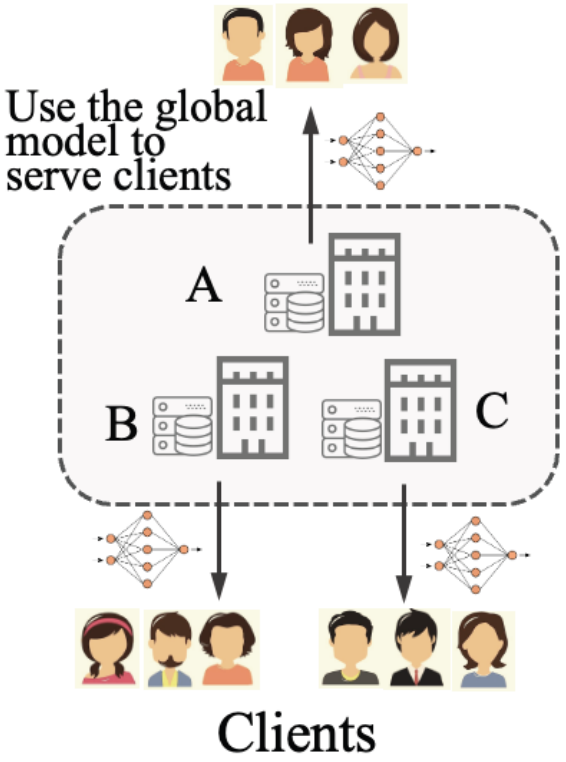
All Federated or Not: Optimizing Personal Model Performance in Cross-silo Federated Learning

Juan Li, Yanmin Zhu, Jie Wu, Weifan Wu, Tianzi Zang, Liu Lu

Outline

- Introduction
- Problem Formulation
- Personal Model Performance Improvement Estimation
- Design of Coalition Partition Algorithm
- Experimental Evaluation
- CONCLUSION

Background



In **cross-silo FL**, Organizations are participants performing local training. They own the global model and use it to serve their clients.

Each organization cares most about its personal model performance.

the performance of the global model on its personal data distribution.

Motivations

One mode of Coalition

- {A,B} Distribution 1
- {C,D} Distribution 2

Organization	Local Accuracy	All Federated Accuracy	Coalition Accuracy
A	0.83	0.80 ↓	0.84 ↑
B	0.75	0.79 ↑	0.82 ↑
C	0.94	0.92 ↓	0.94 ↑
D	0.73	0.91 ↑	0.93 ↑

how to arrange organizations into coalitions to improve their personal model performance?

Problem Formulation

how to arrange organizations into coalitions to improve their personal model performance?

- **Estimate personal model performance improvement**
- **Solve the coalition formation problem**

Personal Model Performance Improvement(utility): The difference between the personal model performance obtained by joining coalition C and by training the model alone for an organization i.

Real value can only be known after FL is done!

Organizations' utility

social welfare

How to find a **stable** partition?

Nash stable partition

No player has an incentive to change its coalition because it cannot get a higher utility

Individually stable partition

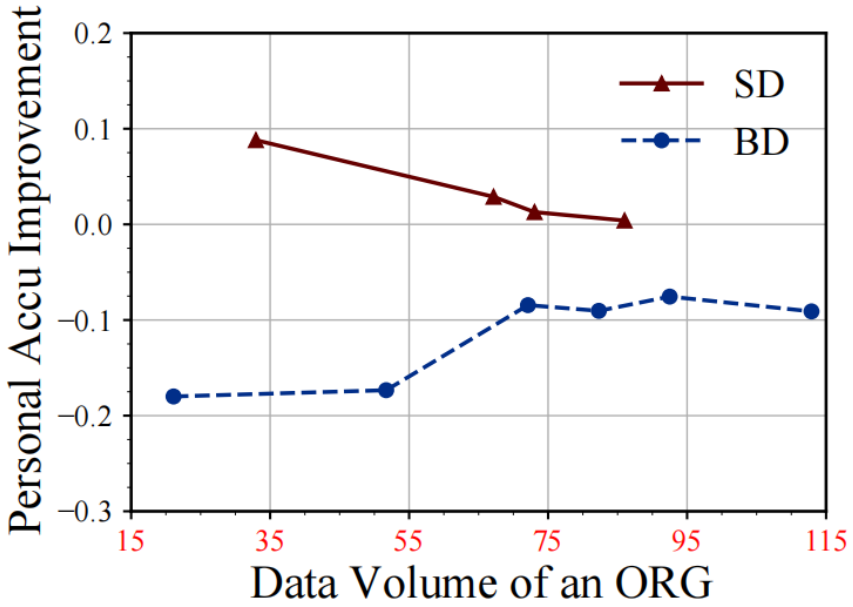
No player can change its coalition to coalition C for a higher utility with the agreement of all members of C

