

特別研究報告

題目

オーバレイネットワークの管理に必要な資源情報を得るための
スケーラブルな計測手法

指導教員

中野 博隆 教授

報告者

森 弘樹

平成 21 年 2 月 16 日

大阪大学 基礎工学部 情報科学科

オーバーレイネットワークの管理に必要な資源情報を得るための
スケーラブルな計測手法

森 弘樹

内容梗概

オーバーレイネットワークは IP ネットワーク上に論理的に構築されたネットワークであるため、性能の維持、向上のためには定期的にオーバーレイパスの資源情報を計測によって得る必要がある。オーバーレイネットワークの構築に必要な情報を得る手法は数多く提案されているが、その多くは小規模なオーバーレイネットワークを対象としており、全てのオーバーレイノード間の経路を計測する手法である。オーバーレイノード数を n とすると、このような手法では $O(n^2)$ のコストが必要であり、オーバーレイノード数が増加した場合には計測に必要なコストの増加が問題となる。

そこで本報告は、IP ネットワークにおけるルータ数に対するオーバーレイノードの割合（オーバーレイノード密度）に対しコストがスケーラブルな計測手法を提案する。具体的には、オーバーレイパス内に他のオーバーレイノードが含まれており、オーバーレイパスが複数のホップ数の小さなオーバーレイパスで構成されている場合に、それらの部分オーバーレイパスの計測結果を用いて資源情報を推測することにより、計測パス数の削減を行う。本手法は、オーバーレイノード密度が増加した場合、このような経路の重複が増加することを利用し、オーバーレイノード密度に対するスケーラビリティを実現している。数学的解析手法、およびさまざまなトポロジに対するシミュレーションによって計測パス数の評価を行った結果、提案手法を用いた場合の計測パス数をフルメッシュ計測と比較して最高で約 1/4000 程度に削減できることを明らかにする。また、シミュレーションよりオーバーレイパスの資源情報の推測に必要な情報交換数の評価を行う。さらに、ネットワークにおけるノード次数の分布を利用したオーバーレイノードの設置戦略を提案し、計測パス数をさらに約 1/5 程度に削減できることを示す。

目次

1	はじめに	5
2	オーバーレイネットワークにおける計測	7
2.1	ルータオーバーレイネットワーク	7
2.2	資源情報の計測における問題点	7
3	計測パス数の削減手法	10
3.1	提案手法	10
3.2	資源情報の推測	11
4	性能評価	13
4.1	ネットワークモデル	13
4.2	計測パス数の解析的評価	13
4.3	シミュレーション結果および考察	14
4.3.1	計測パス数	14
4.3.2	情報交換回数	15
5	計測パス数を削減するためのオーバーレイノードの設置戦略	26
5.1	ハブ優先設置法の解析的評価	26
5.2	シミュレーション評価	29
6	まとめと今後の課題	36
	謝辞	37
	参考文献	38

目次

1	オーバーレイネットワークモデル	8
2	オーバーレイノード密度	8
3	提案方式	11
4	BA トポロジにおける計測パス数	17
5	BA トポロジにおける式 (2) の評価結果	18
6	ランダムトポロジにおける計測パス数	19
7	ランダムトポロジにおける式 (2) の評価結果	20
8	Abilene-inspired トポロジにおける計測パス数	21
9	Abilene-inspired トポロジにおける式 (2) の評価結果	21
10	AT&T トポロジにおける計測パス数	22
11	AT&T トポロジにおける式 (2) の評価結果	22
12	トポロジの違いが与える影響	23
13	BA トポロジにおける資源情報の送信数	24
14	ランダムトポロジにおける資源情報の送信数	25
15	ハブ優先設置法の性能評価	31
16	シミュレーション結果と式 (4) との比較	32
17	シミュレーション結果と式 (7) との比較	33
18	ハブ優先設置法における式 (8) の評価	34
19	BA トポロジにおけるハブのノード次数の総和	35

