

研究速報

ブロックチェーンを用いた信頼ある機密データの
管理及び利活用基盤池川 航史[†] 西島 直[†]Blockchain-Based Trusted Data Management and Utilization Infrastructure
for Sensitive DataKoshi IKEGAWA[†] and Nao NISHIJIMA[†], Nonmembers[†] (株) 日立製作所デジタルプラットフォームイノベーションセンタ, 国分寺市
Center for Technology Innovation – Digital Platform, Hitachi, Ltd., Kokubunji-
shi, 185-8601 Japan
DOI:10.14923/transinfj.2021SGL0001

あらまし 本論文における提案手法は、機密データを外部に出すことなく他の組織に利活用させることを可能とする。また、提案手法は秘匿機能をもつスマートコントラクトを活用し、機密データ利活用時の手続き情報に関しても秘匿化を実現する。

キーワード ブロックチェーン、スマートコントラクト、トラストサービス

1. まえがき

1.1 信頼ある自由なデータ流通 (DFFT)

2019年1月23日に開催された世界経済フォーラムにて、信頼ある自由なデータ流通 (Data Free Flow with Trust, DFFT) [1] の実現が、新しい経済のための最優先事項であると提言された。DFFTの実現に向け、世界経済フォーラム第四次産業革命日本センターは、1) 国境を超えた自由なデータ流通、2) 個人、企業及び都市間の自由なデータ取引市場、及び3) 規制・ルールのアップデートによるトラストの再設計の三つに焦点を当てている。上記3点を考慮してDFFTを実現するにあたり、複数組織が参加するデータ共有ネットワークが必要であり、そのネットワークの管理権限は非中央集権的である必要がある。DFFTを実現するための技術として、上記の要件を満たすブロックチェーン (Blockchain, BC) 技術の活用が期待されている。

1.2 ブロックチェーン技術

BC技術は、破壊的イノベーションとして金融分野や産業分野などへの応用が期待され、注目を集めている。例えば金融分野では、従来は第三者機関を経由して実施されてきた取引をBC技術を用いて利用者間のピアツーピア (peer-to-peer, P2P) の直接取引に代替することで、取引コストの削減が期待できる。BCの特徴は、取引やデータの内容が参加しているBCネットワーク上の全ての組織に共有され、各組織が保有する台帳 (分散台帳) に書き込まれる。データを改ざんす

るためには分散台帳全てを書き換える必要があるため、高い耐改ざん性を有する。

BCを構成するネットワークは大きく2種類に分けられる。一つ目は、暗号資産のように不特定多数の計算機資源が形成する自由参加型 (例: Bitcoin [2] など) である。二つ目は、特定の企業や団体などの組織のみで形成されており、参加には参加組織の許可が必要となる許可型 (例: Hyperledger Fabric [3] など) である。

一方で、医療データや個人情報などの機密データは、個人情報保護法やEU一般データ保護規則 (GDPR) など各国が定める法律に基づいて厳重に扱う必要がある。許可型BCにおいても参加している組織全てにデータが共有されるため、BCに直接機密データを書き込むことは避ける必要がある。

1.3 本研究の貢献

本論文では、許可型BC基盤であるHyperledger Fabricを用いて、信頼性をもって機密データを管理及び利活用するための手法を提案する。提案手法は機密データを利活用したい組織Aがデータを所有する組織Bに対して、データの利活用処理を依頼し、組織Bが処理を実行することによって機密データを外部に出すことなく他の組織に利活用させることを可能とする。また、提案手法は各組織が所有する機密データを利活用する権利の申請や承認等の手続きをChaincode上で管理する際にFabric Private Chaincodeを用いることにより、その手続き自体も秘匿化できる。更に、提案手法を用いた医療分野における機密データ共有を対象に具体的な手法の実適用例を示す。

