

# 時間重要度をもつ位置依存コンテンツ検索のための P2P ネットワークの検討

高橋健太郎<sup>†</sup> 中野 宏一<sup>††</sup> 春本 要<sup>†††</sup> 西尾章治郎<sup>††</sup>

<sup>†</sup> 大阪大学工学部電子情報工学科 〒565-0871 大阪府吹田市山田丘 2-1

<sup>††</sup> 大阪大学大学院情報科学研究科 〒565-0871 大阪府吹田市山田丘 1-5

<sup>†††</sup> 大阪大学大学院工学研究科 〒565-0871 大阪府吹田市山田丘 2-1

E-mail: <sup>†</sup>takahashi.kentaro@ise.eng.osaka-u.ac.jp, <sup>††</sup>{nakano.hirokazu,nishio}@ist.osaka-u.ac.jp,  
<sup>†††</sup>harumoto@eng.osaka-u.ac.jp

あらまし 近年、P2P ネットワークに対する注目が高まってきている。店舗の営業時間やキャンペーン情報といった、位置と時間に依存する地域コンテンツを検索するサービスを実現する上で、ピアの位置情報に基づいたトポロジを構成すれば、検索対象の領域にクエリを効率的に配送できる。しかし、位置情報以外は考慮されていないため、領域内のピア全てにクエリが転送され、膨大なトラフィックが生じる。そこで本研究では、コンテンツの位置と時間帯別の重要度に着目したリンクの構築とルーティングの手法を提案する。位置依存の P2P ネットワークにおいて、 $n$  ホップ以内のピアのもつコンテンツの位置と時間帯別の重要度をリンクテーブルに保持してリンクを構築する。そして、クエリの内容とテーブルの情報からクエリの転送先を選択することで、クエリに対して適合度の高いレスポンスの取得数を維持しつつトラフィックを削減する。また、シミュレーション実験により提案手法を評価する。

キーワード P2P, 情報検索, リンクテーブル, 位置情報, 時間帯別の重要度

## A study about P2P-Network for searching location based contents with importance dependent on time

Kentaro TAKAHASHI<sup>†</sup>, Hirokazu NAKANO<sup>††</sup>, Kaname HARUMOTO<sup>†††</sup>, and Shojiro NISHIO<sup>††</sup>

<sup>†</sup> Division of Electronic and Information Engineering, School of Engineering, Osaka University  
2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

<sup>††</sup> Graduate School of Information Science and Technology, Osaka University  
1-5 Yamadaoka, Suita, Osaka 565-0871, Japan

<sup>†††</sup> Graduate School of Engineering, Osaka University  
2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

E-mail: <sup>†</sup>takahashi.kentaro@ise.eng.osaka-u.ac.jp, <sup>††</sup>{nakano.hirokazu,nishio}@ist.osaka-u.ac.jp,  
<sup>†††</sup>harumoto@eng.osaka-u.ac.jp

**Abstract** Recently, Peer-to-Peer (P2P) applications are becoming popular. On a searching service to regional contents dependent on location and time, for example, store hours and campaign information, a topology constructed based on location information of peer can deliver query to search area efficiently. But regarding only location information, query are sent to all peers in the search area with a huge network traffic. In this paper, we propose a method to construct link regarding importance at each period of time and location of contents, and a query routing strategy. On P2P networks dependent on location, each peer makes link to peers within  $n$ -hop and holds location and time-importance of their contents in link-table. By selecting peers to forward each query according to the link-table and the query, traffic is reduced with keeping number of response messages that match to queries highly. And by simulation experiments, we evaluate the proposed P2P network.

**Key words** Peer-to-Peer, information retrieval, link-table, data of location, importance at each period of time

## 1. ま え が き

近年、携帯端末の普及やネットワーク技術の発展とともに、ユビキタス環境が整いつつある。今後、ネットワーク中の端末数や情報は飛躍的に増加すると考えられる。このような環境では、端末間の通信形態として P2P ネットワークが適している。P2P ネットワークは、ピアと呼ばれる端末がクライアントとサーバの両方の機能をもち、ピア自身が情報を管理し発信できる論理ネットワークアーキテクチャである。P2P ネットワークは従来のサーバ・クライアント型アーキテクチャのようにサーバの負荷が増大しない。また P2P ネットワークではあるピアがサービスを提供できない状態になった場合でも、他のピアに配置されたキャッシュを利用できる可能性があり、可用性の高いデータ共有サービスの実現に非常に有効である。この環境を利用し、店舗の営業時間やキャンペーン情報といった、地域に密接に関連し、かつ、時間に依存した重要度をもつコンテンツを、ユーザが位置と時間帯を指定して検索するようなサービスが実現されると考えられる。

