様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　５ 年　 ６ 月 １３ 日現在

機関番号：８２６３６
研究種目：基盤研究(B)（一般）
研究期間：2020～2022
課題番号：２０Ｈ０４１８６

研究課題名（和文）自己検証・自己回復能力を持つ構造化オーバレイネットワーク

研究課題名（英文）Structured Overlay Network with Self-Verification and Self-Healing Features

研究代表者
寺西　裕一（Teranishi, Yuuichi）

国立研究開発法人情報通信研究機構・ネットワーク研究所・研究マネージャー

研究者番号：３０４０３００９

交付決定額（研究期間全体）：（直接経費）　　6,800,000円

研究成果の概要（和文）：ビザンチン障害耐性を備える新たなキー順序保存型構造化オーバレイネットワークの実現方式を提案した。提案方式は，各ノードによるメッセージの署名検証、ホップあたりk個の冗長経路の動的構成により、不正ルーティング攻撃からの「自己回復」を可能とする。提案方式は、キーを持つノードに対し、単一キーを指定してメッセージを送信するユニキャストと、キーの範囲を指定してメッセージを送信するマルチキャストの双方を可能とする。提案方式のプロトタイプソフトウェアを開発し、シミュレーションおよび広域テストベッド上での実機評価により動作を検証した。また、分散ストレージシステムの一つであるIPFSに組み込んだ動作検証も行った。

研究成果の学術的意義や社会的意義
構造化オーバレイネットワークのビザンチン障害耐性の実現方法に関する研究は従来盛んに行われてきたが、既存研究の多くは、DHT（分散ハッシュテーブル）を前提としており、範囲検索を扱えるものはなかった。また、データ配信に適さないイテレーティブルーティングを前提としているため、マルチキャストを効率的に行えないなど応用範囲が限定されていた。提案方式は、ビザンチン障害耐性を持つ構造化オーバレイネットワークとして範囲検索・マルチキャストを実現可能とする初の方式であり、非集中型ストレージやデータ配信のためのネットワーク基盤としての意義は大きいと考える。

研究成果の概要（英文）：We propose a novel Byzantine fault-tolerant key-order-preserving structured overlay network. The proposed overlay network facilitates "self-recovery" from adversarial routing attacks through the verification of message signatures by each individual node, along with the dynamic routing of k-redundant paths for each hop. The proposed overlay network provides both unicast, wherein a message is delivered to a node based on a singular key, and multicast, wherein a message is delivered to nodes encompassing a range of keys. The prototype software of the proposed overlay network has been developed and its functionality has been corroborated through simulation and evaluation conducted on an extensive testbed. Additionally, the proposed overlay network has been integrated into IPFS, a widely utilized distributed storage system.

研究分野：ネットワーク

キーワード：Peer-To-Peer　オーバーレイネットワーク　ビザンチン障害耐性　マルチキャスト

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

１．研究開始当初の背景

Peer-to-Peer(P2P)型のネットワークは，一部のノード（計算機・サーバ）に障害が生じてもネットワークとして動作を継続できる利点があり，サーバ・クライアントアーキテクチャにおける単一障害点の課題を克服し，可用性が高いシステムを構築する基盤となる技術である．ノードの参加・離脱が自由なオープン環境における P2P ネットワークの多くは，複数のランダムに選択されたノードを隣接ノードとし，フラッディング（ブロードキャスト）によってデータ共有やルーティングを行う「非構造化オーバレイネットワーク」である．非構造化オーバレイネットワークは，障害や悪意を持つノードが，あるノード集合をネットワーク構造的にあるいは転送を妨害して隔離してしまう「エクリプス攻撃」に対する耐性が高い．このネットワーク分断耐性の高さから，信頼できないノードが混在するオープン環境が前提のブロックチェーン等の応用では，非構造オーバレイネットワークが広く利用されてきた．しかし，非構造化オーバレイネットワークは，スケーラビリティ（ノード数が増加しても性能を維持できる性質）に乏しく，参加ノード数の増加により遅延増大やルーティング失敗が生じる確率が高くなる課題がある．ブロックチェーン等の応用では，データ更新完了時間や承認時間に影響が及ぶ．

