DHT を用いたスケーラブルな匿名通信手法の提案

及川 一樹 †　　王 家宏 ‡　　児玉 英一郎 ‡　　高田 豊雄 ‡

†‡ 岩手県立大学大学院ソフトウェア情報学研究科
020-0193 岩手県岩手郡滝沢村滝沢字巣子 152-52
g231g008@edu.soft.iwate-pu.ac.jp†, {wjh, kodama, takata}@iwate-pu.ac.jp‡

あらまし 近年，インスタントメッセンジャーや WWW を利用した，様々なコミュニケーションサービスが登場し，普及している．しかし，それらのサービスを利用するためには，個人情報の入力を必要とする場合があることや，サービス提供者などの特権を持つ者が，プライベートなコミュニケーションの内容を読み取ることが可能であるといった問題がある．これらの問題は特権の存在しないピュア P2P を用いることにより解決することが可能となり，特に分散ハッシュテーブル (DHT) を用いることによりスケーラブルにすることが出来るが，DHT には匿名性が無いため，直接適用することは出来ない．そこで我々は，DHT を用いたピュア P2P 環境においてスケーラブルで効率の良い匿名通信手法を提案する．

A Proposal of Scalable Anonymous Communication Scheme with DHT

Kazuki Oikawa†　　Jiahong Wang‡　　Eiichiro Kodama‡　　Toyoo Takata‡

†‡Graduate School of Software and Information Science, Iwate Prefectural University
152-52, Sugo, Takizawa, Takizawa village, Iwate, 020-0193, Japan
g231g008@edu.soft.iwate-pu.ac.jp†, {wjh, kodama, takata}@iwate-pu.ac.jp‡

Abstract Recently, communication services based on instant messengers and WWW have become more and more popular. One problem is that, however, to use those services, one has to provide his/her personal information. In addition, the privileged users such as service providers can listen in on the communication. These problems can be solved with pure-P2P network where there are no privileged users. Generally, DHT(Distributed Hash Table) is used for the scalability. However, it is known that DHT cannot preserve anonymity, and if the anonymity is an issue, we cannot take advantage of DHT directly. In this paper a scalable and effective anonymous communication scheme using DHT over pure-P2P network is proposed.

1　はじめに

現在，P2P ネットワーク技術は一般的なものとなり，BitTorrent といったファイルの転送を行うアプリケーションや，映像配信を行うアプリケーションなどに広く利用されている．

P2P ネットワーク技術には中央に管理用サーバを持つハイブリッド P2P 型や，すべてのノードが対等な関係にあるピュア P2P 型などの種類がある．ピュア P2P 型は管理用サーバが不要で単一故障点が存在しないという点などで優れており，問題点だったデータの探索効率が悪いという点も，分散ハッシュテーブル (DHT) に代表される構造型ネットワークの登場により解決

され，今後も様々な用途に活用されると考えられている．

また，近年では無料で利用できるインスタントメッセージングサービスや，WWW を利用した様々なコミュニケーションサービスが登場し，普及している．しかし，それらのサービスを利用するためには，サービスの利用に必要のない個人情報の入力を必要とする場合があることや，サービスの提供者などの特権を持つ者がプライベートなコミュニケーションの内容を読み取ることが可能であるといったプライバシー上の問題がある．他にも，採算がとれないなどと言った理由で，サービスの提供を一方的にやめてしまうなどの問題がある．

これらの問題はピュア P2P 型の P2P ネットワーク技術を用いることにより，サービスの提供者といった特権を持つ存在を不要とし，利用者自身によってサービスを構成することが出来るため，上記の問題を解決することが可能となる．また，分散ハッシュテーブルを利用することにより，スケーラブルで効率の良いサービスの提供も可能になると考えられる．