既存の P2P ネットワークにおける代表的な検索手法であるフラッディングベースの検索には、取得できる情報数とトラフィックとにトレードオフの関係がある、この問題に対し、ピアの位置情報を考慮したトポロジを形成する手法が提案されており、これらを用いることで目的の地域へ効率的にクエリを転送することができる。しかし、位置情報以外の要素は考慮されていない。そのため、検索対象の領域内でフラッディングを行った際に、クエリに適合しないコンテンツをもつピアにもクエリが転送され、トラフィックの問題は部分的にしか解決されない。そこで本研究では、位置依存の P2P ネットワークを基盤とし、コンテンツのもつ時間依存の重要度と位置情報に着目したクエリの転送方式を提案する。各ピアは自身のもつコンテンツの位置情報によるトポロジを構築しているものとする。その後、基盤とする P2P ネットワークのリンクに加え、 $n$  ホップ内の他ピアとのリンクを構築するとともに、それらのピアがもつコンテンツの位置情報と時間に依存する重要度をリンクテーブルに保持する。ピアはクエリを受信すると、クエリの探索時間帯とリンクテーブルの情報からクエリを送信するピアを選択し、クエリに対して適合度の低いピアへのクエリ転送を抑制することで、トラフィックを削減する。

以下、本稿では 2. で関連研究について述べ、3. で提案方式の詳細について述べる。4. で提案方式についての評価を行い、5. で提案方式の課題及び拡張案について述べる。最後に 6. で本稿をまとめる。

## 2. 関 連 研 究

P2P ネットワークは一般に Hybrid 型と Pure 型とに分類される。Hybrid 型 P2P ネットワークではインデックスを一元管理するサーバが存在するという点でスケラビリティや可用性に問題があり、本稿では考慮しない。Pure 型 P2P ネットワークにおける検索手法の代表的なものとして、フラッディングによる方式と、分散ハッシュテーブル (DHT: Distributed

HashTable) による方式がある。DHT を用いれば非常に少ないメッセージ数で検索を行うことができる。しかし、DHT では検索時に情報の識別子を指定しなければならないため、一度に複数のコンテンツを取得したい場合や、キーワードの部分一致などによる検索には不向きである。フラッディングによる検索手法は、複雑な条件比較などを含む任意のクエリに対応できる反面、前章で述べたように検索トラフィックの増大の問題がある。これを解決するには、扱うコンテンツに適したトポロジを構成すればよい。

地域コンテンツを扱うには、位置要素が重要となる。位置要素によるトポロジの構成手法としては LL-Net (Location-based Logical Network) [3] や GeoPeer [1] が提案されている。LL-Net では、各ピアは地域をエリアに分割して階層的に捉えることで、近くのエリアには密に、遠くのエリアには疎に外部リンクを構築する。そのため、遠距離のピアへも少ないホップ数でクエリが届く。また、フラッディング範囲をエリア内に限定することでトラフィックの増加を抑制している。しかし位置情報以外は考慮されず、検索対象のエリア内でフラッディングを行った際に、クエリに適合しないコンテンツをもつピアにもクエリが転送されるという問題がある。

一方、GeoPeer では、ピアの位置情報によるドロネー図を用いたトポロジを構成する。トポロジは各ピアが頂点である 3 角形の集合となり、各 3 角形の外接円の内部に他のピアが含まれないという特性がある。この特性を利用し、各ピアはクエリの探索領域に最も近づくリンクを選択してクエリを探索領域まで送ることができる。また、クエリの転送状況に応じて、遠距離のピアまでのホップ数を短縮してクエリが探索領域まで届くようにリンクを構築する。しかし位置的な範囲検索において、隣接ピアの位置情報しかもたないため、LL-Net と同様の問題が起きる。

隣接ピアの保持するコンテンツの情報によってクエリの送信先を決定する手法として、RI (Routing Indices) [2] が提案されている。RI では、クエリをフラッディングするかわりに、隣接するピアの保持する情報数を考慮し、応答の期待できるピアへクエリを転送する方式を提案している。ピアは隣接するピアのもつコンテンツの情報を管理し、クエリを受信すると、検索対象となっている情報を最も多く保持しているピアへクエリを転送する。これを繰り返すことにより、要求される情報を保持するピアへクエリを転送することができる。この手法では応答の期待できるピアにのみクエリを転送するため、不必要なクエリの発生を抑えることができる。しかし、クエリは深さ優先探索で転送されるため、一度に複数のピアからコンテンツを取得したい場合には、応答時間が長くなるという問題がある。また、時間に依存するコンテンツの重要度は 1 時間単位や 1 週間単位で連続的に変化することが考えられ、離散的な分類ができない。RI はカテゴリに分類される情報を対象としており、そのままでは本研究の想定環境では適用できない。

## 3. 提 案 方 式

本章では、位置依存 P2P ネットワークを基盤とし、クエリ

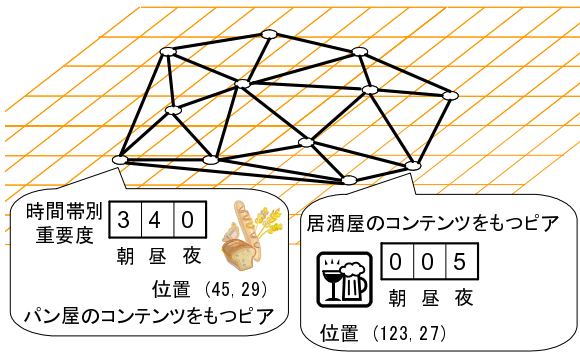


図1 保持するコンテンツの位置情報によるネットワークの例  
Fig.1 Example of network by location of contents that each peer keeps.