一方，隣接ノードが一定の制約のもと決定される「構造化オーバレイネットワーク」は，こうした非構造化オーバレイネットワークにおけるスケーラビリティの課題を解決可能である．多くの構造化オーバレイネットワークは，任意ノードを根とした木構造に基づいて O(log n)（n はノード数）ホップで宛先へ到達可能であり，単一あるいは複数ノードへのルーティングを低遅延で実行可能である．しかし，構造化オーバレイネットワークは，すべてのノードが決められたアルゴリズムに従った動作を行わなければ，ネットワークを維持できず，エクリプス攻撃を含めアルゴリズムを逸脱した任意の動作が行われ得る障害（ビザンチン障害）に対する耐性（Byzantine Fault Tolerance: BFT）をいかに実現するかが課題であった．

２．研究の目的

構造化オーバレイネットワークのうち，ノードが保持する「キー」が順序関係を持ち，キーの順序関係に基づくネットワーク構造をもつ Key-Order Preserving Structured Overlay Network (KOPSON)は，保持するキーの値が連続した複数ノードへのマルチキャストを木構造で効率的に実行可能である．また，ノード間のルーティングにおいて，それらノード間にないキーを持つノードを経由しない「ルーティングのローカリティ」が実現できる．これらの性質は，分散データベース，分散 pub/sub などの応用に適しており，応用範囲が広い．しかし，従来，BFT を有する KOPSON アルゴリズムは知られておらず，オープン環境に KOPSON を適用する上での障壁となってきた．本研究は，KOPSON の性質を維持しつつ，障害を持つ，あるいは，悪意あるノード（faulty ノードと呼ぶ）が存在する状況のもと，オーバレイネットワーク構造の維持と不正ルーティングの回避を可能とする BFT 対応の KOPSON アルゴリズムを確立することを目的とする．

既存研究として，構造化オーバレイネットワークにおける BFT を扱ったものがいくつか存在する[1, 2]．しかし，いずれも，応用として分散ハッシュテーブル(Distributed Hash Table: DHT)のみを想定しており，キーの値が ID 空間上で均一に分散することを前提としている．既存研究はこの性質を用いて経路表内のキーの値を元にルーティングの動作検証をするが，KOPSON では各ノードが任意の値をもつキーを設定可能であり，悪意ノードが攻撃対象のノードの経路表を汚染する攻撃が起こり得るため，適用できない．さらに，既存研究はルーティングとして基本的にルーティング元がルーティング動作を逐一検証するイテレーティブルーティングを対象としており，マルチキャストやブロードキャストに向かない．本研究では，従来の構造化オーバレイネットワークにおける BFT 実現手法と異なるアプローチにより KOPSON が持つ性質を維持しつつリカーシブルーティングを扱う手法を提案する．

近年，セキュアエンクレーブ等の特別なハードウェアを用いて，指定された動作を遠隔ノードが実行したかどうか検証可能とする方式が提案されており，KOPSON の動作検証にも適用できる可能性がある．しかし，すべてのノードが特別なハードウェアを所有する前提は適用範囲を狭めると考えられることから，本研究では特別なハードウェアやファームウェアの存在を前提とせず，ソフトウェアレベルにて自己完結して動作可能とすることを目指す．

３．研究の方法

本研究では，次の基本方針により KOPSON に自己検証・自己回復能力を実現し，BFT に対応することを考える．

１）キーの承認：各ノードは，承認されたキーでのみ KOPSON に参加できるものとする．キーが承認されたものかどうかは公開鍵暗号の枠組により検証可能とする．

２）k-冗長化：KOPSON の各動作を，k 個のノード集合により冗長化して実行させる(k は定数)．k 個のノードのうち，少なくとも１つ正常なノードが存在すれば，ネットワーク構造の維持やルーティングの継続が可能となるようプロトコルを設計する．承認されていないキーを持つノードが生成したメッセージや，プロトコルに矛盾する内容を持つメッセージは廃棄する．