表目次

1	α の変化	27
---	------------------------	----

1 はじめに

近年、コンピュータやインターネットの普及、ユーザの増加により、P2Pアプリケーションなどにおいてオーバーレイネットワークと呼ばれる、IPネットワーク（アンダーレイネットワーク）上に独自に構築された論理的なネットワークが利用されるようになった。図1(a)にアンダーレイネットワークとオーバーレイネットワークの関係を示す。図に示すように、オーバーレイネットワークはIPネットワーク上に論理的に構築されるため、オーバーレイパス（オーバーレイノード間のデータ転送パス）がIPネットワーク上で通過している経路情報などは一般的には不明である。しかし、オーバーレイネットワークの品質は、アンダーレイネットワークであるIPネットワークの品質に大きく影響を受けるため、オーバーレイパスの資源や品質に関する情報を計測によって取得することは、オーバーレイネットワークの性能を維持および向上させる上で非常に重要である。

オーバーレイパスのネットワーク資源を計測によって得る手法に関する研究は数多く行われている [1–3] が、その多くはオーバーレイノード数が高々数十程度の小規模なオーバーレイネットワークを対象としており、計測が必要なオーバーレイパスの計測をフルメッシュ、すなわち、全てのオーバーレイノード間のパスの計測を等しく行うものである。フルメッシュな計測手法は小規模なオーバーレイネットワークにおいては短時間で必要な情報を得られるために有効であるが、オーバーレイノード数を n とすると、計測に $O(n^2)$ のコストが必要となるため、オーバーレイノード数が増加した場合は計測のためのオーバーヘッドが問題となる。例えば Resilient Overlay Networks (RON) [4] においては、オーバーレイノードがフルメッシュに計測を行い、計測結果を全てのノード間で交換することによって、オーバーレイネットワークにおいて経路制御を行うことが提案されている。しかし、参加可能なオーバーレイノード数が50程度に制限されることが指摘されている [5]。また、NICE [6] および HMTP [7] は計測を行うオーバーレイパスを限定することによって計測のオーバーヘッドを $O(n)$ に抑えている。しかし、これらの手法は計測のコストの削減と引き換えに計測範囲を犠牲にしているため、最適なオーバーレイパスが発見できない可能性がある。また、伝播遅延時間のみの計測であればルータ数に対してスケラブルな計測手法が提案されている [8–11]。しかし、これらの手法ではネットワーク帯域やパケット廃棄率などの情報を得ることができない。そのため、十分な計測情報量を維持しつつ、大規模なオーバーレイネットワークにも適用可能な計測手法が必要である。

そこで本報告では、アンダーレイネットワークのIPルータ数に対するオーバーレイノードの割合（オーバーレイノード密度）を考慮した計測手法の提案を行う。提案方式は、密度が高くなったときに多発するオーバーレイパス同士の経路の重複に注目し、大きなホップ数のオーバーレイパスの資源情報をホップ数の小さな複数のオーバーレイパスの資源情報から推測することにより、計測が必要なオーバーレイパス数（以下、計測パス数と呼ぶ）を削減する。また、

提案方式における計測パス数をシミュレーションによって評価することで、提案手法がオーバーレイノード密度に対しスケラブルであることを示す。さらに、計測パス数を数学的解析手法によって導出することによって、提案手法の定性的な性質を明らかにする。また、提案方式が必要とするオーバーレイノード間の情報交換回数を評価し、そのオーバーヘッドがフルメッシュな計測手法と比較して極めて小さいことを示す。

本報告の構成は以下のとおりである。2章では対象とするオーバーレイネットワークモデル、およびオーバーレイノード密度の変化がネットワーク計測へ与える影響に対する説明を行う。3章では提案手法の説明を行い、4章で様々なアンダーレイネットワークに対するシミュレーションを通じて提案手法の評価を行う。5章では計測コストを低減するためのオーバーレイノードの設置場所に関する検討を行なう。最後に6章で本報告のまとめと今後の課題を述べる。

