Efficient Sampling Approaches to Shapley Value Approximation

JIAYAO ZHANG, Zhejiang University, China
QIHENG SUN, Zhejiang University, China
JINFEI LIU*, Zhejiang University; ZJU-Hangzhou Global Scientific and Technological Innovation Center, China
LI XIONG, Emory University, USA
JIAN PEI, Duke University, USA
KUI REN, Zhejiang University, China

SIGMOD 2023

Shapley Value

Lloyd Shapley

2012诺贝尔经济学奖获得者

- Shapley Value：合作博弈论概念
  - 衡量多个参与者对工作的贡献

Let $N = \{z_1, \ldots, z_n\}$ be a finite set of players and let $\mathcal{U}: 2^N \to \mathbb{R}$ be a "game value" function for sets of players. For each player z_i:

$$SV_i = \frac{1}{n} \sum_{S \subseteq N \setminus \{z_i\}} \frac{\mathcal{U}(S \cup \{z_i\}) - \mathcal{U}(S)}{\binom{n-1}{|S|}}$$

- 从空集开始随机地添加参与者，当添加z_i时引起$\mathcal{U}$函数值变化，z_i的Shapley值就是这一变化的期望
- 所有参与者的Shapley值之和为$\mathcal{U}(N)$

Shapley Value

- Shapley value is the unique allocation of payment that satisfies all requirements in Shapley fairness
  - Balance: the payment is fully distributed to all players
  - Symmetry: same marginal contribution, same payment
  - Zero element: no marginal contribution, no payment
  - Additivity: payments on individual tasks sums up to the payment on a combined task

Shapley Value

Permutation	A	B	C
ABC	7-0	17-7	24-17
ACB	7-0	24-15	15-7
BAC	17-4	4-0	24-17
BCA	24-14	4-0	14-4
CAB	15-6	24-15	6-0
CBA	24-14	14-6	6-0
Shapley value	(7+7+13+10+9+10)/6	(10+9+4+4+9+8)/6	(7+8+7+10+6+6)/6

Monte Carlo Sampling

- Approximate the Shapley value of z_i by sampling permutations

  ① Uniformly sample a permutation over $[n]$

  ② Compute marginal contributions by adding z_i over all players preceding it in permutation

  ③ Repeat k times and output average marginal contributions

Permutation	A	B	C
ABC	7-0	17-7	24-17
ACB	7-0	24-15	15-7
BAC	17-4	4-0	24-17
BCA	24-14	4-0	14-4
CAB	15-6	24-15	6-0
CBA	24-14	14-6	6-0
	$(7+13+10)/3$	$(10+4+8)/3$	$(7+7+6)/3$