しかし，分散ハッシュテーブルには匿名性が無い．すなわち，宛先となるキーを知ることができれば，そのキーと IP アドレスの組み合わせが漏洩してしまい，匿名性が確保できない．この問題は，インターネットを利用して知り合った知人とインスタントメッセンジャーのようなアプリケーションで通信することを想定した場合，IP アドレスが漏洩し，住んでいる地域などの情報が相手に知られてしまうため，通信相手が必ずしも信頼できるとは限らない環境においては問題となる．

そこで我々は，文献 [1] の手法を改良した，分散ハッシュテーブルを用いたピュア P2P 環境においてコミュニケーションシステムを実現するために必要な適度な匿名性を持つ匿名通信手法を提案する．本提案手法は，送受信者の匿名性を確保するほか，管理用サーバを必要としないためスケーラビリティに優れる手法となっている．

2 既存研究

分散ハッシュテーブルを利用した匿名通信に関する提案が幾つかなされている [2][3].

これらの提案では，送信者は受信者がいずれかのグループに所属するように，いくつかのグループを作成する．グループ間のルーティングは分散ハッシュテーブルを利用して効率よく行い，グループ間でやりとりされるメッセージを多重暗号化することで，前後のグループ以外の情報を漏らさないことにより，送信者の匿名性を確保している．そしてペイロードは受信者の公開鍵で暗号化し，各グループ内でペイロードをブロードキャストすることで，受信者を他のグループメンバに漏らすことなく，ペイロードの配送が可能になる．

しかし，これらの提案には以下に示す 2 つの問題点が存在する．

受信者の匿名性が送信者に対して確保されない これらの提案では，内部告発など送信者の匿名性のみが重要視される用途向けに設計されているため，受信者の匿名性は考慮されていない．そのため，宛先として分散ハッシュテーブルでノードに割り当てられるハッシュ値をそのまま利用するため，宛先となるハッシュ値よりそのノードの IP アドレスを特定することが可能になってしまう．

ディレクトリサーバを必要とする 中継ノードの公開鍵を取得するために，各ノードに割り当てられるハッシュ値と公開鍵のペアを保存するサーバが必要なほか，中継ノードの選出にもディレクトリサーバを利用するため，分散ハッシュテーブルのスケーラビリティという特徴が生かせず，単一故障点となり得る．

3 提案手法

本節では，1 節で述べた目的に合い，また 2 節で挙げた問題点を解決する手法を提案する．

まず，3.1 節で本提案手法が利用する，分散ハッシュテーブルについて簡単に述べ，3.2 節で本提案手法が確保する匿名性に関して述べる．そして，3.3 節で本提案手法の概要を示した後，3.4 節以降の節で詳細を説明する．

3.1 分散ハッシュテーブル

分散ハッシュテーブル (DHT: Distributed Hash Table) とは，ノードに IP アドレスとは別のアドレスとしてハッシュ値を割り当て，ハッシュ値が近いノードに関する情報は密に，ハッシュ値が遠いノードに関する情報は疎に持つ．これにより，ネットワーク全体の情報を持たなくても，ネットワークサイズを N とした時，$O(\log N)$ 程度のホップ数で，あるハッシュ値を持つノードの探索が可能となり，ピュア P2P 型ながらスケーラビリティに優れるという特徴を持つ．代表的なアルゴリズムとしては，Chord[4] や Kademlia[5] が知られている．

分散ハッシュテーブルは，一般的にはインターネットの様な既存のネットワークの上に新たなアドレス体系を持つネットワークを構築するため，基底となるネットワークのことをアンダーレイネットワークと呼ぶのに対し，オーバーレイネットワークと呼ばれる．