2 オーバレイネットワークにおける計測

2.1 ルータオーバレイネットワーク

本報告ではオーバレイネットワークを IP ネットワーク上に構築した論理的なネットワークと定義する。既存のオーバレイネットワークを利用したサービスとしては Skype [12]、BitTorrent [13]、Veoh [14] などがあるが、それらのサービスは図 1(a) に示すように、IP ネットワークに接続されたエンドホストで動作しているアプリケーション同士が接続することによってオーバレイネットワークが構築される。このようなオーバレイネットワークを本報告においてはエンドホストオーバレイネットワークと呼ぶ。この手法では、オーバレイネットワーク上の全ての通信がエンドホスト間の通信となるため、オーバレイ層におけるルーティング制御であるオーバレイルーティングなどを効果的に行うことができない。一方、図 1(b) に、IP ルータにオーバレイノードとしての機能を搭載することでオーバレイネットワークに参加させる、ルータオーバレイネットワークの例を示す。このように、ネットワーク内のルータ上にオーバレイノードが存在することによって、効率的なオーバレイトラフィックの制御が可能となる。本報告では、このようなルータオーバレイネットワークを前提とし、計測パス数の削減手法を提案する。

2.2 資源情報の計測における問題点

図 2 にルータオーバレイネットワークにおけるオーバレイノード密度が低い場合 (図 2(a)) と高い場合 (図 2(b)) のネットワークトポロジの例を示す。オーバレイノード密度とは IP ルータ数に対するオーバレイノードの割合であり、例えば図 2(b) の場合には $11/15 \simeq 0.73$ である。

オーバレイパスが通過するネットワークの資源情報 (帯域幅、伝播遅延時間、パケット廃棄率など) の計測を行う際には、オーバレイノード密度によって計測の特性が大きく変化する。図 2(a) のようにオーバレイノード密度が低い場合には、オーバレイネットワークは以下のような特徴を持つ。

- 2つのオーバレイノード間のパス上に他のオーバレイノードが存在する確率は低い
- 複数のオーバレイパスの経路が重複している可能性は低い
- 平均経路長が大きいため、各パスのネットワーク計測の精度は低い

このような環境においては、複数のオーバレイパスの計測を同時に実行しても、計測が重複することが少ない。そのため、計測精度を低下させることなく、同時に複数のオーバレイパ