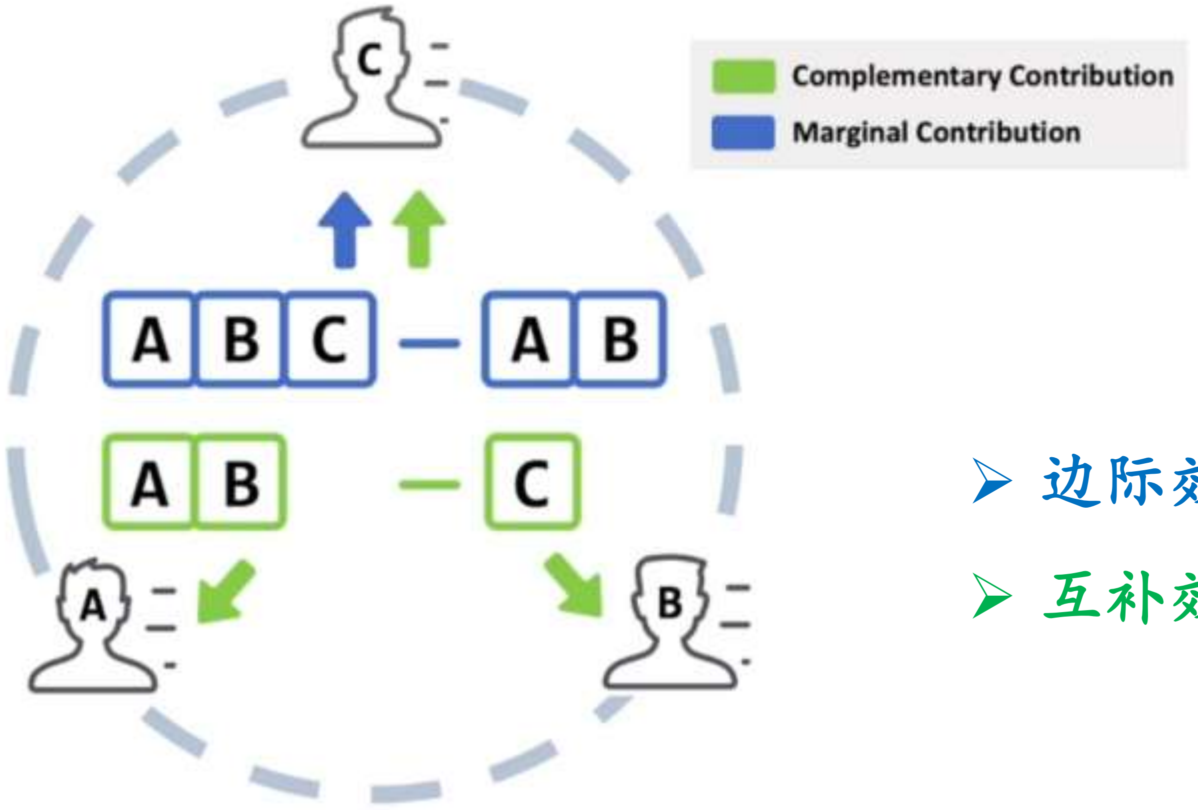
充分利用计算的效应函数值

- 本质是采样边际效应$\boldsymbol{\mathcal{U}(S\cup\{z_i\})-\mathcal{U}(S)}$
- 计算的边际效应越多，结果越精确
- $\forall S\subseteq N,\mathcal{U}(S)$在Shapley Value的计算中被用到的次数：
  - $\forall z\in N\setminus S,\mathcal{U}(S\cup\{z\})-\mathcal{U}(S)$
  - $\forall z\in S,\mathcal{U}(S)-\mathcal{U}(S\setminus\{z\})$
  - 共n个边际效应需要用到$\mathcal{U}(S)$
- 理想情况：算法中任意计算的$\mathcal{U}(S)$都被**充分利用**（n次）

充分利用计算的效应函数值

- 实际情况：**不可能**全部充分利用
- 假设计算了$\mathcal{U}(S)$
  - 若要充分利用$\mathcal{U}(S)$，需要对$\forall \boldsymbol{z} \in \boldsymbol{N} \setminus \boldsymbol{S}$计算$\mathcal{U}(S \cup \{z\})$
  - 若要充分利用$\mathcal{U}(S \cup \{z\})$，需要对$\forall \boldsymbol{z}' \in \boldsymbol{N} \setminus (\boldsymbol{S} \cup \{\boldsymbol{z}\})$计算$\mathcal{U}(S \cup \{z, z'\})$
  - …
  - 需要对所有$\boldsymbol{S} \subseteq \boldsymbol{S}' \subseteq \boldsymbol{N}$计算$\mathcal{U}(S')$
  - 同理，需要对所有$\boldsymbol{S}' \subseteq \boldsymbol{S}$计算$\mathcal{U}(S')$
  - 若要保证计算的效应函数全部充分利用需要计算N的全部2^n个子集

边际效应 vs. 互补效应



- 边际效应：$MC_i(S) = \mathcal{U}(S \cup \{z_i\}) - \mathcal{U}(S)$
- 互补效应：$CC_N(S) = \mathcal{U}(S) - \mathcal{U}(N \setminus S)$

定理

$$SV_i = \frac{1}{n} \sum_{S \subseteq N \setminus \{z_i\}} \frac{MC_i(S)}{\binom{n-1}{|S|}} = \frac{1}{n} \sum_{S \subseteq N \setminus \{z_i\}} \frac{CC_N(S \cup \{z_i\})}{\binom{n-1}{|S|}}$$

边际效应 vs. 互补效应

定理 $$SV_i = \frac{1}{n} \sum_{S \subseteq N \setminus \{z_i\}} \frac{CC_N(S \cup \{z_i\})}{\binom{n-1}{|S|}}$$