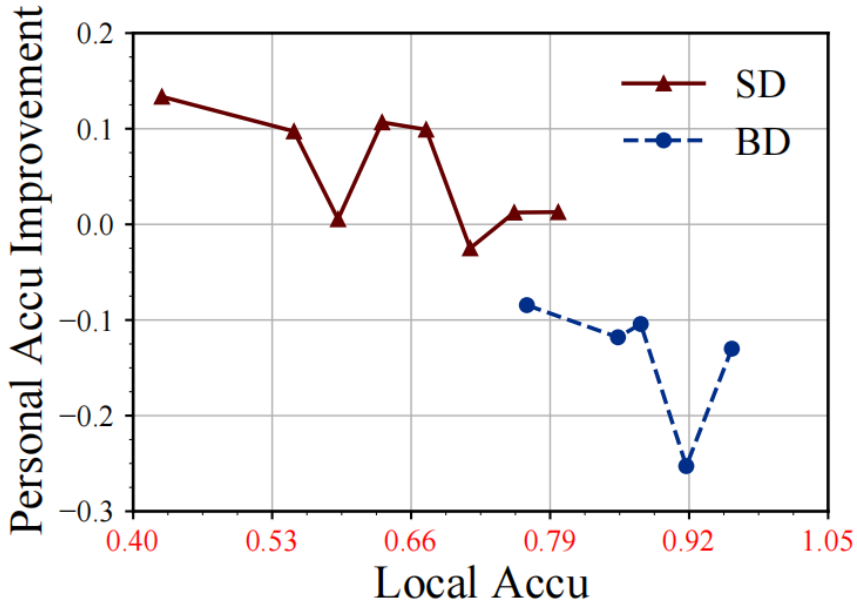
Critical Factors Influencing Utility

Experiment settings

Dataset： MNIST
BD： Bigger data Difference
SD ： Smaller Difference

Factors about organization i:
training data volume,
local model performance





Critical Factors Influencing Utility

Experiment settings

Dataset： MNIST

BD： Bigger data Difference

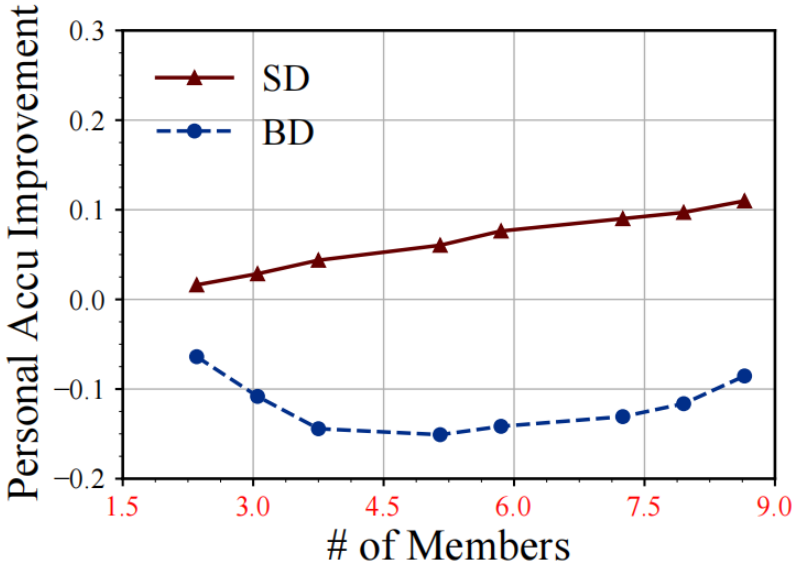
SD ： Smaller Difference

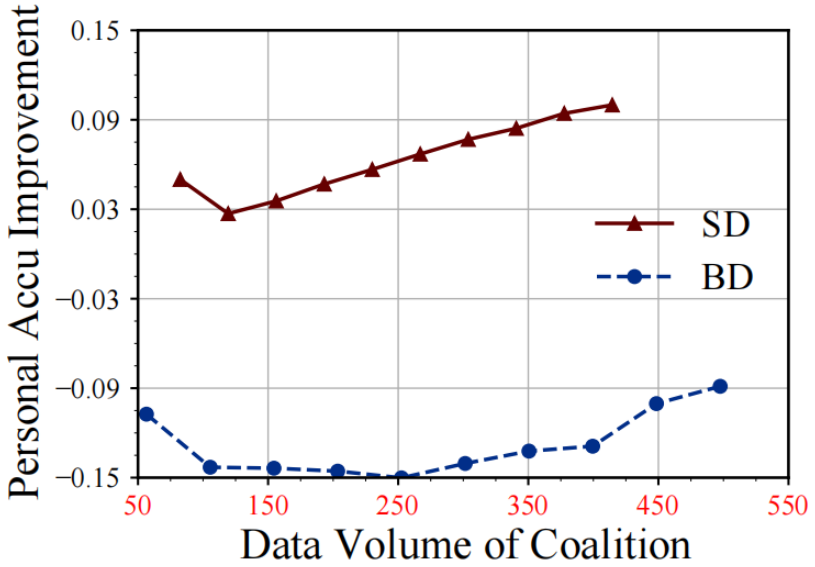
Factors about coalition C:

