $P_U^{(j),1} = P_U^{(j-1)}$.
  $P_{U|Y}^{(j),1} = P_{U|Y}^{(j-1)}$.
  $t = 1$.
  while $t = 1$ or $\left\| P_U^{(j),t} - P_U^{(j),t-1} \right\|_{\ell_1} > \eta_p$ **do**
    Update $P_{U|y_i}$ by solving (6).
    Update P_U by solving (7).
    Update Λ by (8).
    $t = t + 1$.
  Update $P_{Z|U}^{(j)}$ by (4).
  Update $P_{S|U}^{(j)}$ by (5).
  $j = j + 1$.
return $P_{U|Y}$

NUMERAICAL RESULT

Set the prior distribution $p_s = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$ and let $|\mathcal{Y}| = 10, |\mathcal{U}| = 11$.

The conditional distributions $p(y|s)$ under each s are shown below

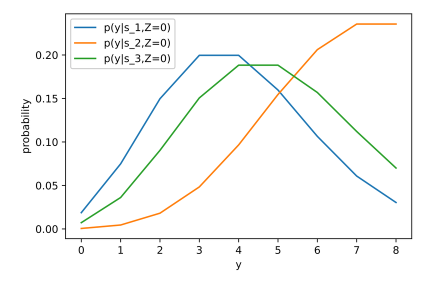

Figure 2: $p(y|s, Z = 0)$

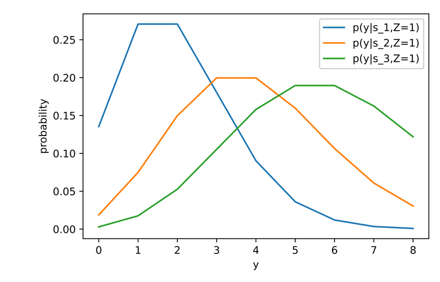

Figure 3: $p(y|s, Z = 1)$

Then we perform both Algorithm 1 and GA to find the optimal transition mapping $p(u|y)$.

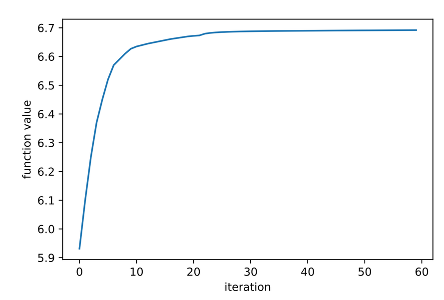

Figure 4: Convergence process of Algorithm 1

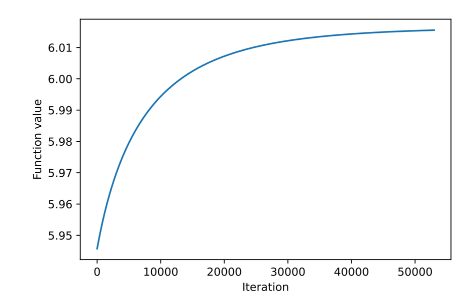

Figure 5: Convergence process of GA

CONCLUSION

We have explored the utility, fairness and privacy trade-off in IAS scenarios under sensitive environments. We have formulated an optimization problem to find the desirable transformation map. We have transformed the formulated non-convex optimization problem and designed an iterative method to solve it. Moreover, we have provided numerical results showing that the proposed method can mitigate the bias and has better performance than GA in the convergence speed, solution quality and algorithm stability.