- 采样互补效应$\boldsymbol{\mathcal{U}(S \cup \{z_i\}) - \mathcal{U}(N \setminus (S \cup \{z_i\}))}$
- $\forall S \subseteq N, \mathcal{U}(S)$在Shapley Value的计算中被用到的次数：
  - $\forall z \in S, \mathcal{U}(S) - \mathcal{U}(N \setminus S)$
  - $\forall z \in N \setminus S, \mathcal{U}(N \setminus S) - \mathcal{U}(S)$
  - 有且仅有n个互补效应需要用到$\mathcal{U}(S), \mathcal{U}(N \setminus S)$
- 算法中任意计算的$\mathcal{U}(S)$都会被**充分利用**

边际效应 vs. 互补效应

定理 $$SV_i = \frac{1}{n}\sum_{S\subseteq N\setminus\{z_i\}} \frac{CC_N(S\cup\{z_i\})}{\binom{n-1}{|S|}}$$

- 采样互补效应$\boldsymbol{\mathcal{U}(S\cup\{z_i\}) - \mathcal{U}(N\setminus(S\cup\{z_i\}))}$
- $\forall S\subseteq N, \mathcal{U}(S)$在Shapley Value的计算中被用到的次数：
    - $\forall z\in S, \mathcal{U}(S) - \mathcal{U}(N\setminus S)$
    - $\forall z\in N\setminus S, \mathcal{U}(N\setminus S) - \mathcal{U}(S)$
    - 有且仅有n个互补效应需要用到$\mathcal{U}(S), \mathcal{U}(N\setminus S)$
- 算法中任意计算的$\mathcal{U}(S)$都会被**充分利用**

基于互补效应的采样算法

An unbiased estimation of Shapley value

① Randomly generate a pair of complementary coalitions S, $N \setminus S$

② Calculate $u = U(S) - U(N \setminus S)$

③ Assign u to n players in N.

- Assign u to $|S|$-th stratum for $z_i \in S$
- Assign $-u$ to $|N \setminus S|$-th stratum for $z_i \in N \setminus S$

④ Draw m samples and average over complementary contribution means as estimated Shapley value.

Example：

- $N = \{z_1, z_2, z_3\}$
- Draw 4 samples
  - $a = U(\{z_1, z_2, z_3\}) - U(\emptyset)$
  - $b = U(\{z_1\}) - U(\{z_2, z_3\})$
  - $c = U(\{z_1, z_2\}) - U(\{z_3\})$
  - $d = U(\{z_1, z_3\}) - U(\{z_2\})$
- $\overline{SV}_1 = \frac{1}{3}\left(b + \frac{c+d}{2} + a\right)$
- $\overline{SV}_2 = \frac{1}{3}\left(-d + \frac{-b+c}{2} + a\right)$
- $\overline{SV}_3 = \frac{1}{3}(-c + \frac{-b+d}{2} + a)$

层次采样方法

$$SV_i = \frac{1}{n} \sum_{S \subseteq N \setminus \{z_i\}} \frac{CC_N(S \cup \{z_i\})}{\binom{n-1}{|S|}}$$

$$= \frac{1}{n} \cdot \sum_{j=0}^{n-1} \frac{1}{\binom{n-1}{j}} \sum_{\substack{S \subseteq N \setminus \{z_i\} \\ |S|=j}} CC_N(S \cup \{z_i\})$$

$$= \frac{1}{n} \cdot \sum_{j=0}^{n-1} SV_{i,j}$$

$$= \mathrm{AVG}(\mathrm{AVG}(CC_N^{i,j}))$$

层次采样算法

- For each $0 \le j \le n$:
  - Repeat $\boldsymbol{m_j}$ times:
    - 采样大小为j的集合$S \subseteq N$
    - For each $z_i \in S$：更新$\overline{SV_{i,j}}$
    - For each $z_i \in N \setminus S$：更新$\overline{SV_{i,n-j}}$
- For each $z_i \in N$: $\overline{SV_i} = \frac{1}{n} \cdot \sum_{j=0}^{n-1} \overline{SV_{i,j}}$