メッセージのルーティングにおいては，ルーティング元が動作検証できないリカーシブルーティングの BFT を実現するため，複数ノードが独立して動作しても，k 個の冗長化された経路を維持するプロトコルとする．また，k 個の隣接ノード間で経路表上のエントリを共有し，意図

的に値が離れたキーを用いるなどのキー順序の不整合があれば排除する．

本研究では，上記方針に基づき，次の４つの機能要素に分類して検討を進めた．

① ノード認証方式

各ノードを承認する Authority の実現方法を実現する．Authority は各ノードが持つ公開鍵に対する署名も行う．

② オーバレイ構造維持方式

冗長化されたルーティングを行えるオーバレイネットワーク構造を提案する．KOPSON として知られる Skip Graph[3]の構造を適用し，キー順にソートされたリング構造，複数レベルの双方向スキップ構造を，冗長性を持って維持する．

③ ルーティング方式

k 個の冗長化された経路を持つリカーシブルーティングによって BFT を実現する効率的なルーティングプロトコルを提案する．単一のキーを宛先とするユニキャスト，及び，キーの範囲を宛先とするマルチキャストのルーティング方式についてそれぞれ提案する．

④ アプリケーション

本研究で提案する KOPSON をブロックチェーン等との親和性が高く，広く用いられている IPFS(InterPlanetary File System)のルーティングレイヤとして組み込み，適用可能性を示す．

４．研究成果

① ノード認証方式

信頼の起点となる Authority が構造化オーバレイネットワークに参加する各ノードを認証する手順を提案した．全てのノードは Authority へアクセスするためのホスト名や公開鍵をあらかじめ知っているものとする．

各ノードは秘密鍵・公開鍵を持ち，各ノードの公開鍵には Authority が署名を付与しておく．

ノード v は KOPSON に参加する際，Authority にアクセスし，

- $key(v)$
- $TMV(v)$
- $<key(v), TMV(v)>$に対する Authority の署名

を含む「認証済参加情報」を得る．$key(v)$はノード v が KOPSON に参加するためのキーである．$TMV(v)$は，Authority が生成するα進数の乱数である（Trustable Membership Vector :TMV と呼ぶ）．メッセージを送信するノードはこのノード認証手順を経て取得した認証済参加情報をメッセージに署名つきで付与する．受信ノードはメッセージを受信するたびに認証済参加情報を検証し，検証に失敗したメッセージを廃棄する．これにより，faulty ノードが意図的にキーを選択して経路表に影響を与える攻撃や，偽造されたメッセージを用いることによる不正動作を回避可能とした．ただし，認証済参加情報をもつ認証済みノードであっても faulty ノードでないことは保証されず，あらかじめ決められたアルゴリズムに従わない場合を想定する必要がある．

② オーバレイ構造維持方式

Skip Graph と同様の階層構造をもつオーバレイネットワーク構造を提案した．

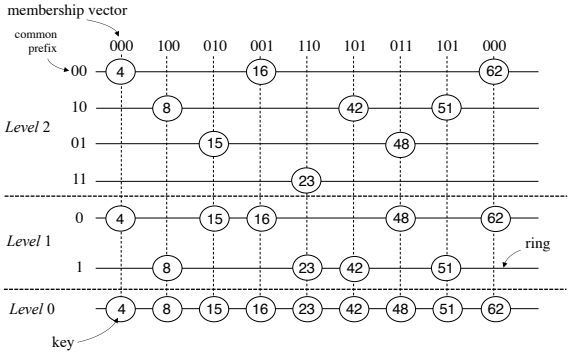

図 1　Skip Graph

Skip Graph は図 1 に示す階層構造を持つ KOPSON である．階層構造の各レベルは，キーの順序に従ってソートされたノードからなるリングから構成されている．リングは環状であり，最大のキーを持つノードの正方向の隣接ノードは最小のキーを持つノードとなる（負方向も同様）．以下，右方向を正方向，左方向を負方向とする．各ノードは十分な桁数のα進数の乱数（MV; Membership Vector）を生成する．レベル i ($i \geq 0$) では，MV の先頭 i 桁と一致するノード同士がリングを形成する．このため，レベル i には最大 αi 個のリングが存在する．レベル 0 のリングには全てのノードが所属する．提案方式では，オーバーレイネットワークの構造の決定に，MV ではなく Authority より与えられた TMV を用いる．これにより，faulty ノードが各レベルの隣接ノードを意図的に選択することが困難となる．