の検索時間帯によく適合するコンテンツを効率的に取得するためのクエリ転送方式について述べる。

### 3.1 想定環境

実際に地域コンテンツを検索するときには、位置情報以外に時間が重要となる。例えば、ユーザが「近くの飲食店」を探したい場合、近くの店のコンテンツ全てがユーザの意図する検索に適合するわけではない。ユーザが実際に訪れる時間に営業時間外の店舗のコンテンツは不必要である。また、昼も夜も営業している店舗では昼と夜で評判が違うこともある。このように、地域コンテンツの有用性は時間帯によって変化する。以降この有用性を重要度と定義する。検索するときには時間帯を指定することで、指定時間帯で重要度の高いコンテンツを選択することが可能となり、便利なサービスを実現できる。

地域コンテンツの検索サービスを実現する上で、コンテンツは店舗がキャンペーン広告のために作成したり、個人が時間を指定して不用品を引き取ってくれる相手を探すために作成するといった場合が考えられ、位置情報と時間帯別の重要度が含まれているものとする。以降その2つの情報をコンテンツの情報とする。

ピアはそれぞれ1つのコンテンツを保持する。図1のように、ピアは位置情報によってネットワークを構築する。ただし、この位置情報とは実空間上においてピアが存在する地理的位置ではなくコンテンツに付与された位置情報である。このため、例えばパン屋のコンテンツが店員の所持する携帯端末に存在するとき、店員が帰宅しても端末はパン屋の位置情報でネットワークに参加する。

図2に示すように、ネットワークのトポロジはドロネー図である。ネットワークの分断はなく、全体で1つの論理ネットワークを形成しているものとする。

リンクは双方向であり、両端のピアが各々相手のアドレスを保持している。リンク先のピアを隣接ピアとする。ピアは、ネットワークに参加している任意のピアとのリンクを構築できるものとし、通信範囲などによる制約はないものとする。またピアがコンテンツを検索する際には、位置と時間を範囲指定したクエリを送信する。クエリの要求内容に適合するコンテンツをもつピアは、送信元にレスポンスを送る。

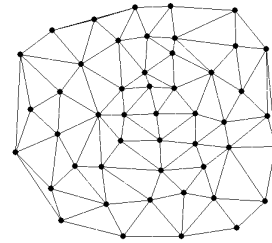


図2 ドロネー図  
Fig.2 Delaunay diagram

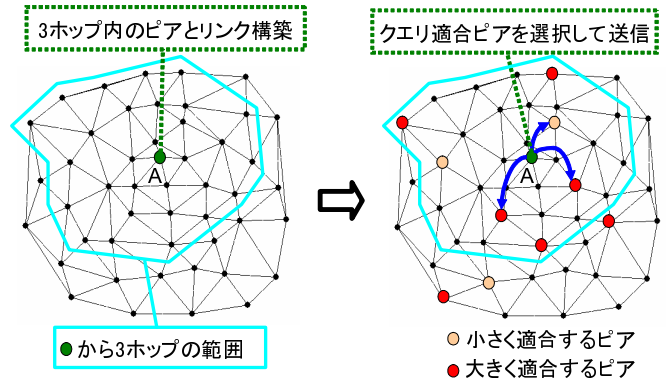


図3 リンクテーブルによるクエリの送信  
Fig.3 Query forwarding by link table

クエリは一意的識別子が付与されており、ピアは受信したクエリの識別子の履歴を作成することで、同じ識別子のクエリが重複して処理されない。また、クエリには経由したピアのアドレスが経路履歴として保持されており、経路履歴にあるピアにはクエリを送信しないことで、クエリのループを防止する。クエリには発行元のピアのアドレスが含まれており、1ホップでレスポンスはクエリ発行元のピアに届く。

### 3.2 リンクテーブル

位置ネットワークのトポロジであるドロネー図において、各ピアの平均リンク数は6である。ここで、クエリに適合するピアが存在しなければ、クエリの送信先をランダムに選択することになり検索効率が低下する。隣接するピアが全てクエリに適合しないときは、それらのピアを経由せずに2ホップ先、3ホップ先のピアへ転送できればトラフィックを削減できる。そのために、ピアは位置情報によるネットワーク上で $n$ ホップ先までのピアともリンクを構築する。その際、ピアは他ピアの保持するコンテンツの位置情報および時間帯別重要度をリンクテーブルに保持する。

図3では、ドロネー図のネットワークを示している。ピアAは、図で示した3ホップ以内のピアともリンクを構築しており、リンク先のピアがもつコンテンツの情報をリンクテーブルに保持している。ピアAがクエリを送信するときには、リンクテーブルとクエリの指定時間帯から送信先を選択して、検索効率を向上させている。ただし、 $n$ が大きくなるとテーブルサイズが増大するため、何らかの手段でテーブルサイズを限定しなければならない。ドロネー図のリンク数は、2ホップ先までのリン