采样总数为$\boldsymbol{m}$，如何合理分配各层次采样数$\boldsymbol{m_j}$以获得**最小估计方差**？

最小化方差方法

$$SV_i = \mathrm{AVG}(\mathrm{AVG}(CC_N^{i,j}))$$

设$var\left[CC_N^{i,j}\right] = \sigma_{i,j}^2$，则有（用期望$E[m_{i,j}]$代替采样数$m_{i,j}$）：

$$\sum_{i=1}^{n} var[\overline{SV_i}] = \frac{1}{n} \sum_{j=\lceil n/2 \rceil}^{n} \frac{\sum_{i=1}^{n}(\frac{\sigma_{i,j}^2}{j} + \frac{\sigma_{i,n-j}^2}{n-j})}{m_j}$$

令方差和最小化，即求解

$$\min \sum_{j=\lceil n/2 \rceil}^{n} \frac{\sum_{i=1}^{n}(\frac{\sigma_{i,j}^2}{j} + \frac{\sigma_{i,n-j}^2}{n-j})}{m_j},$$

$$\text{s.t.} \sum_{j=\lceil n/2 \rceil}^{n} m_j = m.$$

Jerzy Neyman. 1992. On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. In Breakthroughs in statistics. Springer, 123–150.

最小化方差方法

$$\min \sum_{j=\lceil n/2 \rceil}^{n} \frac{\sum_{i=1}^{n}\left(\frac{\sigma_{i,j}^2}{j} + \frac{\sigma_{i,n-j}^2}{n-j}\right)}{m_j},$$

$$\text{s.t.} \sum_{j=\lceil n/2 \rceil}^{n} m_j = m.$$

使用拉格朗日乘数法解得：

$$m_j = \frac{m\sqrt{\sum_{i=1}^{n}\left(\frac{\sigma_{i,j}^2}{j} + \frac{\sigma_{i,n-j}^2}{n-j}\right)}}{\sum_{j=\lceil n/2 \rceil}^{n}\sqrt{\sum_{i=1}^{n}\left(\frac{\sigma_{i,j}^2}{j} + \frac{\sigma_{i,n-j}^2}{n-j}\right)}}.$$

最小方差采样数分配算法

- 第一阶段：先进行小范围采样，估计$\sigma_{i,j}^2$
- 第二阶段：计算**最优的m_j**并进行后续采样

基于误差界的方法

THEOREM (***Empirical Bernstein-Serfling Inequality***) *Given a set of players* $\mathcal{N} = \{z_i, \ldots, z_n\}$, *the range* r *of the utility function, a sample without replacement of* $CC_{\mathcal{N}}^{i,j}$ *of size* $m_{i,j}$ $\{CC_{\mathcal{N}}(\mathcal{S}_1), \ldots, CC_{\mathcal{N}}(\mathcal{S}_{m_{i,j}})\}$, *where* $\mathcal{S}_1, \ldots, \mathcal{S}_{m_{i,j}} \in \mathfrak{S}^{i,j}$ $(1 \le i, j \le n; 1 < m_{i,j} \le \binom{n-1}{j-1})$, *with probability at least* $1 - \delta$ $(\delta > 0)$ *we have*

$$|\frac{1}{m_{i,j}} \sum_{k=1}^{m_{i,j}} CC_{\mathcal{N}}(\mathcal{S}_k) - \mathcal{SV}_{i,j}| = |\overline{\mathcal{SV}_{i,j}} - \mathcal{SV}_{i,j}|$$

$$\le \frac{1}{m_{i,j}} [\sqrt{2(m_{i,j} - 1)\widehat{\sigma^2_{i,j}} \rho_{m_{i,j}} \log(10/(1+\delta))} + \kappa r \log(10/(1+\delta))],$$

where $\kappa = \frac{14}{3} + 3\sqrt{2}$, $\widehat{\sigma^2_{i,j}} = \frac{1}{m_{i,j}-1} \sum_{k=1}^{m_{i,j}} (CC_{\mathcal{N}}(\mathcal{S}_k) - \mathcal{SV}_{i,j})^2$, *and*

$$\rho_{m_{i,j}} = \begin{cases} 1 - \frac{m_{i,j}-1}{\binom{n-1}{j-1}} & \text{if } m_{i,j} \le \binom{n-1}{j-1}/2 \\ \left(1 - \frac{m_{i,j}}{\binom{n-1}{j-1}}\right)(1 + 1/m_{i,j}) & \text{if } m_{i,j} > \binom{n-1}{j-1}/2 \end{cases}.$$