提案方式では，冗長化パラメータ k を用いる．k は，k 個のノードを無作為に選んだときに，少なくとも 1 つは正常なノードが含まれる確率が十分高くなるように選ぶ．各ノード v は Skip Graph と同様，階層化された経路表を持つ．v の経路表のレベル i は，TMV が v の TMV と先頭 i 桁が一致するノードのうち，キーが範囲 $[u, w]$ に含まれるノードのリストである．ただし，u（あるいは w）は v からレベル i リングを左方向（あるいは右方向）にたどったときに，TMV が $i+1$ 桁一致するノードの $k-1$ 個目のノードのキーである．レベル i のノードを，u から w の方向に並べたものをレベル i ノードリストと呼ぶ．例えば，図 2 において $k=3$，v をノード 5 とすると，u はノード 1，w はノード 9 となる．このため，v のレベル i ノードリストは，

ノード 1 からノード 9 までのノードのリストとなる．レベル i ノードリストの各要素（経路表エントリ）は，当該ノードのロケータ（IP アドレスなど），認証済参加情報，公開鍵から構成される．レベルが上がるほど，リングに所属するノード数は減少する．一方，各ノードのレベル i ノードリストには，左方向，右方向とも少なくとも $k-1$ 個のノードが入る．このため，いずれ左方向，右方向でノードリスト内に重複が生じることになる．このレベルを経路表の最大レベル h とする．経路表の高さは $h+1$ である．

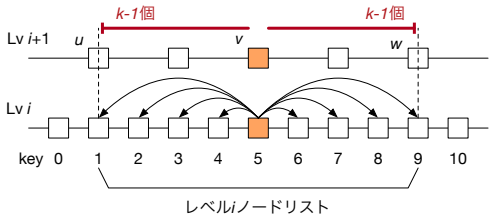

図 2　レベル i ノードリスト

③ ルーティング方式

③-1 ユニキャスト

②の構造上で動作するユニキャストアルゴリズムを提案した．提案手法は，上位レベルから下位レベルへ向かってマルチホップでメッセージを配送する．その際，「経路の冗長化」と「メッセージの検証」によって不正動作に対処する．

提案手法では，KOPSON を用いるアプリケーションが，キーの近傍 k 個のノードにデータや機能を複製配置することを想定し，宛先として指定されたキーs の近傍 k 個のノードを検索する機能を提供する．

任意のノード v の，あるレベル i ノードリストにおける連続する k 個のノード $(n_1, n_2, \ldots, n_k)$ を考える．この k 個が s を中央に含む場合（$s \in [n_{\lceil k/2 \rceil}, n_{\lceil k/2 \rceil+1})$ となるとき中央に含むという），これら k 個のノードのレベル $i-1$ ノードリストは $[n_1, n_k]$ をカバーする．また，レベル $i-1$ でこの範囲に少なくとも k 個のノードが存在するため，レベル $i-1$ ノードリストにも s を中央に含む連続する k 個のノードが存在する．このため，あるノード v が s を中央に含むレベル i ノードリストの連続する k 個のノードにメッセージを送信し，次にこのメッセージを受信した各ノードが s を中央に含むレベル $i-1$ の k 個のノードに送信するという動作を再帰的に繰り返すことで，最終的にはレベル 0 で s を中央に含む k 個のノードにメッセージを送信できる．図 3 は k=4 の場合の転送経路である．