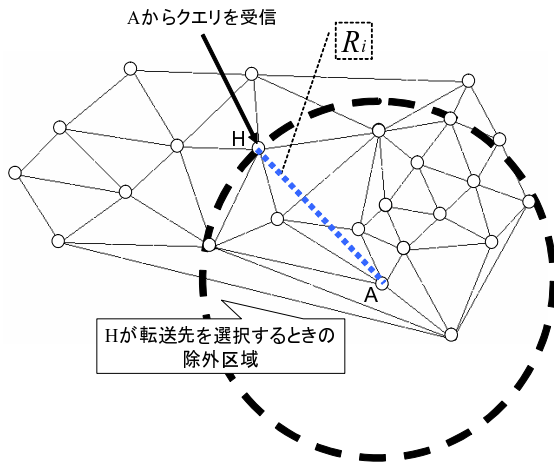


図 4 送信方向の制限

Fig. 4 Limit direction of query propagation

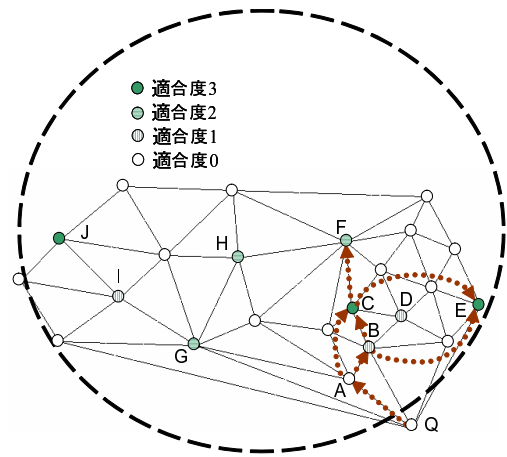


図 5 Distance 方式の転送例

Fig. 5 Example of query propagation in Distance method

ク数が平均 20, 3 ホップ先までのリンク数が平均 40, 4 ホップ先までのリンク数が平均 70 となる。テーブルサイズを現実的な大きさにすることと、トラフィック削減効果の向上とを併せて考慮し、本稿の評価では  $n = 3$  として、ピアは 3 ホップ内全てのコンテンツの情報を保持するものとした。

### 3.3 コンテンツの問合せに対する適合度

地域情報の重要度は時間帯によって変化する。例えば、パン屋のコンテンツは昼の重要度は品揃えが充実するために高くなり、夜は閉店するために重要度は 0 となっている。このようにコンテンツの時間帯別重要度とクエリの指定時間帯から、コンテンツのクエリに対する価値として適合度を定義する。適合度は以下の式で計算する。

$$F = \sum I_t \quad t \in \{c | Q_c = 1\} \quad (1)$$

ここで、 $F$  は適合度、 $I_t$  は時間帯  $t$  でのコンテンツの重要度である。クエリが時間帯  $c$  を指定時間帯にすると、 $Q_c$  は 1 となる。適合度を計算するための  $t$  はクエリの指定時間帯の集合である。つまり、適合度は、クエリの指定時間帯でのコンテンツの時間帯別重要度の総和となる。図 1 の例ではパン屋のコンテンツの時間帯別重要度が朝に 3, 昼に 4, 夜に 0 であるため、朝と昼を指定するクエリに対する適合度は 7 となる。

### 3.4 クエリ送信先の選択

クエリを受信したピアは、リンクテーブル内のコンテンツの情報から最も検索領域に近づくものを選び、そのコンテンツを保持するピアにクエリを送る。そして、クエリが検索領域内に到達すると複数リンクへのクエリ転送が開始される。以降これをセレクトティブフラッディングとする。ただし、クエリは検索領域内から検索領域外へ送られない。

同一ピアにクエリが複数送信されると不要なトラフィックが生じる。そこでクエリの送信方向の制限によって、同一クエリの重複受信を抑制する。各クエリには、セレクトティブフラッディング開始ピアの位置情報が付加される。クエリを受信したピアは、セレクトティブフラッディング開始地点に近づくリンクにはクエリを送信しない。その判断は、クエリ受信ピアとセレ

クティブフラッディング開始ピア間の位置情報による距離を半径  $R_i$  とし、セレクトティブフラッディング開始ピアを中心とする円内に位置するピアを除外することで行う。以降その円内の地域を除外区域とする。図 4 ではピア A がセレクトティブフラッディングを開始して、ピア H がクエリを受信している。このとき、 $R_i$  はピア A と H の距離となり、ピア H がクエリを転送するときの除外区域は点線で示した円内の地域となる。このように、クエリを受信したピアは除外区域に位置しないピアからクエリの転送先を選択する。

クエリを送信するピアは、リンクテーブルとクエリの内容を参照して、リンク先のピアがもつコンテンツの適合度を計算できる。しかし、適合度が 0 のピアにクエリを送信しただけでは、トラフィックの削減効果が期待できない。送信クエリ数を限定し、適合度が高いピアの上位に送信することでトラフィックは削減できるが、高い適合度のピアにクエリが集中して取得できるレスポンス数が低下する。そこで、適合度を基準として、より多くのレスポンスを取得できるように送信先を選択するためのスコアを定義する。

クエリの転送先の選択のためのスコア計算は、ピアのもつコンテンツの位置情報による距離を用いる計算式を Distance 方式、ドロネー図におけるピア間のホップ数を用いる計算式を Hop 方式として、2 つの方式を比較する。それぞれの計算式について以下で述べる。

Distance 方式において、スコアはクエリに対する適合度と、クエリ受信ピアと送信候補ピア間の距離から以下の式で算出する。

$$S = \frac{F}{d} \quad (2)$$

ここで  $S$  は送信候補ピアのスコア、 $F$  はクエリに対するコンテンツの適合度、 $d$  はクエリ受信ピアと送信候補ピアの距離を示している。距離はピアのもつコンテンツの位置情報から決定される。このスコア計算式から、同じ適合度をもつリンクが複数ある場合は、距離に近いものが優先される。これは、リンク先ピアの適合度が全て同じ場合、距離に近いものにクエリを送信しなければ、除外区域に位置する必要なコンテンツをもつピ