THEOREM *Given a set of players* $\mathcal{N} = \{z_1, \ldots, z_n\}$, *if for* $1 \le i, j \le n$, $|\overline{\mathcal{SV}_{i,j}} - \mathcal{SV}_{i,j}| \le \epsilon_{i,j}$ *holds with probability at least* $1 - \delta$ $(0 < \delta \le 1, \epsilon_{i,j} > 0)$, *then* $\sum_{i=1}^{n} |\overline{\mathcal{SV}_i} - \mathcal{SV}_i| \le \epsilon$ *holds with probability at least* $(1-\delta)^{2n}$, *where* $\epsilon = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \epsilon_{i,j}$.

Rémi Bardenet and Odalric-Ambrym Maillard. 2015. Concentration inequalities for sampling without replacement. Bernoulli 21, 3 (2015), 1361–1385.

基于误差界的方法

采样集合S时，会更新$\{\overline{SV_{i,|S|}}|z_i \in S\} \cup \{\overline{SV_{i,|N\setminus S|}}|z_i \in N \setminus S\}$

直观上，最好选择误差最大的估计值来更新，即采样：

$$\arg\max_{CC_{\mathcal{N}}(\mathcal{S})} \{\sum_{z_i \in \mathcal{S}} \epsilon_{i,|\mathcal{S}|} + \sum_{z_i \in \mathcal{N}\setminus\mathcal{S}} \epsilon_{i,|\mathcal{N}\setminus\mathcal{S}|}\}.$$

令$\Delta_{i,n-j} = \epsilon_{i,n-j} - \epsilon_{i,j}$，则有：

$$\begin{aligned}&\arg\max_{CC_{\mathcal{N}}(\mathcal{S})} \{\sum_{z_i \in \mathcal{S}} \epsilon_{i,|\mathcal{S}|} + \sum_{z_i \in \mathcal{N}\setminus\mathcal{S}} (\epsilon_{i,|\mathcal{S}|} + \Delta_{i,|\mathcal{N}\setminus\mathcal{S}|})\}\\ =&\arg\max_{CC_{\mathcal{N}}(\mathcal{S})} \{\sum_{z_i \in \mathcal{N}} \epsilon_{i,|\mathcal{S}|} + \sum_{z_i \in \mathcal{N}\setminus\mathcal{S}} \Delta_{i,|\mathcal{N}\setminus\mathcal{S}|}\}.\end{aligned}$$

则问题转变为计算：

$$\arg\max_{CC_{\mathcal{N}}(\mathcal{S})} \sum_{z_i \in \mathcal{N}\setminus\mathcal{S}} \Delta_{i,n-j}, where\ |\mathcal{S}| = j.$$

→ 可以贪心求解

基于误差界的方法

- $N = \{z_1, z_2, z_3, z_4, z_5\}$

- Current bound matrix about $\epsilon_{i,j}$:

$$\mathbf{M} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 2 & 5 & 1 \\ 4 & 1 & 5 & 3 & 2 \\ 5 & 2 & 4 & 1 & 3 \end{bmatrix} \qquad \Delta = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 3 & 4 \\ 4 & 1 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 4 & 3 \\ 4 & 3 \\ 5 & 2 \\ 3 & 5 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} -3 & -1 \\ -3 & -1 \\ -2 & 2 \\ 1 & -4 \\ 4 & -2 \end{bmatrix}$$

$$\mathbf{M} = \begin{bmatrix} 4 & 3 & 3 & 4 & 5 \\ 4 & 3 & 3 & 4 & 5 \\ 5 & 2 & 2 & 5 & 1 \\ 3 & 5 & 5 & 3 & 2 \\ 1 & 4 & 4 & 1 & 3 \end{bmatrix} + \begin{bmatrix} -3 & -1 & 0 & 0 & 0 \\ -3 & -1 & 0 & 0 & 0 \\ -2 & 2 & 0 & 0 & 0 \\ 1 & -4 & 0 & 0 & 0 \\ 4 & -2 & 0 & 0 & 0 \end{bmatrix}$$