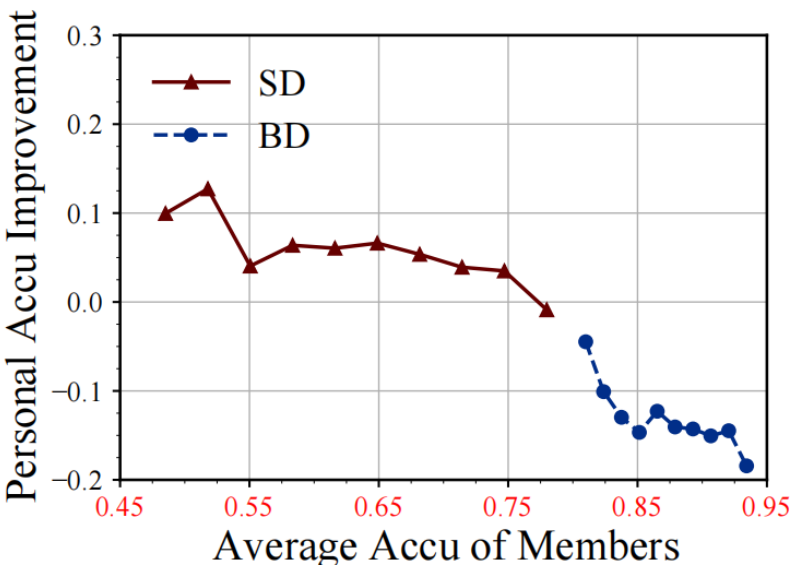
the number of members of C,

the total data volume of C,

the average local accuracy of all members of C





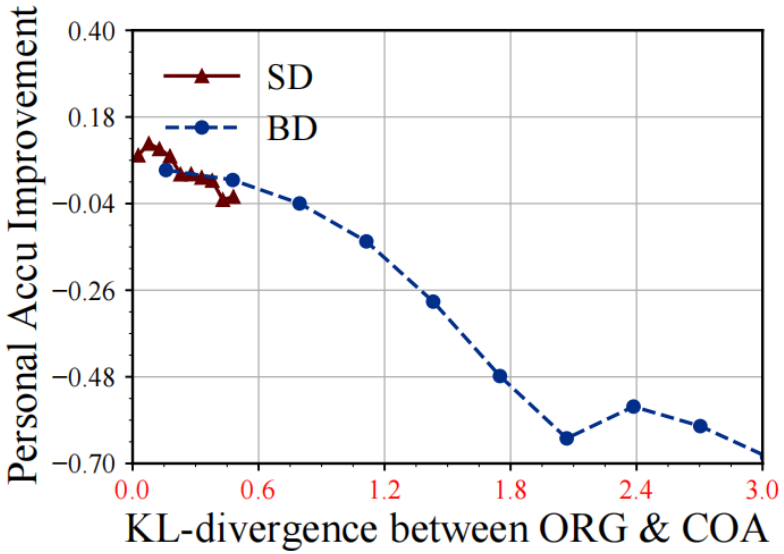


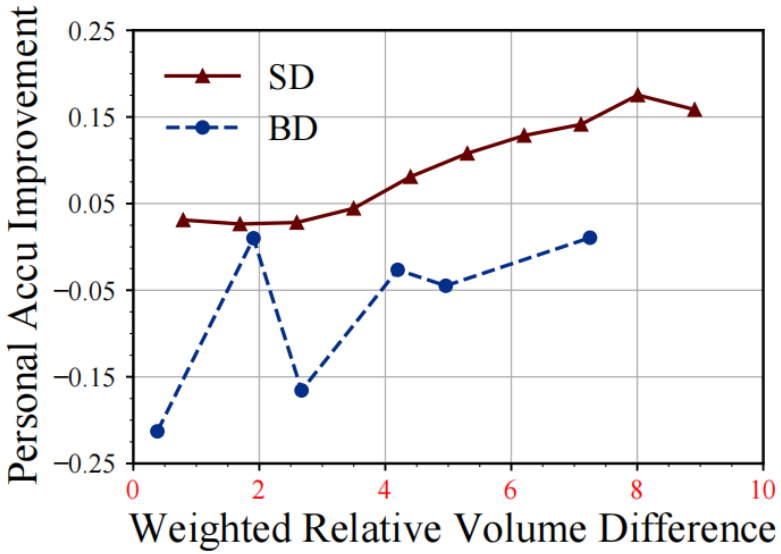
Critical Factors Influencing Utility

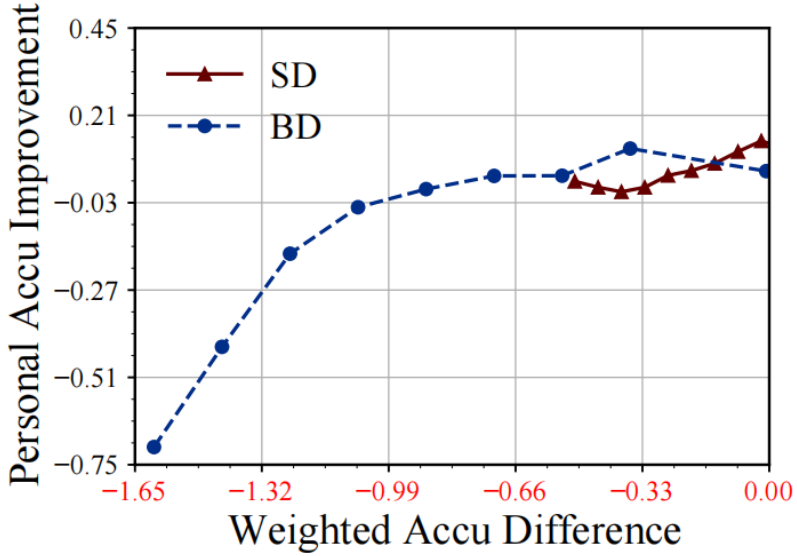
Experiment settings

Dataset： MNIST
BD： Bigger data Difference