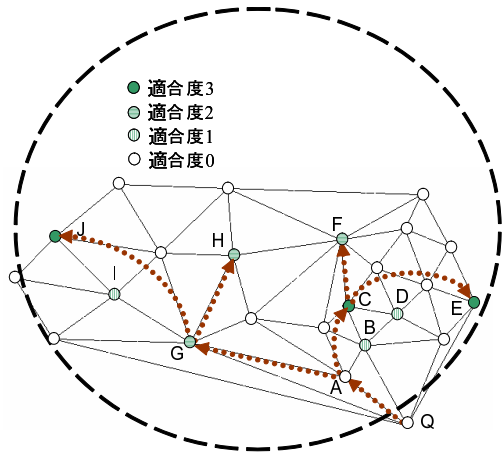


図 6 Hop 方式の転送例  
Fig. 6 Example of query propagation in Hop method

アにクエリが届かなくなるためである。

このスコア計算式では、距離の遠近による影響が大きい。図 5 では、ピア A がピア Q から点線の円の地域を探索範囲とするクエリを受信して、セレクティブフラッディングを開始する場合であり、上位 2 個のピアに送信するものとする。ピア C, E, J のもつコンテンツのクエリに対する適合度が 3、ピア F, G, H が適合度 2、ピア B, D, I が適合度 1 とする。このとき、A は式 2 によって 3 ホップ以内のリンク先ピアのスコアを計算して、B, C にクエリを送信する。また、A が送信したクエリを受信したピアも、それぞれ同様にスコアを計算してクエリを送信する。このように、コンテンツ間の距離を用いるとコンテンツ位置の分布密度による影響が大きく、分布密度の高い地域にクエリが転送されやすい。そのため探索領域内で、特定の地域にクエリが集中することがある。

スコア計算において、クエリ受信ピアと送信候補ピア間の距離ではなく、ホップ数を用いると距離のスコアに対する影響を限定できる。そこで、

$$S = \frac{F}{h} \quad (3)$$

の式を Hop 方式として検討した。ここで、 $h$  はクエリ受信ピアから送信候補ピアまでのホップ数である。ピアは位置情報によるドロネー図において 3 ホップ先までとリンクするため、スコア計算時に適合度を割る数は 1, 2, 3 のいずれかに限定される。そのため、位置情報の影響が小さくなると考えられる。図 6 では、図 5 と同様の状態のときに、スコア計算式 (3) としている。A はリンク先ピアのスコアを計算して、ピア C, G にクエリを送信している。また、クエリを受信したピアもそれぞれ同様にスコアを計算してクエリを送信する。このように、Hop 方式では距離の影響が小さくなり、クエリの探索領域内でピアの分布する位置に偏りがあっても、クエリが均等に広がる。

クエリを送信するピアは、リンク先のピアそれぞれのクエリに対するスコアを算出した上で、スコアの上位  $s$  個のピアにクエリを送信する。このとき、適合度が 0 ではない送信可能ピアが 2 つしかなければ  $s = 3$  でも 2 ピアにしか送信しない。クエリに対するテーブル内のコンテンツの適合度が全て 0 であ

表 1 評価環境の設定値

Table 1 Simulation parameters

設定項目	設定値
ピア数	400
適合度別のコンテンツ数	一様
エリアサイズ	1000 平方
クエリ数	400
探索領域	エリア全体
TTL	100

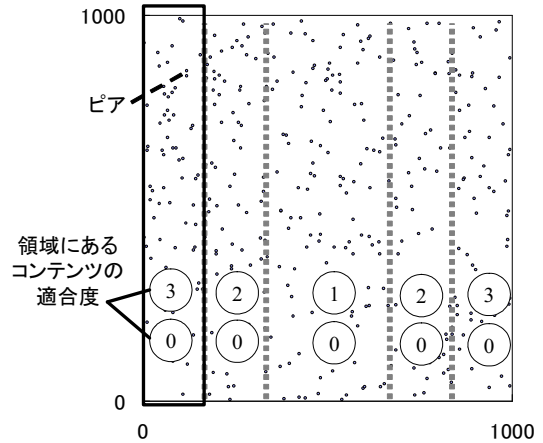


図 7 評価環境 B の領域における適合度の分布  
Fig. 7 Distribution of adaptation in area under environment B

ば、テーブル内からクエリの経路履歴に含まれていないピアをランダムに 1 つ選択してクエリを送信する。テーブル内に送信可能なピアがなければ、クエリの転送は終了する。

#### 4. 評価

本章では、提案手法の評価のために行ったシミュレーション実験の評価環境と評価結果を示す。

##### 4.1 評価環境

シミュレーション評価における評価環境の共通設定を表 1 に示す。シミュレーション領域を 1000 平方の地域として、1 日は設定を単純にするために 10 のタイムスロットとする。地域コンテンツの数の参考のために、Yahoo! エリア検索 [4] において「北千里駅から半径 0.5km」で検索したところ、検索結果は 200 件程度であった。想定するサービスでは、個人によるコンテンツの作成も考えられるため、ピア数は 2 倍の 400 と設定した。