2. 本研究に関連する技術及び技術課題

本節にて関連研究及び技術として、提案手法にて用いるHyperledger Fabric (以下、Fabric) を紹介する。次に、Fabricにおけるプライバシー保護に関する技術及び研究を紹介する。最後に、Fabric以外のBC技術におけるプライバシー保護に関する研究及び技術を紹介する。

2.1 Hyperledger Fabricの概要

Fabricは、Hyperledger Foundationにより管理されているオープンソースの許可型BC基盤の一つである。図1に、Fabricを用いてBCネットワークを立ち上げた際の概観及び合意形成の流れを示す。

PeerはLedger, State DB, 及びChaincodeの実行環境を含み各組織がそれぞれ一つ以上保有する。Ordererはトランザクションの順番を決め、Peerにブロックを配布する役割を担う。Ledgerはトランザクションを

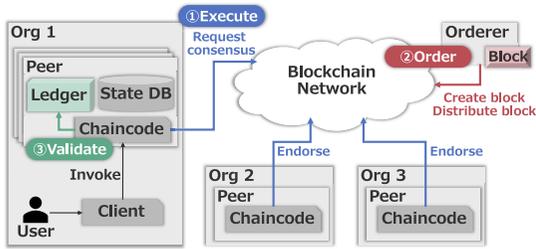


図1 Fabric を用いて BC ネットワークを立ち上げた際の概観及び合意形成の流れ

包したブロックの集合体であり分散台帳である。State DB は Ledger が管理しているトランザクションを実行した際の結果の最新情報を保存するデータベースである。Chaincode は Fabric におけるスマートコントラクトであり、Client から呼び出される。複数組織の Peer 上の Chaincode の実行結果が一致することによって信頼性を確保している。Client はユーザが操作するインターフェースであり、Chaincode の呼び出しを行う。

Fabric では Execute-Order-Validate と呼ばれる 3 段階構成の合意形成アルゴリズムを用いてトランザクション書き込み及び実行の合意を取得している。以下に図 1 を用いて合意形成の流れを示す。1. Execute: ユーザが Client を介して Chaincode を呼び出しトランザクションの実行を要求すると、BC ネットワークに参加する他の組織に署名を要求し、他の組織の Peer が、Chaincode 上でトランザクションの検証及び署名を行う。2. Order: Orderer は署名を集めたトランザクションを受け取り、トランザクション同士の順序関係を解決してブロックとして梱包し、そのブロックを各組織の Peer に渡す。3. Validate: Orderer から渡されたブロックを各 Peer が検証し台帳に書き込む。ブロック内に書き込まれたトランザクションを実行し、State DB に結果を書き込む。

2.2 Fabric におけるプライバシー保護技術

本節にて、Fabric におけるプライバシー保護機能である Channel 機能、Private Data Collection 機能、及び Fabric Private Chaincode 機能を紹介する。

a) Channel

Fabric の BC ネットワーク上でトランザクション及びデータの読み書きを行うために、各組織は Channel に参加する必要がある。Channel とは特定の組織間のプライベートな通信レイヤーであり、その Channel に参加していない BC ネットワーク内の他の組織は台帳

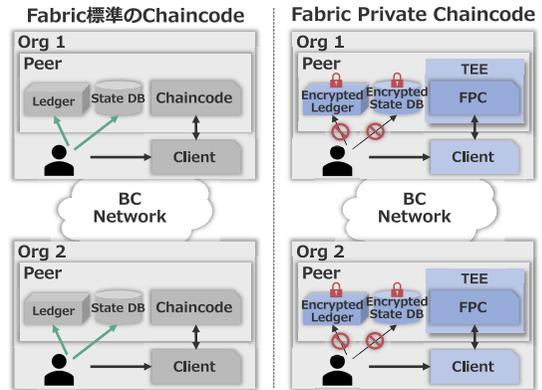


図2 Fabric 標準の Chaincode と FPC の違い

を見ることができない。各 Channel は、Channel に参加している組織のみが読み書きできる独立した分散台帳で構成されている。BC ネットワークに参加する組織数が N の場合、 $N(N-1)$ 個の Channel を作成する必要がある、Channel の管理コストがかかるという課題がある。

b) Private Data Collection (PDC)