SD ： Smaller Difference

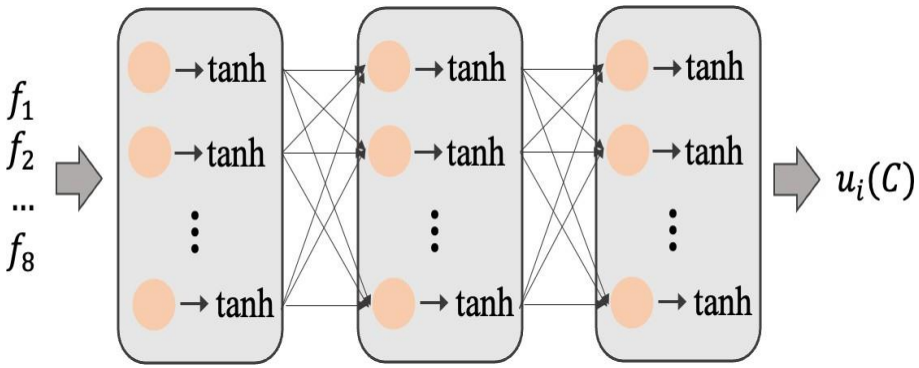
Factors about difference between organization i and coalition C:
KL-divergence,
the weighted relative volume difference,
the weighted accuracy difference







Utility Fitting by Neutral Networks



Input: 8 critical factors {f1, ..., f8} we have selected.

Output: Estimated utility for organization i by joining coalition C.