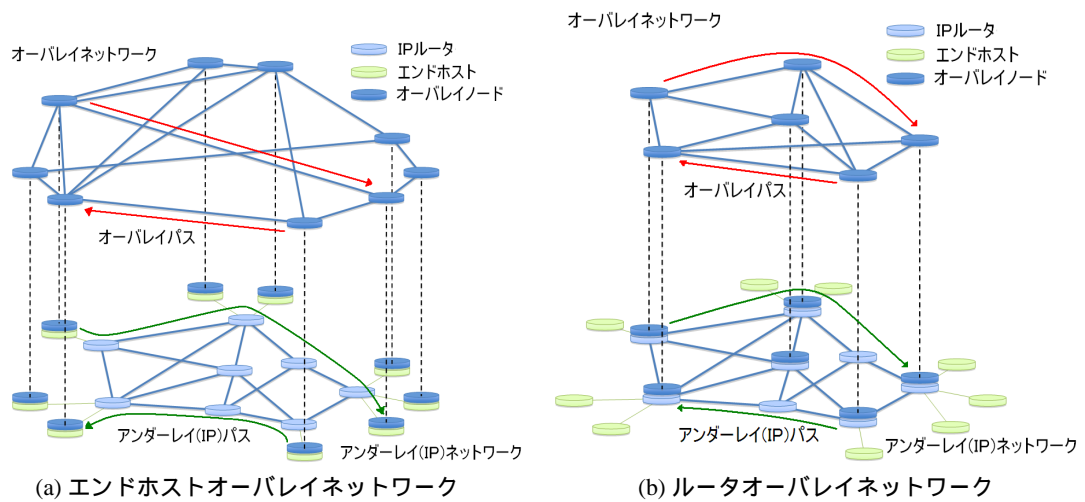


図 1: オーレイネットワークモデル

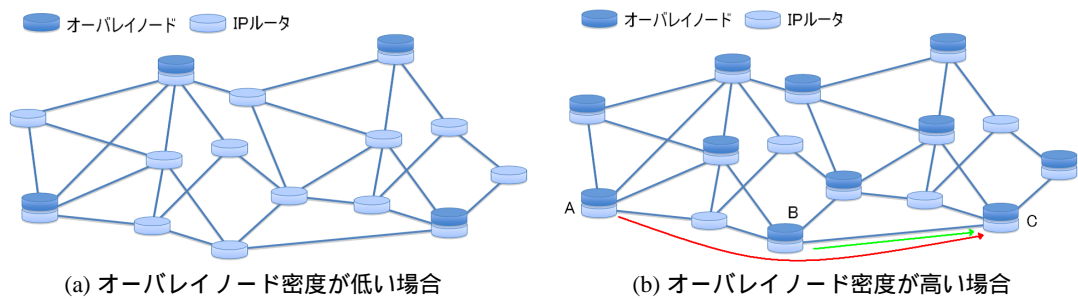


図 2: オーレイノード密度

スの計測を行うことが可能となる。また、一度の計測に用いる時間を長く確保することが可能となるため、計測精度を補うことができる。

一方、図 2(b) のようにオーレイノード密度の高いオーレイネットワークは以下のような特徴を持つ。

- 2つのオーレイノード間のパス上に他のオーレイノードが存在する可能性が高い
- 複数のオーレイパスの経路が重複している可能性は高い
- 平均経路長が小さいため、計測の重複が発生しなければネットワーク計測の精度は高い

このような環境においては、経路の重複が多く、共有された IP パスでは計測トラヒックによる負荷が増大する。また、このことは計測精度の低下を引き起こす。例えば、図 2(b)において、オーレイノード A、B がそれぞれオーレイパス AC、BC の計測を同時に行う場

合、オーバーレイパス BC の IP リンクにおいて両オーバーレイノードの計測トラヒックが互いに干渉することによって、計測結果に誤差が生じる。

オーバーレイネットワークにおける資源情報の計測の衝突に関する研究としては [1, 15] などがある。例えば、[15] ではネットワーク資源の計測をスケジューリングすることによって、計測の衝突を防ぐ手法を提案している。しかし、これらの手法は各オーバーレイノードがアンダーレイネットワークのトポロジの情報を得られることを前提としている。アンダーレイネットワークのルータ数が少ない場合にはこれら手法は有効であるが、アンダーレイネットワークのルータ数が多くなると、トポロジ情報を得るためのオーバーヘッドが問題となる。また、オーバーレイノード数が増加した場合においても収集すべき情報量がノード数の二乗に比例するという問題がある。

ここで、資源情報の計測の衝突は、ホップ数の小さいオーバーレイパス AB および BC のみを計測し、その情報を元にホップ数の大きなオーバーレイパス AC の資源情報の推測を行うことにより回避可能であると考えられる。そこで本報告では、オーバーレイノード密度が高くなるほどこのようなオーバーレイパスの重複が発生する可能性が高くなることを利用して、経路の重複による計測精度の低下を防ぎつつ、計測パス数の削減を行う手法の提案を行う。

3 計測パス数の削減手法

本章では 2.2 節で説明したように、ホップ数の大きなオーバレイパスがホップ数の小さな複数のオーバレイパスから構成されている場合に計測パス数の削減を行う手法を提案する。なお、各オーバレイノードは他の全てのオーバレイノードの IP アドレスを知っているものとし、他のオーバレイノードに対して traceroute コマンドを実行できるものとする。また、自ノードを通過する traceroute パケットをキャプチャすることにより、traceroute コマンドの送信ノード、受信ノードの IP アドレスを記録できるものとする。

