クラウド環境におけるゲノム秘匿検索に向けた システムデザイン及び暗号スキームの比較と分析

理学専攻・情報科学コース　山田　優輝　(学籍番号：1840674)

1　はじめに

近年ゲノムデータの統計解析が可能になり，医療分野に留まらず様々な分野でゲノムデータ利用の実用化が期待されている．ゲノムデータを統計処理する際には大型のストレージと計算機が必要になるため，クラウドを用いたゲノムデータ委託システムが普及していくと考えられる．この際プライバシ保護が必須となるが，暗号化されたデータ同士での演算が可能な完全準同型暗号を用い，クラウドに秘密鍵を渡すこと無く秘匿演算を行う秘匿検索手法が近年盛んに研究されている．完全準同型暗号を用いたゲノム秘匿検索システムを実現する際には，システムデザインやアルゴリズムだけでなくデータ構造や暗号スキーム，また採用する暗号ライブラリなど様々な要素が影響し合うため，これらを総合的に分析・評価する必要がある．本研究では，クラウド上で行われる完全準同型暗号を用いたゲノム秘匿検索演算の性能を左右する要素について，主にシステムデザインと暗号スキーム・暗号ライブラリに着目して分析する．

2　完全準同型暗号

完全準同型暗号 FHE (Fully Homomorphic Encryption) は加法と乗法の両方において準同型性を有する暗号化手法である．FHE は公開鍵暗号方式の機能を持つが，秘密鍵を用いることなく暗号文同士の演算を行い，平文同士の演算を暗号化した値を導くことが出来るため，ユーザは平文上で行うのと同様に暗号文同士での加法演算・乗法演算を行うことが出来る．課題としては計算量が大きいことの他に，暗号文に含まれるノイズが演算の度に増加し，閾値を越えると正しく復号することが出来なくなる，というものが挙げられる．演算の回数を限定するか，bootstrap と呼ばれるノイズをリセットする手法を導入することで復号を保証することが出来る．しかし bootstrap は計算量が非常に大きいなど難点は残る．

3　システムデザイン

石巻らによる先行研究 [1] 及び [2] で提案されたシステムデザインをそれぞれデザイン 1，デザイン 2 とする．デザイン 1 では一文字分の問い合わせを繰り返すことで最終的な結果を得る．これにより一度あたりのサーバ上での FHE の計算量を削減することが出来るが，クエリ長が増大するとともにクライアントでの計算負荷も増加してしまうという問題もある．また，毎回の通信で大容量のデータが転送されるため，通信ネットワークに負荷をかけ，計算資源の乏しいクライアントには適さないと考えることが出来る．これに対してデザイン 2 では，クエリの文字列長に関わらず一往復の通信で検索を行うことが出来る．これはデザイン 1 よりも計算資源の乏しいクライアントには適しているが，暗号文の容量やサーバ上での毎回の FHE 演算の計算量はデザイン 1 と比べて増大してしまう．

デザイン 2 を可能にするためには，bootstrap の導入もしくはより多くの演算を可能にするための大きなパラメータの指定が必要となる．Bootstrap を用いる場合，暗号文のノイズをリセットすることが出来る一方で，bootstrap 自体が計算量の大きな演算であるため，大きなオーバーヘッドが発生し，計算コストが増大するという問題がある．また，大きなパラメータを指定した場合，bootstrap によるオーバーヘッドは発生しないものの，毎回の FHE 演算の計算コストが増大してしまうという問題点がある．また，クエリ長によって必要な回路の深さが変わるため，それに応じてパラメータも変更する必要がある．

4　実験及び分析

システムデザインと暗号スキーム・暗号ライブラリの二点に着目し，それぞれについて比較実験を行う．システムデザインについては，デザイン 2 に bootstrap を導入したデザインをデザイン 2-1，大きなパラメータを指定するデザインをデザイン 2-2 とし，これにデザイン 1 を加えた三種類のデザインについて比較を行う．暗号ライブラリについては HElib[3] が提供する BGV[4] と PAILSADE[5] が提供する BFV[6] とを比較する．

実験に用いた環境を表 1 に，パラメータを表 2 に示す．

表 1　実験環境

Server	OS	CentOS 6.10
	CPU	Intel®Xeon®Processor E5-2643 v3 (3.4GHz) 6 Cores × 2 Sockets
	Main Memory	512GB
	SSD	80GB
	HDD	2TB

表 2　パラメータ

	デザイン 1	デザイン 2-1	デザイン 2-2
BGV (HElib)	*level = 9*	*level = 22*	*level = 9* * クエリ長
BFV (PALISADE)	*NumMults = 15*	N/A	*NumMults = min(50, 9* * クエリ長*)*

はじめにシステムデザインによる比較実験を行う．クエリ長によるサーバ上での実行時間の比較結果を図 1 に，ポジション数によるサーバ上での実行時間の比較結果を図 2 に示す．

次に暗号スキーム・暗号ライブラリによる比較実験を行う．クエリ長によるサーバ上での実行時間の比較結果を図 3 に，ポジション数によるサーバ上での実行時間の比較結果を図 4 に示す．