Finding Individually Stable Partition

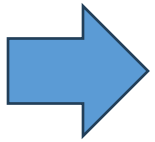
it is A's turn

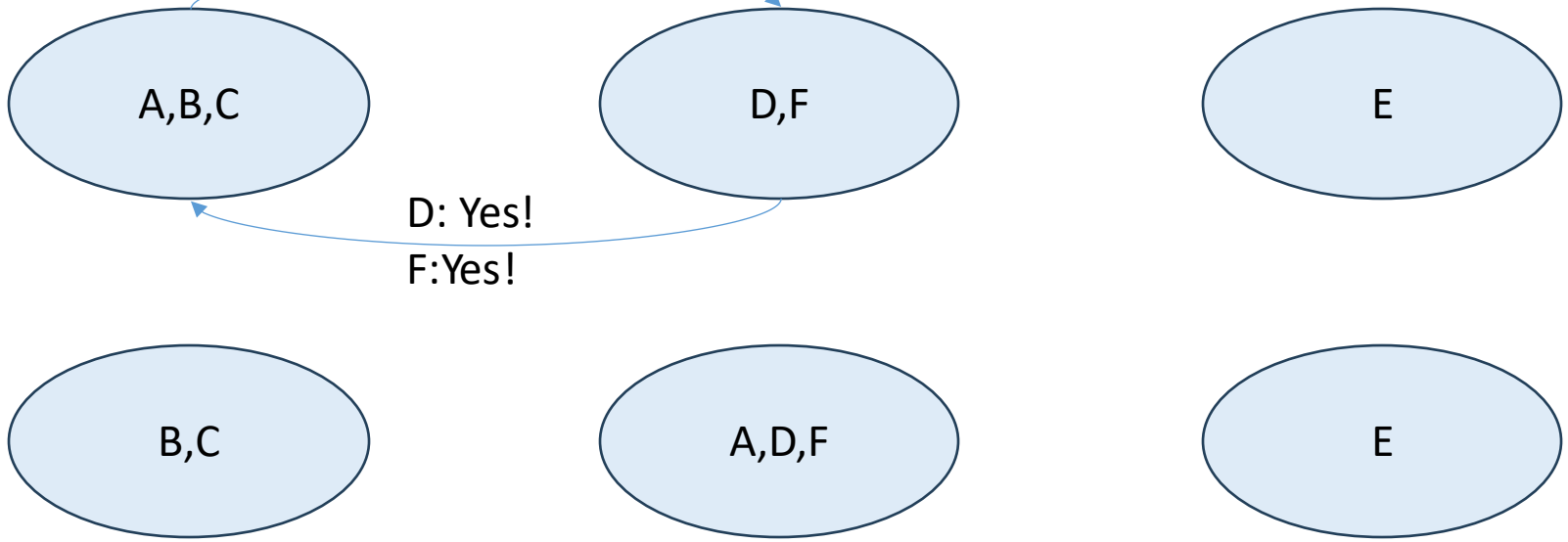
A find its estimated utility

$u_A(\{A, D, F\}) > u_A(\{A, B, C\})$

D:$u_D(\{A, D, F\}) \geq u_D(\{D, F\})$

F:$u_F(\{A, D, F\}) \geq u_F(\{D, F\})$



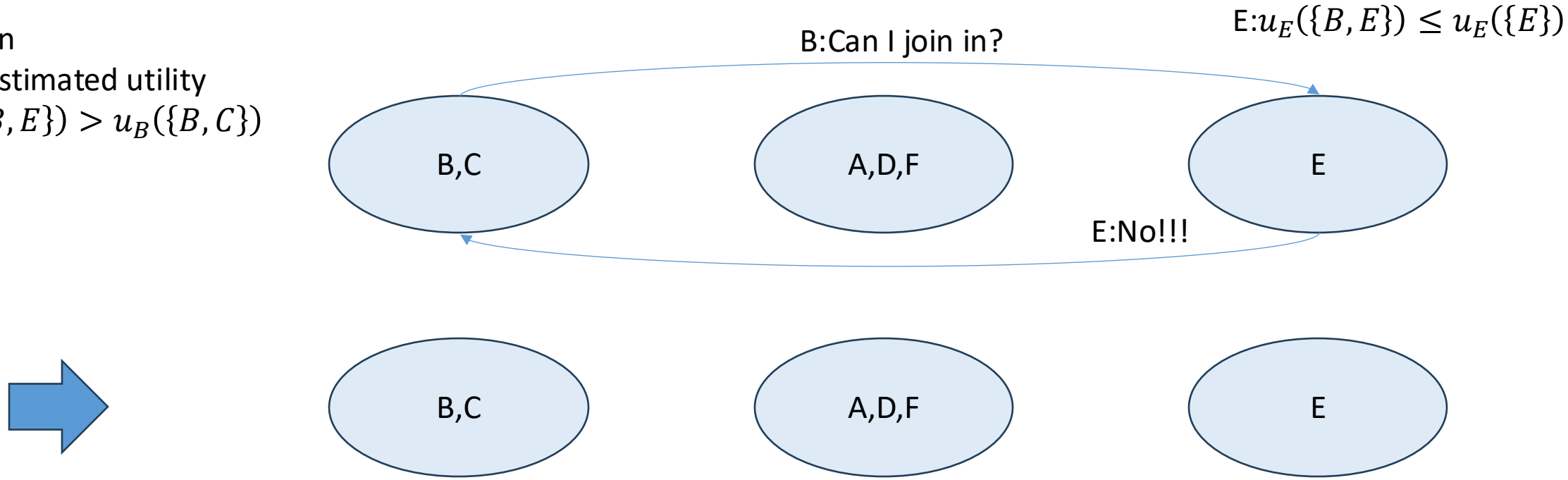


Finding Individually Stable Partition