3.1 提案手法

ここでは、図 3 におけるオーバレイノード A が、オーバレイパス AD の資源情報を取得する状況を想定し、提案手法の動作を説明する。提案方式は各ノードが独立に実行することができる。

- ステップ 1: 他のオーバレイノードへ traceroute コマンドを実行
送信ノード A が受信ノード D に対して traceroute コマンドを実行する。
- ステップ 2: traceroute パケットをキャプチャ
他のオーバレイノードからの traceroute コマンドを受け取ったオーバレイノード B および C は traceroute の発信元、送信先を記録する。
- ステップ 3: 経路の重複の検出
オーバレイノード A、B、および C は traceroute コマンドおよび、自ノードを通過した traceroute パケットのキャプチャ結果から、オーバレイパスの重複を検出する。
- ステップ 4: 計測を行うオーバレイパスの決定
オーバレイパス中に他のオーバレイノードが存在しない場合は通常の計測を行う。
- ステップ 5: オーバレイパスの資源情報の推測
オーバレイパス中に他のオーバレイノード（図 3 の例では B、C）が存在している場合には計測を行わず、ステップ 4 の結果と、他のオーバレイノードとの情報交換によって得られた情報を基にオーバレイパスの資源情報の推測を行う。まず、オーバレイノード B はオーバレイパス BC の計測結果とオーバレイノード C から受け取ったオーバレイパス CD の資源情報からオーバレイパス BD の資源情報の推測を行い、その結果をオーバレイノード A へ送信する。オーバレイノード A は自らが計測を行ったオーバレイパス AB の資源情報とオーバレイノード B から受け取ったオーバレイパス BD の資源情報からオーバレイパス AD の資源情報の推測を行う。

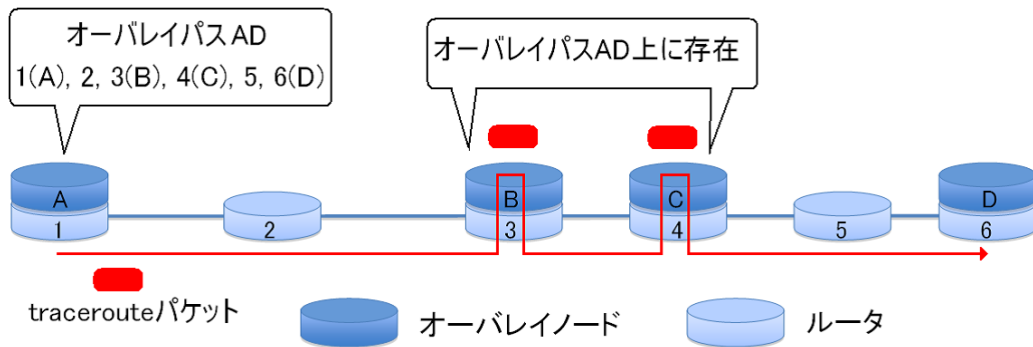


図 3: 提案方式

3.2 資源情報の推測

図 3 のようにオーバーレイパス (AD) 内に他のオーバーレイパス (AB、BC、CD) が含まれる場合における資源情報の推測方法を説明する。オーバーレイノード i, j 間の帯域、伝播遅延時間、パケット廃棄率をそれぞれ w_{ij} 、 l_{ij} 、 p_{ij} とすると、下記のようにオーバーレイパス AD の資源情報をそれぞれ推定することができる。

$$\text{帯域 } w_{AD} = \min(w_{AB}, w_{BC}, w_{CD})$$

$$\text{伝播遅延時間 } l_{AD} = l_{AB} + l_{BC} + l_{CD}$$

$$\text{パケット廃棄率 } p_{AD} = 1 - (1 - p_{AB})(1 - p_{BC})(1 - p_{CD})$$

このように推測を行うことにより、図 3 の状況においては本来 12 回必要であった計測を 6 回に削減することが可能である。また、オーバーレイパス AD の計測を行わないことにより、計測の重複による計測精度の低下を防ぐことが可能である。