ピアはそれぞれ 1 つのコンテンツをもち、コンテンツは位置情報と時間帯別重要度が設定されている。なお、あるクエリに対して適合度が 1 以上となるコンテンツ数を一様にするために、各時間帯における重要度は 0 か 1 のみとした。

実際の地域コンテンツの分布では、時間帯別重要度の偏りが考えられる。例えば、繁華街に沿った店舗は居酒屋などが多く、夕方から深夜にかけて重要度が高くなるコンテンツが集まることがありうる。このとき、夜を指定するクエリに対する適合度の分布は、住宅街には低いものが集中し繁華街には高いものが集中する。

検索領域内コンテンツのクエリに対する適合度の分布の偏りによる影響を評価するため、評価環境 A, B で提案手法を評価する。環境 A では、発生するクエリに対して適合するコンテンツの分布は一様である。環境 B では、シミュレーションで発生するクエリの探索時間を固定してコンテンツの重要度設定を調整することで、クエリに対する適合度の分布を図 7 のように設定した。図 7 ではエリアを縦に分割して、クエリに高い適合度のコンテンツをもつピアが両端に分布している。両環境で、適合度が 1 以上となるコンテンツは一様に存在しており、環境 A と比較することで適合度の分布による影響を評価する。

初期状態で、ピアは保持するコンテンツの位置情報によってドロネー図となるトポロジを構成する。その後、提案手法ではピアが 3 ホップ以内の他ピアとリンクを構築して、リンク先のピアがもつコンテンツの情報をテーブルに保持する。クエリは各ピアが 1 つ発行する。クエリの探索地域におけるクエリの転送のみを評価するため、クエリの探索地域はエリア全体とする。

- コンテンツの時間帯別重要度とクエリの設定

環境 A において、重要度は時刻 0 から 9 のうち、ランダムに選択した連続する 4 つの時間帯で設定されている。クエリはランダムに連続した 3 つの時間帯を指定して検索する。

環境 B において、重要度は時間帯 0 から 2 のうち、連続する 0~3 つの時間帯で設定されている。クエリは全ての時間帯を指定して検索する。

環境 A, B で共通して、重要度の設定値は選択された時間帯で 1, ほかの時間帯では 0 となる。コンテンツの重要度とクエリの設定から、クエリに対するコンテンツの適合度は 0, 1, 2, 3 のいずれかである。あるクエリに対して適合度が 1 以上のコンテンツの数が一樣なことで、取得したレスポンス数の適合度別の差が提案手法によるものと判断できる。

以上のようなシステム環境において、Distance 方式, Hop 方式のそれぞれで性能評価を行う。

#### 4.2 評価項目

提案手法について 3 種類の評価を行う。比較手法としてフラッディングを用いる。通常、フラッディングではクエリを受信したピアは、隣接したピア全てにクエリを転送する。しかし、シミュレーション環境においてピアは位置情報による構造型のネットワークを構築しているため、通常フラッディングでは同一ピアがクエリを複数受信することが多くなり効率が悪い。そこで、比較対象のフラッディングには、提案手法と同様に送信方向の制限を設けた。この制限のために、グラフに示しているフラッディングの結果は通常フラッディングと比べてトラフィックは 45%削減されているものである。制限を設けても領域内のピア全てにクエリが到達するため、適合するコンテンツを全て取得でき、適合度別のリコール率は全て 1 となっている。評価項目は以下の通りである。

評価 1 ネットワークトラフィックの変化

評価 2 レスポンスのクエリ適合度に応じたりコール率

評価 3 トラフィックの削減効果に伴うリコール率への影響

評価 1 では、論理ネットワークにおいて提案手法を適用し、トラフィックの変化を確認する。クエリあたりの平均ホップ数

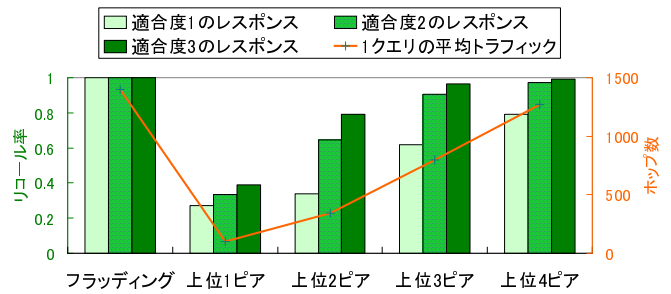


図 8 環境 A: Distance 方式のレスポンスの適合度別リコール率  
Fig. 8 Environment A: Comparison of recall ratios and the number of messages by adaptation of responses (Distance method)

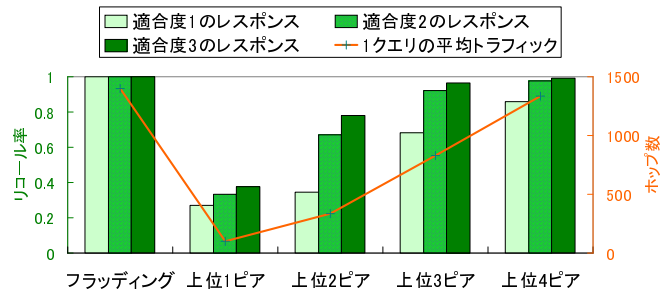


図 9 環境 A: Hop 方式のレスポンスの適合度別リコール率  
Fig. 9 Environment A: Comparison of recall ratios and the number of messages by adaptation of responses (Hop method)