it is B's turn

B find its estimated utility

$u_B(\{B,E\}) > u_B(\{B,C\})$

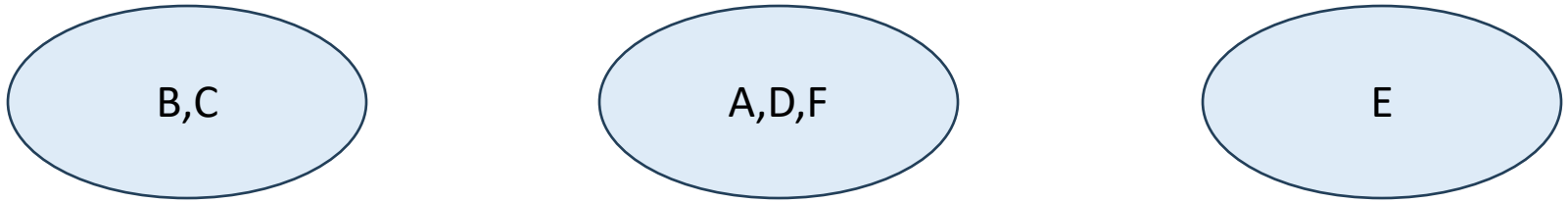


Design of Coalition Partition Algorithm

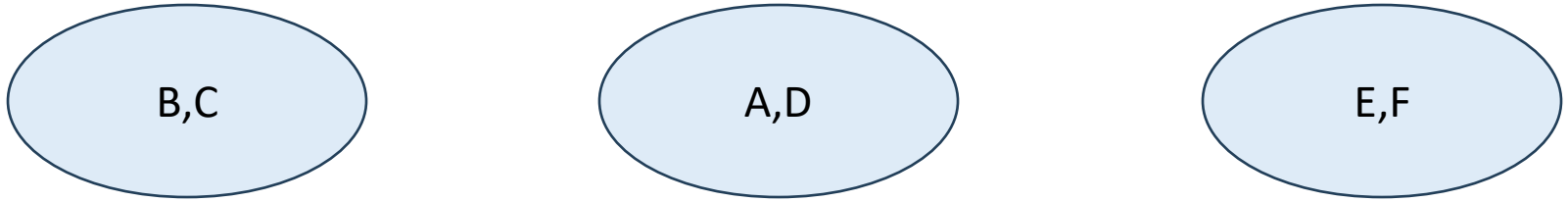
Finding Individually Stable Partition

it is C's turn
C find its estimated utility is max when it remains still

C: Current is perfect!