提案手法を用いてオーバーレイパスの資源情報を推測するためには、資源情報をオーバーレイノード間で交換することが必要である。資源情報の送信回数は資源情報の推測を行うオーバーレイパス数と等しく、(フルメッシュな場合の計測パス数) - (提案手法で計測を行うオーバーレイパス数) で表される。しかし、オーバーレイノードへの資源情報の送信をできる限りまとめて行うことにより、資源情報の送信回数を大幅に削減させることができる。具体的には図 3 において、オーバーレイノード B はオーバーレイノード A がオーバーレイパス AD の資源情報の推測に用いる、オーバーレイパス BD の資源情報を送信する。同時に、オーバーレイノード A はオーバーレイパス AC の資源情報を推測するために、オーバーレイパス BC の資源情報も必要としている。この場合にオーバーレイノード B はオーバーレイパス BC、BD の資源情報をオーバーレイノード A へまとめて送信することによって資源情報の交換回数を削減することがで

きる。

4 性能評価

本章では提案手法を用いた場合にフルメッシュ的な計測手法と比較して、どの程度の計測パス数の削減効果があるかの評価を行う。

4.1 ネットワークモデル

シミュレーションにおいては、以下のようなネットワークトポロジをアンダーレイネットワークとして利用した。

BA トポロジ [16] において提案された生成規則を用いたトポロジであり、ノード次数分布がべき乗則に従うという特徴を持つ。BA トポロジはトポロジジェネレータ BRITE [17] を用いて作成する。ルータ数は 171、1000、10000 とし、リンク数のルータ数に対する倍率 m を 1、2、4 とする。

ランダムトポロジ Waxman モデル [18] に基づくトポロジであり、ランダムネットワークとも呼ばれる。BA トポロジと同様、ランダムトポロジは BRITE を用いて作成する。ルータ数およびリンク数は BA トポロジと同様である。

Abilene-inspired トポロジ [19] において示された米国の学術ネットワーク Abilene に似たトポロジを持つネットワークであり、ノード次数分布がべき乗則に従う。コアルータ、ミドルルータ、エッジルータから構成されており、ルータ数 171、リンク数は 174 である。

AT&T トポロジ [20] で紹介されている AT&T のルータレベルのトポロジであり、ルータ数 523、リンク数 1304 である。

上述のアンダーレイネットワーク上にオーバーレイネットワークを構築する際に用いるオーバーレイノードは、指定したオーバーレイノード密度に応じてランダムに選択したルータ上に設置する。オーバーレイノード間のオーバーレイパスの経路はダイクストラ法 [21] によって決定する。

4.2 計測パス数の解析的評価

本節では、提案手法が必要とする計測パス数を数学的解析によって導出することにより、オーバーレイノード密度が計測パス数に与える影響について議論する。

アンダーレイネットワーク、およびオーバーレイノードの配置が与えられた時、計測パス数 M は以下のようにして求めることができる。ここで、 R_{max} をオーバーレイパスの最大経路長

(通過する IP ルータ間のリンク数)、 $P(r)$ を全経路に対する経路長が r であるパスの割合、 n をルータ数、 d をオーバーレイノード密度とする。

経路長が r であるオーバーレイパスの総数は $P(r) \times nd(nd - 1)$ で表される。オーバーレイパス上のルータがオーバーレイノードである確率は各ルータに対して、(送信ノード、受信ノード以外のオーバーレイノード数) / (送信ノード、受信ノード、およびこれまでに通過したオーバーレイパス上のルータでオーバーレイノードではないことが確認されたルータ以外のネットワーク内の全ルータ数) で表される。オーバーレイパスが計測されるのは経路上のルータが全てオーバーレイノードでない場合であるため、経路長が r であるオーバーレイパスが計測される確率は $\prod_{i=1}^{r-1} \left(1 - \frac{nd-2}{n-1-i}\right)$ となる。よって計測パス数 M は式 (1) となる。

$$M = \sum_{r=1}^{R_{max}} \left(P(r) nd(nd - 1) \prod_{i=1}^{r-1} \left(1 - \frac{nd - 2}{n - 1 - i}\right) \right) \quad (1)$$