PDC 機能は、Benhamouda らの研究 [4] にて提案され Fabric に実装された、他の組織に対して秘匿したい機密データを台帳外の外部記憶装置に保存する機能である。PDC 機能を用いて保管されているデータを他の組織に共有する際は BC ネットワークを介さず P2P の直接通信で送受信を行う。一方で機密データを P2P 通信で他の組織に送信するため、他の組織に送られたデータがどのように活用されるか、提供者は分からないという課題がある。

c) Fabric Private Chaincode (FPC)

FPC [5] は、Trusted Execution Environment (TEE) を用いて Chaincode の実行を可能にする Fabric の拡張技術である。TEE とは、CPU に搭載されたハードウェアセキュリティ機能であり、Intel や AMD、ARM などの CPU ベンダ各社が提供している。FPC の実装に用いられた Intel 社製の CPU に搭載されている TEE である Software Guard Extensions (SGX) [6] は、主記憶装置上に暗号化された保護領域を生成し、その領域にプログラムやデータを読み込むことによって機密データを保護しつつプログラムを実行することを可能にする。

図 2 に Fabric の機能として利用可能な標準の Chaincode と FPC の違いを示す。通常の Fabric の機能として実行される Chaincode は、トランザクション及び

データが平文の状態では保存されており、Chaincode を介さずとも分散台帳や State DB を直接アクセスすることで内容を確認することができる。一方で、FPC を用いた場合は分散台帳及び State DB へ書き込まれたトランザクション及びデータは暗号化されて分散台帳や State DB へ書き込まれる。よって、それらに直接アクセスしても取引やデータを取り出すことはできない。FPC は TEE 上で実行されるため、外部からの攻撃等を受けず安全かつ高信頼にトランザクション及びデータの暗号化・復号化が可能となる。よって、FPC を用いて管理された分散台帳及び State DB の内容を取得するには FPC を介してデータの取得をする必要がある。FPC の関連研究として、Desai らの研究 [7] では、オークションを実現するスマートコントラクトに FPC を用いることにより、入札の情報などを秘匿し談合などの不正な入札防ぎ手段を提案している。

2.3 その他 BC 技術におけるプライバシー保護技術

BC におけるプライバシー保護に関する研究及び技術が数多く存在する。Kosba らの研究 [8] では、BC 上の金融取引を平文で保存しない分散型スマートコントラクトシステムを提案した。また、Hyperledger Avalon は、Enterprise Ethereum Alliance が公開している Trusted Compute Specifications [9] を実装したものであり、TEE を用いて BC (On-chain) の処理をスマートコントラクト上から BC 外の専用の計算機資源 (Off-chain 領域) に安全に移すことを目的としている。Avalon を用いた関連研究として、Ranathunga らは機械学習と BC を活用した高信頼なクロスレイヤの計算モデリングを実現する手法を提示した [10]。更に、LayerX 社が開発した Anonify [11] は LayerX が独自に開発した BC 基盤を用いて Off-chain 領域におけるプログラムを正しく実行できることを保証している。加えて、Cheng らの研究 [12] では、Ethereum と TEE を組み合わせ、プライバシー保護を実現するとともに高速な処理かつ低遅延を実現している。

3. 提案手法

BC を用いて機密データを管理及び利活用しようとする流れがある。医療分野において、複数の組織で機密データであるゲノムデータを管理及び利活用する研究 [13] がある。この研究は機密データを各組織が保有する外部記憶装置に格納し、そのデータのメタデータのみを BC に格納することにより、機密データの複数組織管理を実現している。しかし、この管理及び利活用基盤では、データの利活用に関係しない組織が解