3.2 確保する匿名性

本提案手法では分散ハッシュテーブルを用いることを前提としているため，分散ハッシュテーブルで用いられるものと同じ体系のオーバーレイネットワークアドレスを利用した二者間通信に対して匿名性を提供する．すなわち，接続先として IP アドレスの代わりにオーバーレイネットワークアドレスを指定して行う二者間の通信に対して匿名性を提供するということである．また，匿名性を確保するに当たり，第三者から隠す情報は，アンダーレイネットワークのアドレスとする．これは，オーバーレイネットワークアドレスは自由に変更を行うことが出来る一方，アンダーレイネットワークアドレスは自由に変更することが難しく，また，場合によっては参加者の所在地などのプライバシーに関わるような情報が漏れるためである．よって，本提案手法では，オーバーレイネットワークのアドレスと，アンダーレイネットワークのアドレスを結びつけることが困難な状態であることを，匿名性が確保されている状態であると定義する．

そして，前述した匿名性の定義の下，送受信者の匿名性を確保する．これは，通信相手を含めた他のノードに対して，自身のオーバーレイネットワークアドレスが知られる可能性はあるが，アンダーレイネットワークアドレスを特定することは困難であると言うことを意味する．

しかし，オーバーレイネットワークアドレスは知られる可能性があるため，通信を行っている二者のオーバーレイネットワークアドレスの組み合わせも特定される可能性もあるが，オーバーレイネットワークアドレスは容易に変更することができ，また複数持つことも可能であるため，使い捨てのアドレスを用いて通信を行うなど，上位層からの利用方法を工夫すれば，この点は問題にならないと考えられる．

また，本提案手法は 1 節で示したように，公権力を持つ者を相手とした厳密な匿名性を確保することを目的としていない．よって，大規模な結託に対する耐性や，全ネットワークの盗聴に対する耐性は持ち合わせないが，小規模な結託や盗聴に対しては匿名性を維持し続けることができると考えられる．

3.3 概要

まず，本提案手法では第三者から送受信者の匿名性を確保するほかにも，送受信者間でも匿名性が保たれるようにするために，オニオンルーティング [6] に似た働きをする多重暗号経路を送受信者両方側で構築する．

そして，分散ハッシュテーブルを利用して送受信者の多重暗号経路の終端ノードを探索することで，スケーラブルに送受信者の多重暗号経路間を繋げ，二者間の双方向通信を可能にする．

また，中継ノードの選出や，多重暗号経路の構築に必要な各ノードの公開鍵の取得に関しては，従来の手法のように鍵管理サーバを利用するのではなく，分散ハッシュテーブルのルーティングテーブルを利用する．このようにすることで，追加コストを必要としないで中継ノードに対応する公開鍵が取得できるほか，生存率の高い中継ノードの選出を行うことが可能となる．

本提案手法では，伝搬遅延やパケット損失率に伴う性能評価を行うため，ノード間の通信に UDP を利用することを前提とするが，多少の変更を加えることで TCP を利用することも可能となる．

3.4 ノードの公開鍵と中継ノードの選出

本提案手法を利用して匿名通信を行うノードはすべて，分散ハッシュテーブルを利用した 1 つのネットワークに属する．そして，分散ハッシュテーブルに参加する際に，楕円曲線暗号の秘密鍵/公開鍵のペアを生成し，ここで生成した公開鍵を分散ハッシュテーブルにおけるハッシュ値の代わりとして利用する．このようにすることで分散ハッシュテーブルにおいてノードの公開鍵を利用した探索が可能となる．

一般的な分散ハッシュテーブルでは，SHA-1 のハッシュ値を利用しているため，160bit の長さを想定しているが，楕円曲線暗号の公開鍵は十分な安全性を確保しても 256bit 程度と RSA 暗号などと比べ非常に短いため，先に挙げた分散ハッシュテーブルのアルゴリズムを変形することなく，自然に適用することができると考えられる．

そして，中継ノードを選出する際には，分散ハッシュテーブルのルーティングテーブルを用いる．分散ハッシュテーブルにおいてノードに割り当てられているハッシュ値は，上記の通り楕円曲線暗号の公開鍵であるため，ルーティングテーブルには，IP アドレスとそのノードに対応する公開鍵のペアが登録されている．また，一般的な分散ハッシュテーブルのアルゴリズムでは，ルーティングテーブルは適度にメンテナンスされ，ネットワークから離脱したノードや故障したノードなどの情報が含まれている可能性が低くなっている．そのため，ルーティングテーブルより中継ノードを選出した場合，離脱や故障したノードが含まれる確率が低いため，後述する多重暗号経路を構築できない可能性は十分低くなると考えられる．