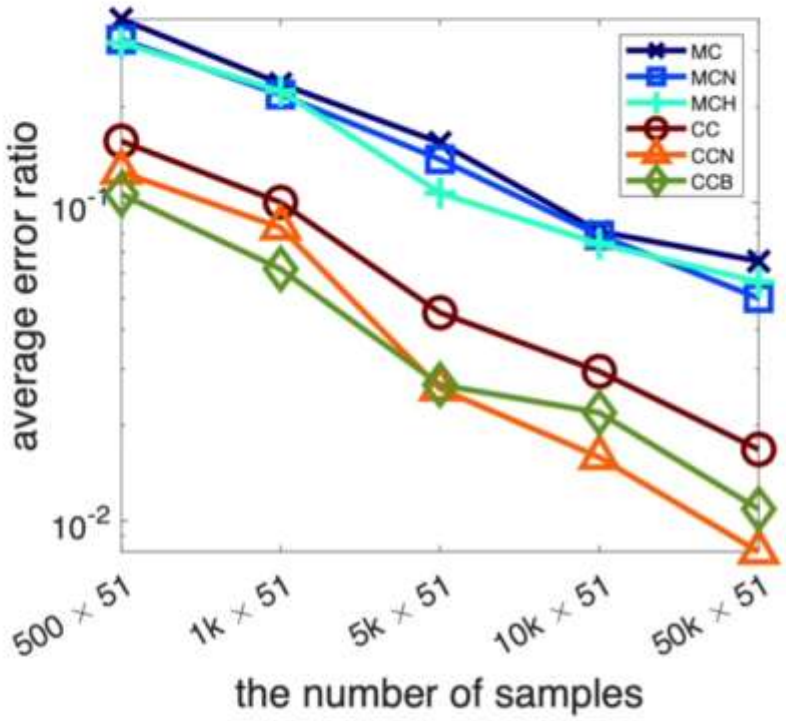
①k=3, top 2 $\Delta_{i,2}$ is -1. Select $S = \{z_1, z_3\}$ and $\epsilon_{\{\{z_1,z_3\},\{z_2,z_4,z_5\}\}} = 18$.

②k=4, top 1 $\Delta_{i,1}$ is 4. Select $S = \{z_5\}$ and $\epsilon_{\{\{z_5\},\{z_1,z_2,z_3,z_4\}\}} = 21$

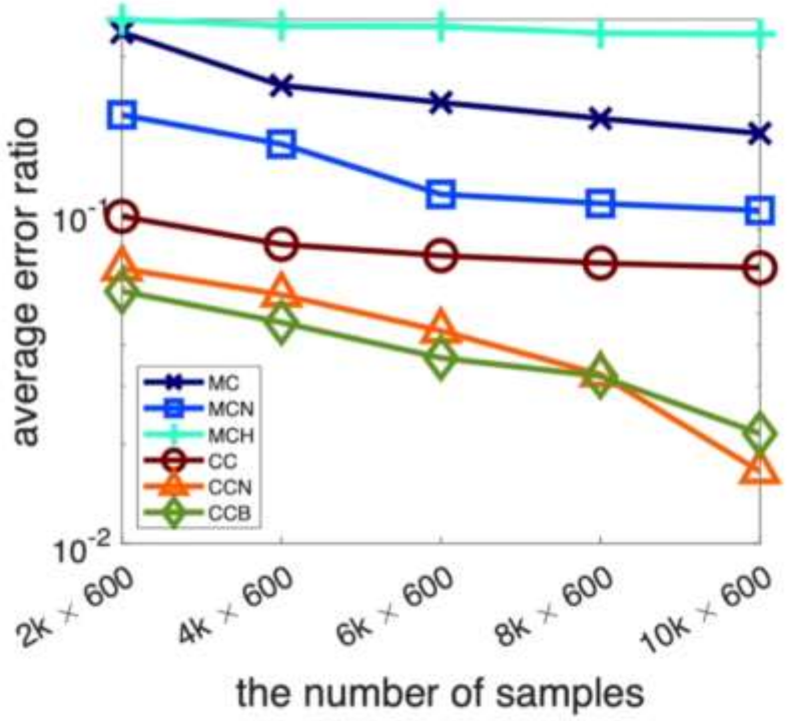
③k=5, $\epsilon_{\{\{z_1,z_4,z_2,z_3,z_4,z_5\},\emptyset\}} = 16$