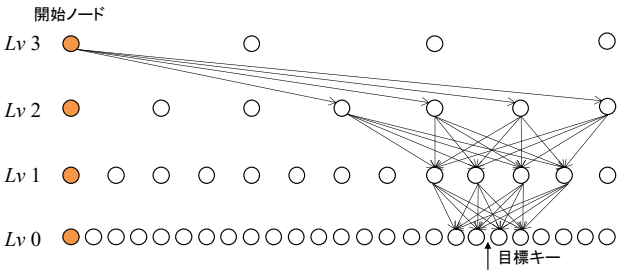

図 3　転送経路（ユニキャスト）

各ノードは独立してレベル $i-1$ の転送先の k 個のノードを選ぶが，これらの k 個のノード集合はすべて同じとなる．したがって，k 個のすべてのノードが正常である場合，レベル $i-1$ の k 個のノードはそれぞれレベル i の k 個のノードから同じメッセージを重複して受信する．このため，k 個の一部が faulty ノードであったとしても，少なくと 1 つのノードが正常ならばレベル $i-1$ の転送処理を継続できる．

あるレベルにおける転送先ノード集合のうち少なくとも 1 つのノードが faulty ノードでなければ，1 ホップのメッセージの転送は成功する．ノード全体で faulty ノードが占める割合を f とすると，1 ホップの転送に成功する確率は $(1-f^k)$ である．ホップ数が増えるにしたがい転送の成功率は低下し，i ホップの検索が成功する確率（i ホップ目の k 個のノードのうち少なくとも 1 つがメッセージを受信する確率）は $(1-f^k)^i$ となる．最大ホップ数は経路表の最大レベル h に等しく，メッセージ到達率は $(1-f^k)^h$ よりも大きくなる．

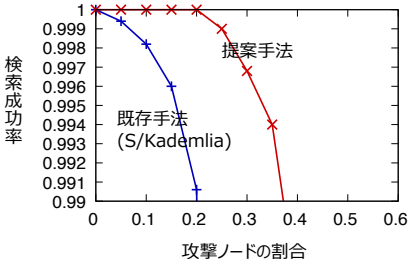

図 4　メッセージ到達率

提案手法のシミュレーション評価を行ない，ビザンチン障害ノードが存在する環境のもとであっても，既存の DHT 向け構造化オーバレイネットワークである S/Kademlia[4]と比較しても高い検索成功率・性能が得られることが分かった（図 4）．

③-2 マルチキャスト

ユニキャストに加え，②の構造上で動作するマルチキャストアルゴリズムも提案した．マルチキャストにおいても，ユニキャストと同様，「経路の冗長化」と「メッセージの検証」によって不正動作に対処する．提案手法は，上位レベルから下位レベルへ向かってマルチホップでメッセージを配送する．

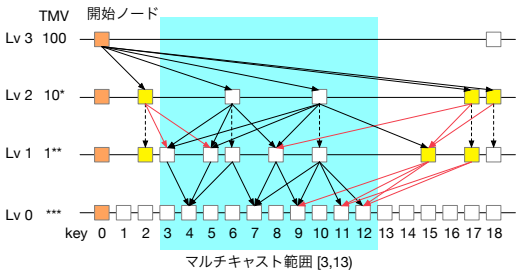

図 5　転送経路（マルチキャスト）

提案手法では，あるノード v に対してメッセージを送信するノード（上流ノードと呼ぶ）を k 個にすることで配送経路を冗長化する．k 個のノードすべてが faulty である確率は十分低いため，v は高確率でメッセージを受信できる．k 個の冗長配送経路をマルチキャスト範囲の端まで確保するため，提案手法ではマルチキャスト範囲の外側のノードを利用する（補助ノードと呼ぶ）．マルチキャスト範囲内のノードと同様，補助ノードもメッセージの配送に関与する．図 5 は，k=3 の場合の提案手法によるマルチキャスト配送経路の例である．図において，黄色のノードは補助ノードであり，赤色の線は補助ノードによるメッセージ送信を表す．最初のホップ(key=0 のノードからの直接配信)を除き，各受信ノードが上流ノードを 3 つ確保している．提案方式におけるメッセージ到達率はホップ数が大きくなるにつれ小さくなる．メッセージ到達率の期待値は，ユニキャストの場合と同様，最大ホップ数（経路表の最大レベル h）のときの各受信ノードへのメッセージ到達率は$(1-f^k)^h$以上となる．