…. After D,E,F made their choices in turns as well , **One round** has been completed, and the current partition is:



Now it's A's turn again.
A finds it wants to join {E,F}, and {E,F} agree.
So a new round begin, and the process aboveis **repeated until in one round no one moves.**

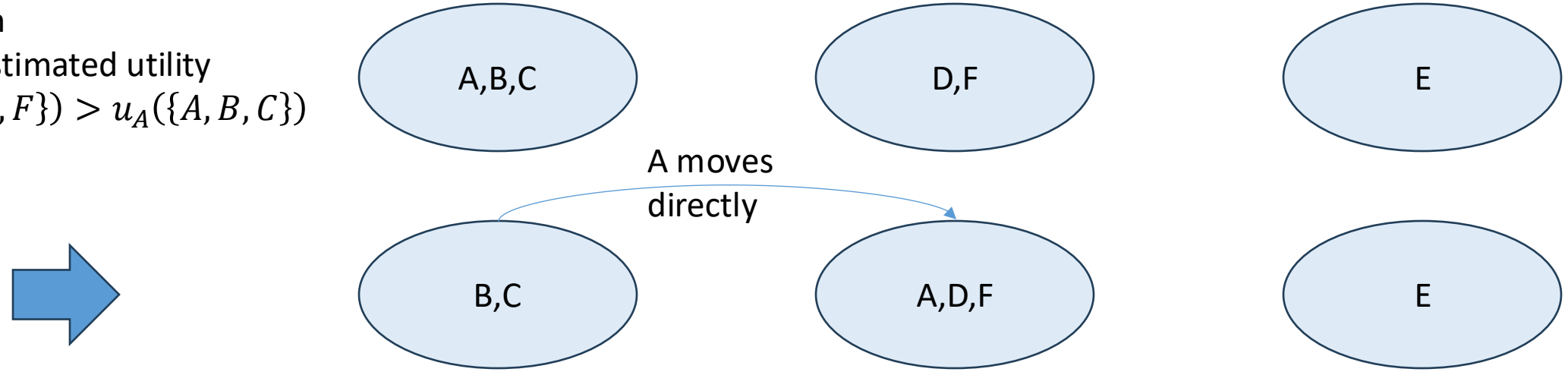
When finding Nash Stable Partition

Finding Nash Stable Partition

When finding Nash Stable Partition, An organization A joining a coalition **does not need to obtain the agreement** of other organizations in this coalition.

If members of the coalition find that they have a better choice after joining an organization, they can also move directly to the ideal coalition in its turn.

it is A's turn
A find its estimated utility
$u_A(\{A, D, F\}) > u_A(\{A, B, C\})$



It is a little difficult to find a **Nash Stable Partition** than find a **Individually Stable Partition.**

Improve Social Welfare of Stable Partitions

If there are multiple stable solutions which one is the **optimal**?
How to maximize the **social welfare**?

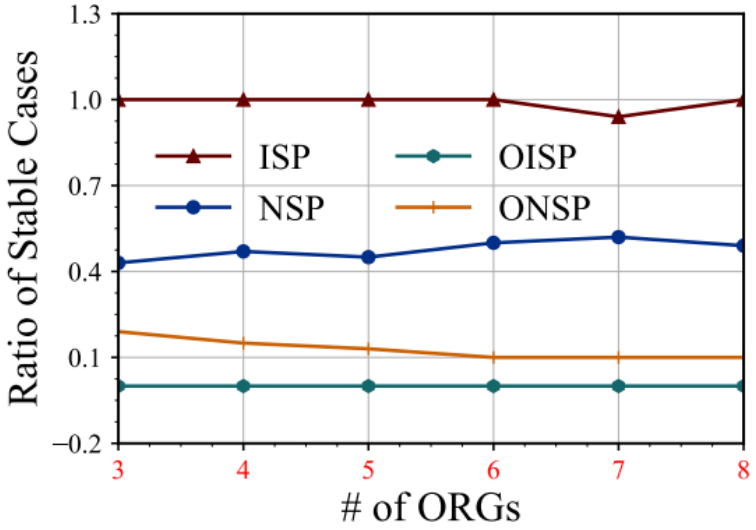
set the **initial coalition structure** as the optimal structure

We find the optimal partition with the highest social welfare based on dynamic programming.