④Return $CC_N(\{z_5\})$

实验结果：采样数-误差率曲线



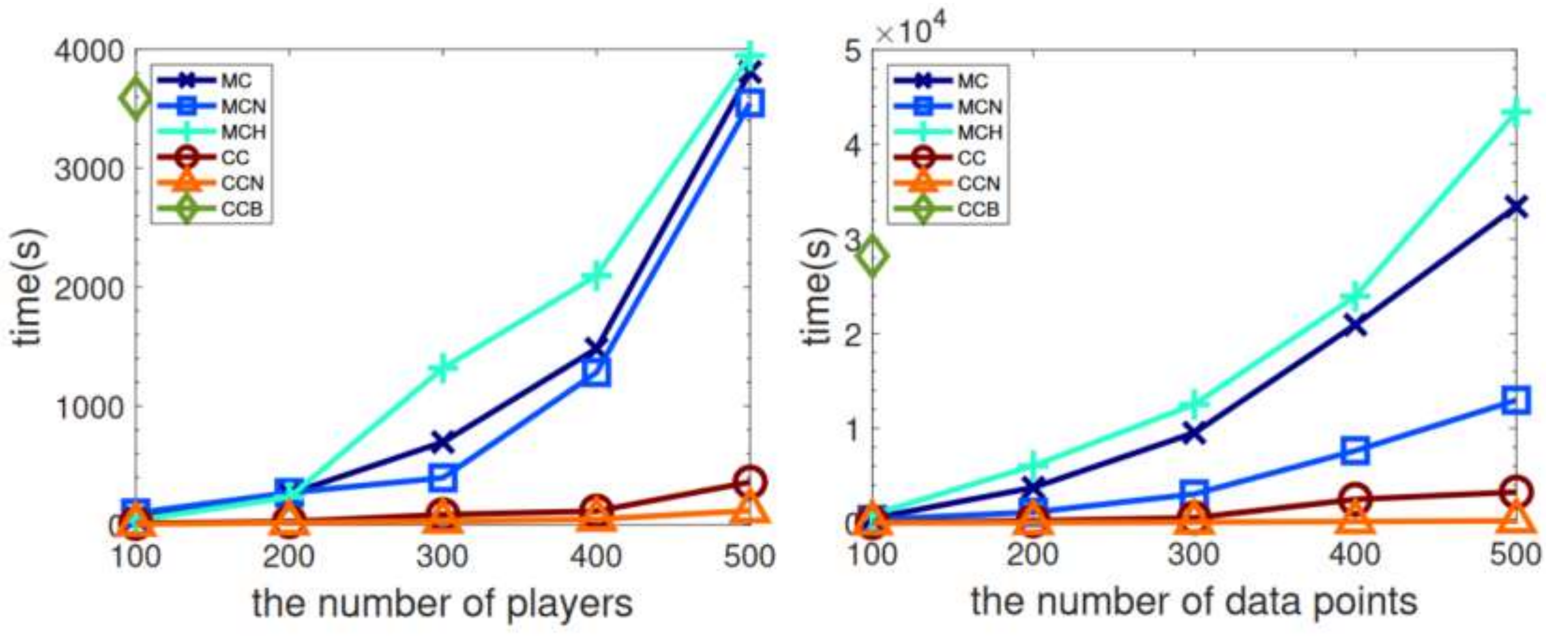
The voting game



The data valuation task

- 基于**边界效益**的方法
  - **MC**：基于随机排列的经典蒙特卡洛采样
  - **MCN**：基于最优分配的层次采样
  - **MCH**：基于Hoeffding界的层次采样
- 基于**互补效益**的方法
  - **CC**：普通随机子集采样
  - **CCN**：基于最小方差的层次采样
  - **CCB**：基于误差界的层次采样

实验结果：误差率降至10%所需的时间



The voting game

The data valuation task

- 基于边界效益的方法
  - MC：基于随机排列的经典蒙特卡洛采样
  - MCN：基于最优分配的层次采样
  - MCH：基于Hoeffding界的层次采样
- 基于互补效益的方法
  - CC：普通随机子集采样
  - CCN：基于最小方差的层次采样
  - CCB：基于误差界的层次采样