セキュアマルチパーティ計算による PS 学習法について
Profit Sharing method for Secure Multi-party Calculation

牧野 隼輝 * 宮島 廣美 * 重井 徳貴 * 宮島 洋文 **
(*鹿児島大学大学院理工学研究科) (**長崎大学大学院医歯薬学総合研究科)

1 はじめに

データマイニングに関して, データを秘匿したまま処理を実現する研究が行われている. 特に個々の秘匿データを分散して記憶や計算処理を実現する研究が注目されている [1, 2]. 本稿では, 個々の秘匿データを分散した状態で計算処理や記憶を行う SMC に着目し, 強化学習に用いる学習データを別々のサーバに分割送信して計算処理や記憶を行う手法を提案し, その有効性を示す.

2 セキュアマルチパーティ計算 (SMC)

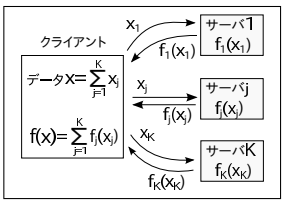

図 1: SMC の概念図

データを秘匿したまま処理を行う方法として, 本稿で用いる Secure Multi-party Calculation(SMC) について説明する. ここでは, クライアントと K 個のサーバからなるモデルによる, 実数データ x に対する関数 $f(x)$ の計算を用いて SMC の計算処理を説明する [2]. i を正の整数として $Z_i = \{1, \cdots, i\}$ とする. はじめに, クライアント側で実数データ x をランダムに K 個のデータ $x_j (j \in Z_K)$ に分割する. j 番目のデータ x_j は j 番目のサーバに送信される. サーバでは送られたデータを用いて $f_j(x_j)$ の処理が行われる. 最終的な結果はクライアント側で $f_j(x_j)$ を加算することで得られ, $\sum_{j=1}^{K} f_j(x_j)$ と表される. 問題は $f(x)$ と $\sum_{j=1}^{K} f_j(x_j)$ がどのようにデータ分割と計算を行えば一致するかということである.

3 Profit Sharing

Profit Sharing(PS) は, 強化学習の代表的手法の 1 つである. S を状態の集合, A を行動の集合とすると, 状態 $\mathbf{s} \in S$ と行動 $a \in A$ に対応する重み $\omega(\mathbf{s}, a)$ を学習により決定する. 学習時の状態 $\mathbf{s}_i$ において行動 a_i をとって状態遷移して目的を達成した場合, r を報酬, γ を減衰率, エピソード終了時点での行動回数を n とすると, $i \in Z_n$ に対する重み $\omega(\mathbf{s}_i, a_i)$ は次式にて更新される [3].

$$\omega(\mathbf{s}_i, a_i) \leftarrow \omega(\mathbf{s}_i, a_i) + r \times \gamma^{n-i} \quad (1)$$

4 SMC による入力データの分割を用いた強化学習

強化学習によって得られる重みは K 個のサーバに分割して保存され, サーバ毎に分割された重み $\omega_j(\mathbf{s}, a)$ の総和を取ることで元のデータとして扱うことができ, $\omega(\mathbf{s}_i, a_i) = \sum_{j=1}^{K} \omega_j(\mathbf{s}_i, a_i)$ のように表される. これにより, いずれかのサーバからデータが漏洩した場合でも全体の機密性を保つことができる. 本稿では, 学習データの集合を分割する手法と, 更新に用いる学習パラメータを分割する手法の 2 つについて考える.

4.1 経路の送信先変更によるデータ分割手法

学習によって得られるデータ集合を部分集合に分けて各サーバへ送信する. 各サーバでの学習では, 受け取ったデータ集合を用いて重みを式 (2) によって更新することで学習を行う.

$$\omega_j(\mathbf{s}_i, a_i) \leftarrow \omega_j(\mathbf{s}_i, a_i) + r \times \gamma^{n-i} \quad (2)$$

4.2 正規分布を用いた報酬の分割によるデータ分割手法

学習の際, 重みの更新量を決定するパラメータである報酬の値を分割する. 強化学習においては, データの数値自体だけでなく大小関係などデータの傾向も重要な情報であるため, データ分割の割合決定に正規分布を用いることで各サーバ別に偏った学習を行わせる. 各サーバでの学習では, 報酬分割の割合 $\beta_j(\mathbf{s}_i, a_i)$ と受け取ったデータ集合を用いて重みを式 (3) によって更新することで学習を行う.

$$\omega_j(\mathbf{s}_i, a_i) \leftarrow \omega_j(\mathbf{s}_i, a_i) + \beta_j(\mathbf{s}_i, a_i) \times r \times \gamma^{n-i} \quad (3)$$

5 シミュレーション

本稿では数値実験として学習者がスタート地点からゴールまでの最短経路を学習する迷路問題を用いる. 結果は従来手法, 経路の送信先を変更する手法, 正規分布によって報酬を分割する手法の 3 つについて比較検討する. 実験環境として図 2 のベンチマーク問題として用いられる sutton の迷路を設定した. 状態数 $|S| = 54$, 行動数 $|A| = 4$, サーバ数 $K = 3$, 最大学習回数 10000 回とし, 20 回試行の平均値を学習結果とする.

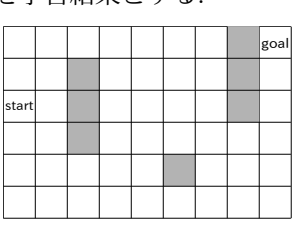

図 2: sutton の迷路

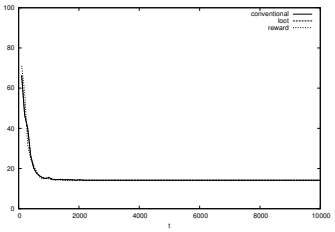

図 3: 移動回数の推移

表 1: 学習後のテスト結果

	従来法		集合分割		報酬分割	
	成功	移動回数	成功	移動回数	成功	移動回数
統合	20	406.0	20	406.0	20	404.1
サーバ 1	-	-	20	2623.1	0	-
サーバ 2	-	-	20	26052.2	0	-
サーバ 3	-	-	20	731.4	0	-

図 3 は学習回数に対する目的達成までの移動回数を示しており, 提案手法は従来手法とほぼ同等の性能を保てていることがわかる. また, 表 1 は学習後の結果を固定し, 図 2 の全ての地点を始点としたとき, 全始点で目的を達成できた試行数, 及びそのときの移動回数の総和を示している. 提案手法については, 各サーバのみの結果を用いた評価を同様に示した. これにより提案手法 (特に報酬分割手法) は, データの秘匿性を保持しつつ, 従来法と同程度の結果を実現している.

6 まとめ

本稿では, 強化学習の 1 つである PS 学習について 2 つのデータ分割手法を用いて SMC を実現し, その有効性を示した. 今後の課題として, 他のデータ分割手法の提案などによる更なる機密性の向上を行いたい.

参考文献

[1] 宮西 他, 電子情報通信学会技術研究報告, Vol.114, No.49, P19-24 (2014)
[2] H.Miyajima et al., Proc. of the IMECS, Vol 1, pp.381-386, 2016.
[3] 木村 他, 計測と制御, Vol.38, No.10, pp.618-623, 計測自動制御学会 (1999)