3.5 多重暗号経路

多重暗号経路とはオニオンルーティングを変形させたもので，オニオンルーティングと同じように，送信元を隠蔽する機能を持つ．

まず，多重暗号経路を構築するノード (始点ノード) は，3.4 節で示した方法を用いて，中継ノード $R_1 \cdots R_k$ を選出する．そして，ルーティングテーブルに登録されている各中継ノードの公開鍵を $Pub(R_1) \cdots Pub(R_k)$ とする．次に，各中継ノードとの共通鍵を生成するために，楕円曲線暗号の秘密鍵 $E_1 \cdots E_k$ とそれに対応する公開鍵 $P_1 \cdots P_k$ をランダムに生成後，楕円 DiffieHellman 鍵共有アルゴリズムを利用し，$E_1 \cdots E_k$ と $Pub(R_1) \cdots Pub(R_k)$ から共通鍵 $S_1 \cdots S_k$ を作成する．そして，以下の式に基づいてメッセージ M_1 を作成する．

$$M_i = \begin{cases} P_i, \langle R_{i+1}, M_{i+1} \rangle_{S_i} & (i \neq k) \\ P_k, \langle Payload \rangle_{S_k} & (i = k) \end{cases}$$

ここで，カンマは連接記号とし，$\langle PlainText \rangle_S$ は，共通鍵 S で $PlainText$ を暗号化することを表す．

そして，最初の中継ノード R_1 に経路判別用のラベル L_1 とパディングを付与したメッセージ $\{L_1, M_1, Padding\}$ を送信する．このラベルは，同じ連続した二つ以上の中継ノードを経由する多重暗号経路が複数存在した場合でも，複数の経路を識別するために利用され，IEEE802.1Q (VLAN タギング) と似た働きをするものである．

メッセージを受信した中継ノード R_i では，$Pub(R_i)$ に対応する秘密鍵と受信した M_i に含まれる P_i を用いて楕円 DiffieHellman 鍵共有を行い S_i を生成する．そして，M_i の暗号化された部分を S_i を用いて復号することで R_{i+1} と M_{i+1} を取り出し，次の中継ノード R_{i+1} へ新しいラベル L_{i+1} を付与したメッセージ $\{L_{i+1}, M_{i+1}, Padding\}$ を送信する．この処理を終端ノード R_k に達するまで繰り返す．

以上のように，多重に暗号化されたメッセージを複数のノードで中継することで，各中継ノードは前後の中継ノードの情報しか得られず，また，前の中継ノードが中継ノードなのかそれとも始点ノードなのかを判別できなくすることで，送信元を隠蔽する．そして各中継ノード R_i は前後の中継ノードとそれに対応するラベル ($\langle R_{i-1}, L_i \rangle$ と $\langle R_{i+1}, L_{i+1} \rangle$) をキーとして共通鍵 S_i を記憶することで，2 回目以降の通信では楕円 DiffieHellman 鍵交換を行うことなく，直接その経路に対応した共通鍵 S_i を用いて，メッセージの復号を行う．また，終端ノードから始点ノードへ逆向きにメッセージを送信する場合は，中継ノード ($R_k \cdots R_1$) を経由するたびに共通鍵 ($S_k \cdots S_1$) を用いて暗号化を行い，始点ノードでは $S_1 \cdots S_k$ を用いて順にメッセージを k 回復号することで，平文を手に入れる．

多重暗号経路では以上のような処理により，中継ノードに通信内容を漏らすことなく，始点ノードと終端ノード R_k 間での通信が可能になる．以降の節において “多重暗号経路を経由して送信” と言う場合は，上記の手順を用いて始点ノードと終端ノード間で通信を行うことを指す．