この式をオーバーレイノード密度 d に関して偏微分を行うと以下のようになる。

$$\frac{\partial M}{\partial d} = \sum_{r=1}^{R_{max}} \left(P(r) n \left(2nd - 1 - nd(nd - 1) \sum_{j=1}^{r-1} \frac{1}{n(1-d) + 1 - j} \right) \times \prod_{i=1}^{r-1} \left(1 - \frac{nd - 2}{n - i - 1}\right) \right) \quad (2)$$

これらの式から、計測パス数は、オーバーレイノード密度 d の値に複雑な影響を受けることがわかる。詳細な評価は次節にて行う。

4.3 シミュレーション結果および考察

4.3.1 計測パス数

図 4、6、8、10 に、それぞれ BA トポロジ、ランダムトポロジ、Abilene-inspired トポロジ、及び AT&T トポロジに対して提案方式を適用した場合の、オーバーレイノード密度に対する計測パス数の解析結果およびシミュレーション結果を示している。図中の fullmesh は、フルメッシュに計測を行う場合に必要となる計測パス数を示している。また、BA トポロジおよびランダムトポロジについては、リンク数 (係数 m) を変化させた場合の結果を示している。これらの図から、ルータ数、リンク数およびネットワークトポロジに関係なく、提案手法はフルメッシュ計測に比べて計測パス数を大幅に削減しており、オーバーレイノード密度が高くなるにつれ、提案方式がより有効になっていることがわかる。また、式 (1) で与えられる計測パス数の解析結果が、シミュレーション結果とほぼ一致しており、解析の妥当性が

確認できる。図 5、7、9、11 は、それぞれのトポロジに関して、式 (2) の評価結果を示したものである。これらの図から、オーバーレイノード密度が増加すると、計測パス数は増加から減少に転じることがわかる。これは、計測パス数の増減はオーバーレイノード密度の増加によるオーバーレイパス数の増加と、ネットワーク中のオーバーレイノードの増加によりオーバーレイパス上に他のオーバーレイノードが存在する確率が大きくなり、提案手法が適用されることによる計測パス数の減少の大小関係によって定まり、オーバーレイノード密度が高くなるほど後者の影響が大きくなるためであると考えられる。

次に BA トポロジ及びランダムトポロジにおいて、リンク数が与える影響に着目すると、 m が増加するにつれて計測パス数が増加するとともに、計測パス数が最大となるオーバーレイノード密度が大きくなることからわかる。これは、 m が大きくなるとネットワーク内のリンク数が増加し、オーバーレイパスの経路長が短くなることによって、オーバーレイパス上に他のオーバーレイノードが存在する確率が低くなるため、提案手法による計測パス数の削減効果が小さくなったものと考えられる。また、アンダーレイネットワークのルータ数の増加による計測パス数の変化に着目すると、ルータが増加するほど計測パスの削減率は高くなること、および、計測パス数が最大となるオーバーレイノード密度が小さくなることがわかった。これは、ルータ数の増加によりオーバーレイパスの経路長が大きくなったことによって、提案手法の計測パス数の削減効果が大きくなったためであると考えられる。