析依頼や結果情報を閲覧することができる。ゲノム情報を扱う研究者にとって、研究対象としているゲノムデータに関する情報は秘匿したい。

本論文では、Fabric を用いて、信頼性をもって機密データを管理、共有、及び利活用するための手法を提案する。提案手法は機密データを利活用したい組織 A がデータを所有する組織 B に対してデータの利活用処理を依頼し、組織 B が処理を実行することによって機密データを外部に出すことなく他の組織に利活用させることを可能にした。また、提案手法は各組織が所有する機密データを利活用する権利の申請や承認等の手続きを Chaincode 上で管理する際に Fabric Private Chaincode を採用し、その手続きの内容自体も秘匿化した。

3.1 提案手法の構成

図 3 に、本論文で提案するデータ管理及び利活用基盤の構成を示す。本基盤は、Fabric と FPC を利用して BC ネットワークを構築した。この BC ネットワークを構成する各組織の Peer には、一つの Fabric 標準の Chaincode と二つの FPC を用いた Private Chaincode がインストールされている。Fabric 標準の Chaincode としては、この BC ネットワークに所属する全ての組織に共有すべきデータのメタデータを管理する (a) データカタログ管理 Chaincode をインストールする。(b) データ権限管理 Chaincode と (c) タスク管理 Chaincode は、データを交換する組織間でのみトランザクション及びデータを共有すべきであるため、FPC を用いた Private Chaincode としてインストールする。

3.2 提案手法の処理の流れ

本基盤を利用しデータの管理及び利活用をする流れを示す。表 1 に State DB へ書き込まれたデータの内の

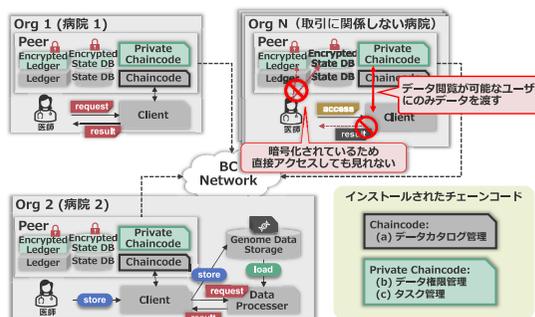


図3 提案手法の概観：Fabric と FPC を用いたデータ管理及び利活用基盤

表1 State DB に書き込まれたデータの内容

(a) データカタログ管理 Chaincode が管理する State DB

Data name (key)	Owner (1)	Hash (2)	Other (3)
Genome Data 001	Org 2	00aa11bb22cc...	Data info

(b) データ権限管理 Chaincode が管理する State DB

Data name (key)	Access Request (1)	Access Approval (2)
Genome Data 001	Org 1 Doctor	Org 1 Doctor

(c) タスク管理 Chaincode が管理する State DB

Data name (key)	Requester (1)	Task (2)	Result (3)
Genome Data 001	Org 1 Doctor	Analyze conditions	Results

容を示す。各 Chaincode が管理する State DB は Data name をキーとしてデータを管理している。医療分野におけるゲノムデータの管理及び利活用を例に処理の流れを紹介する。

はじめに、組織 2 の医師がゲノムデータを外部記憶装置に保存する。同時に、組織 2 の医師はデータカタログ管理 Chaincode を用いてゲノムのメタデータを BC に書き込む (表 1(a))。メタデータとは、データの名前、所有者情報、ゲノムデータのハッシュ値、その他データに関する情報などである。次に、組織 1 の医師はデータカタログ管理 Chaincode を呼び出し、ゲノムのメタデータを参照し、活用したいデータを検索する。組織 1 の医師が活用したいデータを見つけた場合、BC 上のデータ権限管理 Chaincode を用いてデータ所有者にアクセス権を求めるリクエストを書き込む (表 1(b)(1))。組織 2 の医師は、自身もつデータに対するアクセス権の要求を確認し、承認する場合、医師はその情報を BC に書き込む (表 1(b)(2))。組織 1 の医師がアクセス権を得ると、その医師は組織 2 にゲノムデータの解析を依頼する (表 1(c)(2))。組織 1 の医師が書いたリクエストを読み、組織 B が所有する Data Processor が解析処理を開始する。Data Processor は、指定されたゲノムデータを外部記憶装置から読み込む。最後に、Data Processor の処理結果を BC に書き込む (表 1(c)(3))。