3.6 匿名通信コネクション

匿名通信を行う二者 (A と B) は，以下の手順を行って多重暗号経路を作成し，匿名通信を行う準備をする．以下では A の場合について述べているが同様の手順によって B も多重暗号経路を作成する．

1. A は宛先として用いる楕円曲線暗号の秘密鍵を作成し，それに対応する公開鍵を $Pub(A)$ とする．
2. ペイロードに $Pub(A)$ を指定して多重暗号経路を作成する．ここで作成した多重暗号経路を $MCR(A)$ とし，$MCR(A)$ の終端ノードを TN_A とする．
3. TN_A は $MCR(A)$ 作成時にペイロードとして受け取った $Pub(A)$ を用いて，分散ハッシュテーブルに対し，$Pub(A)$ をキーに自身の情報 TN_A を保存する．

ここで作成する多重暗号経路は 1 つの公開鍵に対して複数個用意することが可能であり，複数個用意することにより，中継ノードが故障した場合でも通信を継続することができるようになる．

そして，A が B に対して匿名通信コネクションを確立する際には以下の手順を行う．

1. A は何らかの方法を用いて B の公開鍵である $Pub(B)$ を取得する
2. $MCR(A)$ を通して TN_A に対して分散ハッシュテーブルで $Pub(B)$ を探索するように依頼．TN_A は探索の結果，TN_B の情報を取得し，A に返信する
3. A は、$MCR(A)$ を通して TN_A に TN_B へのメッセージ転送を依頼することで，B にメッセージを送信し，B も同様の方法で A に対してメッセージを送信する．このように A と B の間でメッセージをやりとりすることで，匿名通信コネクション用の共通鍵を共有する

以降の通信では，上記の手順 3. と同じ方法でメッセージをやりとりすることで通信を行うが，コネクション用共通鍵で暗号化されたメッセージに，自身の多重暗号経路の終端ノード一覧を付随させることで，分散ハッシュテーブルへの問い合わせを抑制し，通信遅延を軽減させる．

4 性能評価

ここでは，匿名通信コネクションの耐故障性とスケーラビリティに関しての評価実験を行う．耐故障性の評価では，パケットの損失や，ノードの加入/離脱が発生するような環境においても匿名通信を継続できるかどうかを評価し，スケーラビリティの評価では，ノード数が増えても性能に影響しないかどうかを評価する．

また，評価には，単一のマシン上で動作し，UDP パケットの配送遅延や損失率を設定可能な独自のシミュレータを用い、UDP のパラメータである RTT や再送タイムアウト，最大再送数はそれぞれ 50 ミリ秒，200 ミリ秒，2 回とし，匿名コネクションのパラメータである中継ノード数や再送タイムアウト，最大再送数はそれぞれ 3 ノード，3 秒，0 回とした．

そして，再送タイムアウトに達するまでに応答がない場合は，パケットが損失したものと判断し，パケットを再送し，最大再送数で指定された回数再送しても応答がない場合は，配送失敗とした．なお，分散ハッシュテーブルのアルゴリズムには，Kademlia を利用した．

4.1 耐故障性に関する評価

ノード数を 1000，パケット損失率を 5%，ノードの離脱/参加間隔を 0.5 秒とした場合のコネクション確立成功率やコネクション確立所要時間，メッセージの到達率，RTT の変化について評価した結果を図 1 に示す．

多重暗号経路数が 1 つの場合では，コネクション確立の成功率やメッセージの到達率が必ずしも高くないが，多重暗号経路 2 本を同時に利用することにより，どちらも 100%に達した．RTT もパケットロスなどが発生しない場合には，理論値では 350 ミリ秒となるが，上述のような環境においても，多重暗号経路を 2 本同時に利用