を基準とする。

評価 2 では、論理ネットワークにおいて、取得したレスポンスの適合度ごとのリコール率 (取得率) を確認する。

評価 3 では、提案手法によって削減したトラフィックとリコール率の関係を確認する。ここで、提案手法の有効性を議論するために、定量的な比較のための指標値が必要である。評価指標は、フラッディングと比較したトラフィック減少率と適合度上位のリコール減少率から定義する。指標値は、トラフィック減少率/適合度上位 2 つのリコール率の減少値で求まる値とする。この値は、リコール率が 1%減少したことでトラフィックが何%減少したかを示す。この指標値が高いほど、トラフィックの削減効果にともなうリコール率の減少が抑制されている。

以下ではシミュレーション実験の結果を示し、環境ごとの評価項目の結果を示す。

#### 4.3 評価結果

環境 A, B における実験結果の評価項目 1, 2 の結果のグラフを、方式ごとに図 8 から 11 に示す。環境 A では、Distance 方式と Hop 方式の差はほとんど生じていない。クエリの送信先  $s$  を増加させていくことで、リコール率とトラフィックは共に増加しているが、例えば Distance 方式において  $s = 1$  から 2 に変化させたときにトラフィックが 3 倍以上になったにも関わらずリコール率は 2 倍程度しか増加していない。このことから、複数のリンクにクエリを送信するときには、クエリを重複して受信するピアが生じていると確認できる。これはクエリを受信したピアはそれぞれが独立して転送先を選択することが原

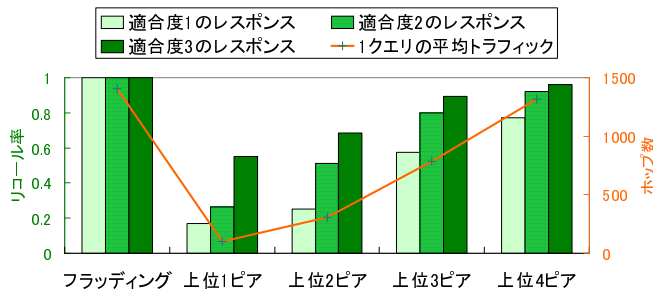


図 10 環境 B: Distance 方式のレスポンスの適合度別リコール率

Fig. 10 Environment B: Comparison of recall ratios and the number of messages by adaptation of responses (Distance method)

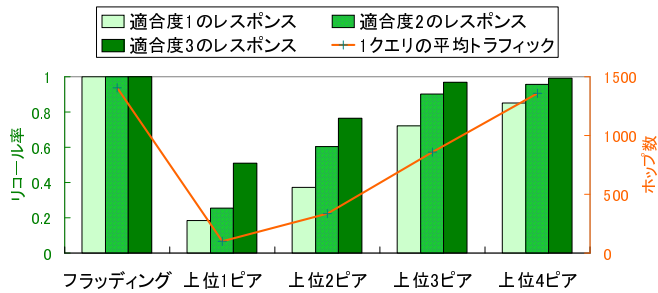


図 11 環境 B: Hop 方式のレスポンスの適合度別リコール率

Fig. 11 Environment B: Comparison of recall ratios and the number of messages by adaptation of responses (Hop method)

因である。しかし、フラッディングと比較するとトラフィックの削減効果は明らかであり提案手法の有効性は確認できる。

環境 B では適合度 3 のコンテンツは両端に分布しており、取得しにくい設定を意図した。しかし、環境 A と比較すると、 $s = 1$  としたときの適合度 3 のリコール率は Distance 方式で 16%、Hop 方式で 13% 増加している。これは、ドローネー図で外周のピアは遠距離のピアとリンクしやすく、領域の端に到達したクエリは外周ピアを経由して反対側へ少ないホップ数で到達するためである。こうして、クエリは中央のピアよりも両端に位置するピアへ送られることが多くなり、グラフのような結果になったと考えられる。環境 A と比較して、複数ピアへクエリを送信するときの Distance 方式のリコール率は低下している。これは、複数のピアへクエリを送信するとクエリの経路履歴が異なるものとなり、クエリを受信したピアが既に受信済みのピアを転送先を選択することが原因である。適合するピアが密集した地域では、位置的に近いピアのスコアが高くなると、密集した地域内で多くの重複受信が生じて地域外へクエリが広がりやすくなること分かる。Hop 方式では環境 A と比較して、グラフからはトラフィックとリコール率に大きな変化が見られない。

環境 A, B の結果の指標値のグラフを図 12, 13 に示す。環境 A では、両方式間の差は小さいが、Hop 方式で  $s = 3$  とするときの指標値が最大となっている。環境 B では、Distance 方式の指標値が低下しており、 $s = 3$  における指標値は環境 A から半分以下になっている。この低下は、クエリが検索地域の端へ

送られやすい環境 B では領域全体へクエリが配送されにくいために生じる。一方、Hop 方式では  $s = 3$  としたときの指標値の低下は 1.4 程度に抑えられている。この結果から探索領域全体へクエリが広がりやすい Hop 方式の有効性が確認できる。