4. むすび

本論文では、Hyperledger Fabric を用いて機密データを共有する場合に焦点を当て、高信頼な機密データを管理及び利活用するための技術を提案した。提案手法は Fabric と Fabric を拡張する形で使用可能な秘匿機能を有するスマートコントラクトを組み合わせて、複数組織間での情報管理と機密データの秘匿を両立することを可能にした。提案手法を実現するための構成を示

し、医療分野におけるゲノム共有のユースケースを例示し処理の流れを示した。今後は、本システムの性能に関する評価や実証実験、実運用を進める。

文 献

- [1] World Economic Forum 2019, "Data Free Flow with Trust (DFFT): Paths towards free and trusted data flows," June 2020.
- [2] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," March 2009. <https://bitcoin.org/bitcoin.pdf> (cited 2022-09-09).
- [3] E. Androulaki, A. Barger, V. Bortnikov, et al., "Hyperledger Fabric: A distributed operating system for permissioned blockchains," Proceedings of the Thirteenth EuroSys Conference, pp.1-15, EuroSys '18, Association for Computing Machinery, New York, NY, USA, April 2018. <https://doi.org/10.1145/3190508.3190538>
- [4] F. Benhamouda, S. Halevi, and T. Halevi, "Supporting private data on hyperledger fabric with secure multiparty computation," 2018 IEEE Int. Conf. Cloud Engineering (IC2E), pp.357-363, 2018.
- [5] M. Brandenburger, C. Cachin, R. Kapitza, and A. Sorniotti, "Blockchain and Trusted Computing: Problems, pitfalls, and a solution for hyperledger fabric," 2018.
- [6] Intel, "Intel Software Guard Extensions (SGX)," <https://www.intel.co.jp/content/www/jp/ja/architecture-and-technology/software-guard-extensions.html> (cited 2021-12-24).
- [7] H. Desai and M. Kantarcioglu, "SECAUCTEE: Securing auction smart contracts using trusted execution environments," 2020 IEEE Int. Conf. Blockchain, pp.448-455, Dec. 2021.
- [8] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamathou, "Hawk: The blockchain model of cryptography and privacy-preserving smart contracts," 2016 IEEE Symposium on Security and Privacy (SP), pp.839-858, 2016.
- [9] EEA, "Enterprise ethereum alliance off-chain trusted compute specification v1.1," July 2021. <https://entethalliance.github.io/trusted-computing/spec.html> (cited 2021-12-24).
- [10] T. Ranathunga, A. McGibney, and S. Rea, "The convergence of blockchain and machine learning for decentralized trust management in iot ecosystems," Proc. 19th ACM Conf. Embedded Networked Sensor Systems, pp.499-504, SenSys '21, Association for Computing Machinery, New York, NY, USA, 2021. <https://doi.org/10.1145/3485730.3493375>
- [11] LayerX, "Anonify." <https://layerx.co.jp/labs/product/anonify> (cited 2022-03-14).
- [12] R. Cheng, F. Zhang, J. Kos, W. He, N. Hynes, N. Johnson, A. Juels, A. Miller, and D. Song, "Ekiden: A platform for confidentiality-preserving, trustworthy, and performant smart contracts," 2019 IEEE European Symposium on Security and Privacy (EuroS&P), pp.185-200, 2019.
- [13] K. Ikegawa, N. Nishijima, Y. Ozawa, et al., "Secure and traceable system for genomic data sharing using hyperledger fabric blockchain," 2020 年日本バイオインフォマティクス学会年会 第 9 回 生命医薬情報学連合大会 (IIBMP2020), Sept. 2020.

(2021 年 12 月 24 日受付, 2022 年 4 月 29 日再受付, 6 月 22 日早期公開)