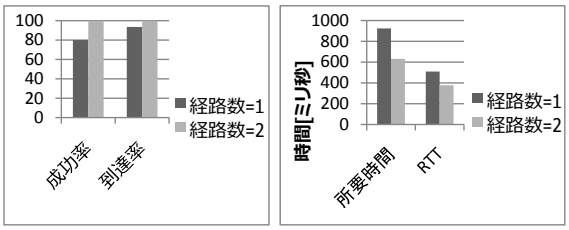

図 1: 耐故障性に関する評価

することにより平均 377 ミリ秒程度に抑えられるなど，パケットの損失やノードの加入/離脱がある環境でも問題なく利用できると考えられる．

また，今回想定した環境よりもパケット損失率が高い場合や，ノードが頻繁に出入りするような環境においても，同時に利用する多重暗号経路数を増やすことにより柔軟に対応できると考えられる．

4.2 スケーラビリティに関する評価

パケット損失やノードの離脱が無い環境において，ノード数を変化させた場合のコネクション確立所要時間や RTT の変化について評価した結果を図 2 に示す．

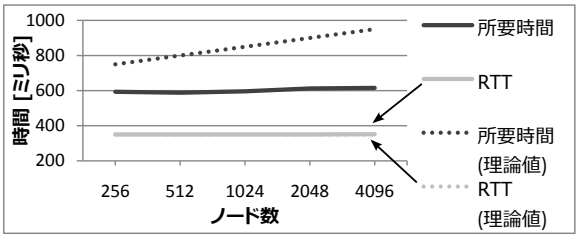

図 2: スケーラビリティに関する評価

コネクションの確立所要時間 (ET) や，RTT は，ノード数を N，中継ノード数を k，UDP の RTT を RTT_{UDP} とすると，

$$ET = (2k + 1 + \log_2 N) \times RTT_{UDP} \quad (1)$$

$$RTT = (2k + 1) \times RTT_{UDP} \quad (2)$$

となるため，ノード数に依存する部分はコネクションの確立部分のみとなる．図 2 に示すとおり，RTT の評価結果は概ね理論値通りとなっている．コネクションの確立所要時間は評価結果と理論値に差があるが，これは式 (1) に含まれる $\log_2 N$ の項は，分散ハッシュテーブルの探索における最悪ホップ数となっているためである．以上より，本提案手法は十分にスケーラブルであると思われる．

5 おわりに

本研究では，ピュア P2P 環境におけるコミュニケーションシステムの実現に適したスケーラブルな匿名通信手法を提案した．また，耐故障性やスケーラビリティに関する評価を行い，実環境においても動作に耐えうる性能であることを確認した．

今後の課題として，本稿の匿名通信を利用した応用アプリケーションとして，インスタントメッセンジャーや掲示板などの実現がある．

参考文献

[1] 及川一樹, 王家宏, 児玉英一郎, 高田豊雄. 受信者の匿名性を送信者に対しても確保可能な DHT を利用した匿名通信路構築手法の提案. In *SCIS2009*. 2F3-3, 2009.

[2] Li Zhuang, Feng Zhou, Ben Y. Zhao, and Antony Owstron. Cashmere: Resilient Anonymous Routing. In *NSDI'05*. pp. 301–314, USENIX Association, 2005.

[3] 近藤正基, 齋藤彰一, 松尾啓志. DHT を用いた双方向匿名通信路の提案. 情報処理学会研究報告. CSEC, [コンピュータセキュリティ], Vol. 2008, No. 71, pp. 195–202, 2008.

[4] Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, and Hari Balakrishnan. Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications. In *SIGCOMM '01*. pp. 149–160, ACM, 2001.

[5] Petar Maymounkov and David Mazières. Kademlia: A Peer-to-peer Information System Based on the XOR Metric. In *IPTPS '01*. pp. 53–65, Springer-Verlag, 2002.

[6] David M. Goldschlag, Michael G. Reed, and Paul F. Syverson. Hiding Routing Information. In *Proceedings of the First International Workshop on Information Hiding*. pp. 137–150, Springer-Verlag, 1996.