④ アプリケーション

提案方式のアプリケーションとして，IPFS の動作確認を行なった．IPFS (InterPlanetary File System)は分散型のファイルストレージを実現するオープンソースプロジェクトの一つである．IPFS はデータを非集中型アーキテクチャにより管理することが可能であり，ブロックチェーンのオフチェーンストレージとしての用途と親和性が高く，応用が進んでいる．

IPFS のルーティングレイヤは Kademlia によって実現されている．本研究では，この IPFS のルーティングレイヤを提案方式に置き換えるプロトタイプ実装を行った．プロトタイプを用いて Google Cloud Platform 上で 32 拠点のデータセンターを用いた動作検証を行い，正しい動作を確認した．

参考文献

[1] Fiat, A., Saia, J. and Young, M.: Making Chord Robust to Byzantine Attacks, Proc. of European Symposium on Algorithms, pp. 803–814 (2005)
[2] Needels, K. and Kwon, M.: Secure Routing in Peer-to-Peer Distributed Hash Tables, Proc. of the 2009 ACM Symposium on Applied Computing, pp. 54–58 (2009).
[3] Aspnes, J. and Shah, G.: Skip graphs, ACM Transaction on Algorithms, Vol. 3, No. 4, pp. 1–25 (2007)
[4] Baumgart, I and Mies, S., “S/Kademlia: A Practicable Approach Towards Secure Key-based Routing,” In IEEE ICPADS, pp.1-8 (2007).

５．主な発表論文等

〔雑誌論文〕　計1件（うち査読付論文　1件／うち国際共著　0件／うちオープンアクセス　1件）

１．著者名 Ishihara Shintaro, Akiyama Toyokazu	４．巻 28
２．論文標題 Towards a Dataflow Platform in a Hierarchical Network: A Proposal for a Dataflow Component Management Method	５．発行年 2020年
３．雑誌名 Journal of Information Processing	６．最初と最後の頁 599, 610
掲載論文のDOI（デジタルオブジェクト識別子） 10.2197/ipsjjip.28.599	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 －

〔学会発表〕　計4件（うち招待講演　0件／うち国際学会　0件）

１．発表者名 秋山 豊和, 寺西 裕一, 安倍 広多
２．発表標題 ビザンチン障害耐性を有する構造化オーバレイネットワーク ByzSkip における冗長化パラメータkの動的制御の検討
３．学会等名 RIXX-PIoT workshop 2023
４．発表年 2023年

１．発表者名 本多 徹, 寺西 裕一, 安倍 広多
２．発表標題 ビザンチン障害耐性を有するキー順序保存型構造化オーバレイネットワークのためのマルチキャスト手法の提案
３．学会等名 第29回マルチメディア通信と分散処理ワークショップ論文集
４．発表年 2021年

１．発表者名 寺西 裕一, 秋山 豊和, 安倍 広多
２．発表標題 ビザンチン障害耐性を備えるキー順序保存型構造化オーバーレイネットワークの実現に向けて
３．学会等名 情報処理学会　第186回　マルチメディア通信と分散処理（DPS）研究会
４．発表年 2021年

１．発表者名 寺西 裕一，秋山 豊和，安倍 広多
２．発表標題 ビザンチン障害耐性を備えるキー順序保存型構造化オーバーレイネットワークの検討
３．学会等名 第12回　広域センサーネットワークとオーバレイネットワークに関するワークショップ
４．発表年 2021年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

―

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
研究分担者	安倍　広多 (Abe Kota) (40291603)	大阪公立大学・大学院情報学研究科・教授 (24405)	
研究分担者	秋山　豊和 (Akiyama Toyokazu) (80324862)	京都産業大学・情報理工学部・教授 (34304)	

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関