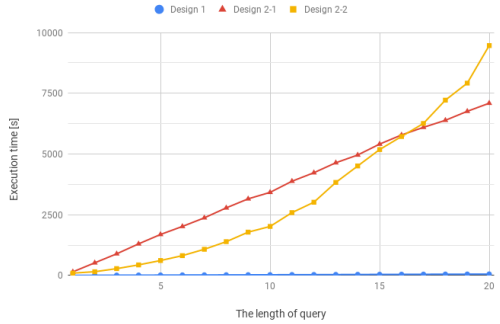

図 1　クエリ長によるシステムデザインごとのサーバ上での実行時間の比較

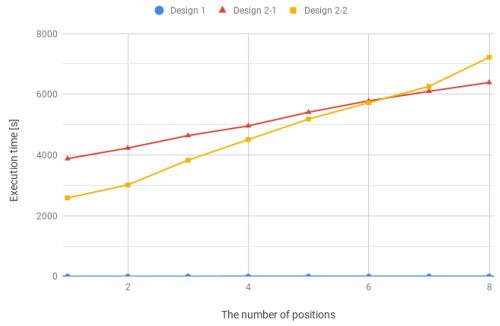

図 2　ポジション数によるシステムデザインごとのサーバ上での実行時間の比較

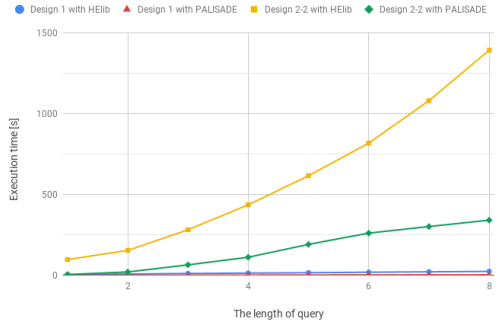

図 3　クエリ長による暗号ライブラリごとのサーバ上での実行時間の比較

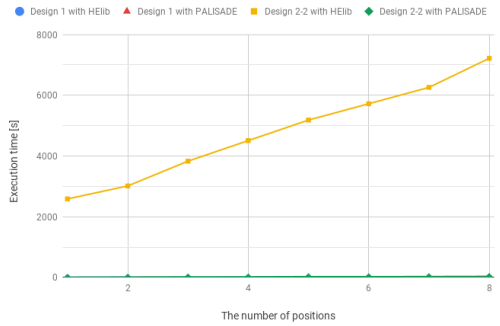

図 4　ポジション数による暗号ライブラリごとのサーバ上での実行時間の比較

クエリ長によって適するデザインが変わること，また本システムについては BGV よりも BFV の方が適していることが読み取れる．

5　結論

先行研究に基づき，完全準同型暗号を用いたゲノム秘匿検索システムを三種類のデザイン及び二種類の暗合スキーム・ライブラリを用いて実装し，クラウドコンピューティングを想定した環境下で実験を行った．また得られた実験結果について，システムデザインと暗号スキーム・ライブラリの観点から分析を行った．その結果，クエリ長やポジション数などによって適するデザインが変わること，本アプリケーションでは PALISADE により提供される BFV スキームが良い性能を示すことが確認された．

本研究に対する発展課題としては，デザイン 1 とデザイン 2 を組み合わせたシステムデザインを提案すること，分散環境下での比較を行うこと，クライアントとして実際に使用するであろう端末を用意し通信帯域を絞って通信時間の計測を行うことなどが挙げられる．デザイン 1 とデザイン 2 を組み合わせたシステムデザインについては一定程度のクエリ長までは一度の問い合わせで検索を終えるが，クエリ長が長くなった場合は問い合わせを分割することで，サーバでの計算量を削減しつつクライアントへの負荷も軽く出来ると考えられる．また，分散環境下での比較比較実験を行った場合，本研究で採用している計算手順ではデザイン 2-1 では bootstrap を分散できず，デザイン 2-2 の方が高パフォーマンスとなることが予想される．

謝辞

本研究は JST CREST JPMJCR1503 の支援を受けております．

参考文献

[1] Y. Ishimaki, K. Shimizu, K. Nuida, and H. Yamana, “Poster: Privacy-preserving string search for genome sequences using fully homomorphic encryption,” in *IEEE Symposium on Security and Privacy*, 2016.

[2] Y. Ishimaki, H. Imabayashi, K. Shimizu, and H. Yamana, “Privacy-preserving string search for genome sequences with fhe bootstrapping optimization,” in *2016 IEEE International Conference on Big Data (Big Data)*, IEEE, 2016, pp. 3989–3991.

[3] homenc, *Helib: An implementation of homomorphic encryption*, `https://github.com/homenc/HElib/`, visited on 12/2019.

[4] Z. Brakerski, C. Gentry, and V. Vaikuntanathan, “Fully homomorphic encryption without bootstrapping,” *IACR Cryptology ePrint Archive*, vol. 2011, p. 277, 2011.

[5] PALISADE, *Palisade homomorphic encryption software library*, `https://palisade-crypto.org/software-library/`, visited on 12/2019.

[6] J. Fan and F. Vercauteren, “Somewhat practical fully homomorphic encryption,” *IACR Cryptology ePrint Archive*, vol. 2012, p. 144, 2012.