それぞれの環境におけるシミュレーション結果のグラフから、提案手法では適合度の高いレスポンスほど高いリコール率を維持できていることが分かる。クエリを送信する数を増やすと、リコール率とトラフィックはともに増加するが、上位 4 ピアに送るとき提案手法はトラフィックがフラッディングに近いにも関わらずリコール率がフラッディングを下回っている。これは提案手法では 3 ホップ先までクエリを送信するため、クエリの経路履歴の違いによってはクエリを受信済みのピアに別の経路履歴をもつクエリが到達する。そのために、フラッディングよりトラフィックが増加することがある。また、ピアがクエリを送信する際に、近くのピアをスキップすることで、除外区域による影響から、クエリが到達できないピアが生じる。よって、提案手法では、送信先ピア数の選択によっては性能がフラッディングより悪化する問題があるとわかる。この問題は、ピアがクエリを転送する際に、転送先からの経路を予測することで緩和できる可能性がある。例えば図 5 の転送例において、ピア A はスコアの上位ピア B, C に送信している。このとき、ピア A が B に送信すれば B が C に転送すると予測して C に送信しないことで、トラフィックを削減できる。トラフィックをさらに削減するためのクエリの経路を考慮した転送手法は、今後の課題である。

## 5. 提案方式の拡張

本稿では、コンテンツの重要度を時間帯別に設定してシミュレーションを行った。しかし実際の地域コンテンツでは、1 日周期の変化だけではなく、週末になるほど重要度が高くなる、月末には重要度が高くなる、といった違う周期での変化が存在する。さらに、ある時点を境に突然重要度が変化するのではなく、連続的な変化を考慮する必要もある。時間帯を限定せず、より連続的な視点で考察することが今後の課題である。

また、本稿ではコンテンツベースでのネットワークとして考えたが、今後ピアが複数の情報をもつことを考慮した手法に拡張する必要がある。ピアが複数のコンテンツをもつときには、各コンテンツ毎にリンクテーブルをもち仮想的に複数のピアとして参加するのが単純であるが、リンクテーブルを集約して利用することが理想である。ピアがもつ情報量を抑えて有効なルーティングをするために、どのような情報をもつべきなのかについて今後検証する予定である。

## 6. ま と め

本稿では、コンテンツの位置と時間帯別の重要度に着目したリンクの構築とクエリのルーティング手法を提案した。またシミュレーション実験による評価を行った。その結果から、リンク先のピアがもつコンテンツの情報をもつことで、価値の高いレスポンスのリコール率を上げつつトラフィックの削減を行うことに成功した。しかし、クエリのルーティング手法について

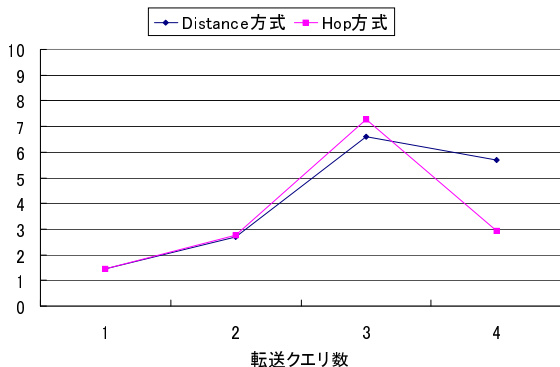


図 12 評価環境 A の手法別指標値

Fig. 12 Environment A: index of each method

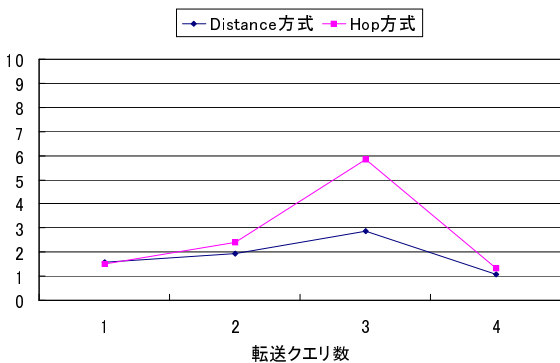


図 13 評価環境 B の手法別指標値

Fig. 13 Environment B: index of each method

はまだ改善の余地がある．今後は，複数コンテンツをもつためのテーブルの集約と重要度の連続性を考慮した手法への拡張を行う予定である．

## 謝 辞

本研究は，平成 18 年度総務省委託研究「ユビキタスネットワーク認証・エージェント技術の研究開発」の一環として実施したものである．ここに記して謝意を表す．

## 文 献

- [1] F. Araujo and L. Rodrigues, "GeoPeer: A location-aware peer-to-peer system," in *Proc. The 3rd IEEE International Symposium on Network Computing and Applications (NCA '04)*, pp. 39-46, Sept. 2004.
- [2] A. Crespo and H. Garcia-Molina: "Routing indices for peer-to-peer systems," in *Proc. the 22nd International Conference on Distributed Computing Systems 2002*.
- [3] 金子 雄, 春本 要, 福村真哉, 下條真司, 西尾章治郎: "ユビキタス環境における端末の位置情報に基づく P2P ネットワーク," 情報処理学会論文誌: データベース, vol. 46, no. SIG18 (TOD 28), pp.1-15.
- [4] Yahoo Japan Corporation, "Yahoo! エリア検索," <URL: <http://area-search.yahoo.co.jp>>.