Experimental Evaluation

The convergence performance

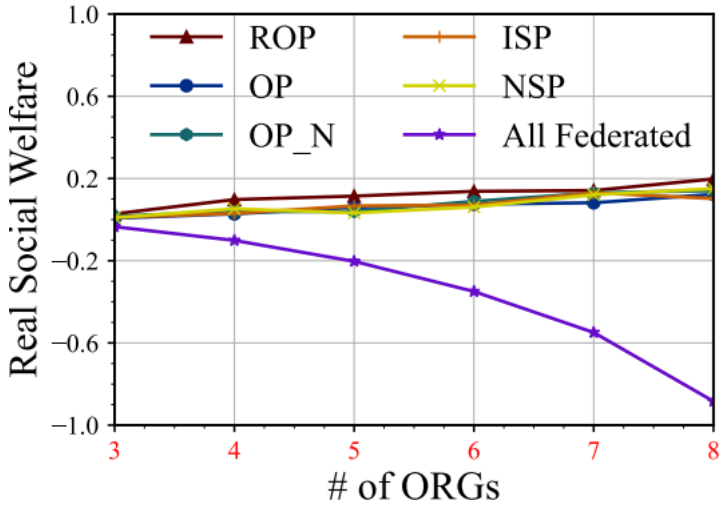
- **ISP** : Individually Stable Partitions
- **NSP**: Nash Stable Partitions
- **ONSP**: Other Nash Stable Partitions
- **OISP**: Other Individually Stable Partitions



It is better to pursue ISP as there exists one and only one ISP in most cases.

Social satisfaction

- **OP:** the Optimal Partitions found by us
- **OP_N**: the average social welfare of OPs when Nash stable partitions exist
- **ROP** :the Optimal Partition found according to the Real individual utilities
- **All Federated** :all organizations form one coalition

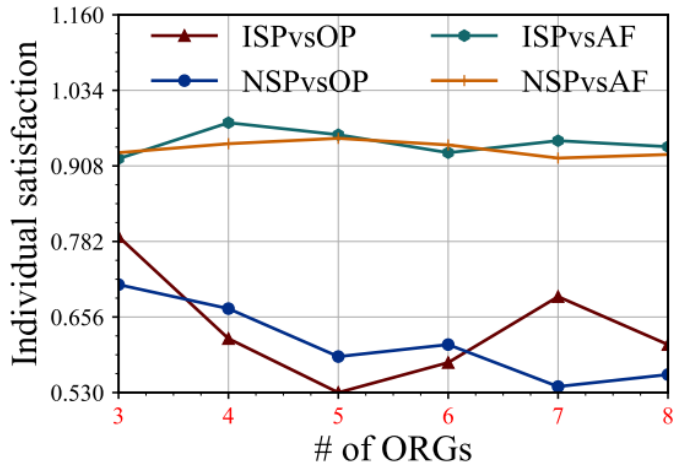


The social welfares of OP and ISP are close to ROP

Experimental Evaluation

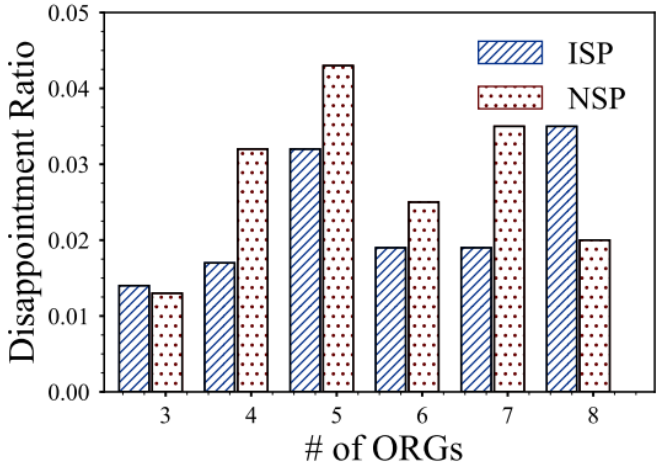
Individual satisfaction

Individual satisfaction of partition Π_1 vs Π_2 is defined as the the ratio of the number of the organizations whose utility in partition Π_1 is more than its utility in partition Π_2 to the total number of the organizations.



ISP and NSP found by our algorithms are better than both OP and AF (All Federated together)

If an organization **is told that it can get a positive utility** as long as it follows the stable solution, **but it gets a negative utility** after FL is done, the organization would **be disappointed**.



The ratio of disappointed organizations is less than 5%.

Conclusion

- We solve a coalition formation problem in cross-silo federated learning to optimize the personal model performance for each organization.
- We first make use of previous FL results to train a neural network which estimates the utility function of each organization for the next time of FL.
- Based on the estimation results, we help organizations form stable coalitions by a distributed algorithm. The found stable coalition structure is close to the optimal one.
- The solution performs well with respect to both real social welfare and individual satisfaction.

Thank you!