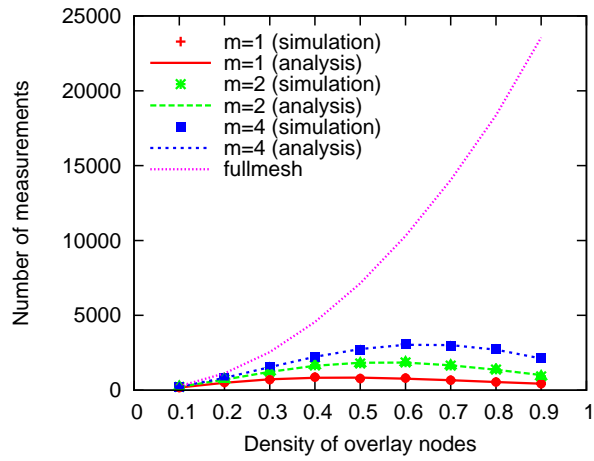
次に、トポロジの違いが提案方式の性能に与える影響について評価を行なう。図 12 は Abilene-inspired トポロジとルータ数、リンク数がほぼ同じである BA トポロジ、ランダムトポロジ ($m=1$) の計測パス数を比較したものである。図から、ランダムトポロジにおける提案方式の性質は、他の 2 つのトポロジとは異なっていることがわかる。これは以下のような理由によるものと考えられる。BA トポロジおよび、Abilene-inspired トポロジはノード度数分布がべき乗則に従っており、ノード度数の大きなハブとなるルータが存在する。ハブは多くのオーバーレイパス上に存在する確率が高い一方で、その数が少ないため、オーバーレイノードである可能性は小さい。その結果、多くのオーバーレイパス上にオーバーレイノードではないルータが存在する可能性が高くなり、パスの計測が行われる可能性が大きくなるためであると考えられる。

4.3.2 情報交換回数

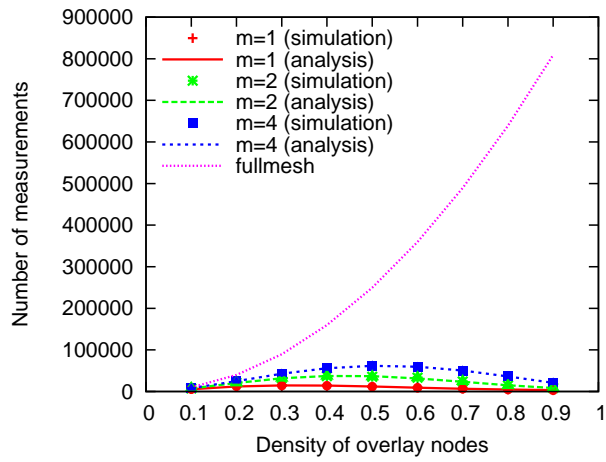
提案手法を用いてオーバーレイパスの資源情報を推測するためには、計測を行っていない区間の情報を他のオーバーレイノードとの情報交換によって得る必要がある。提案手法においては、traceroute パケットが通過したオーバーレイノードが、traceroute コマンドの送信元に必要な情報を送信する。そこで本節では、提案手法を用いた場合のオーバーレイノード密度と資源

情報の送信回数との関係性を評価する。

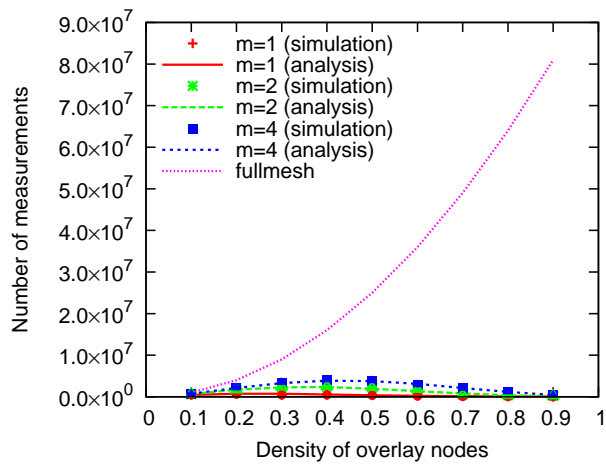
図 13 および 14 に、BA トポロジおよびランダムトポロジにおける、オーバーレイノード密度と情報送信回数との関係性を示している。図から、送信回数は m が大きくなるほど増加していることがわかる。これは、送信回数を削減するためにオーバーレイパスの計測数の削減と同様に、オーバーレイパスの経路の重複を利用しているためである。また BA トポロジと比べて、ランダムトポロジにおける情報送信回数は多くなっている。これは、提案手法を用いた場合の計測パス数が少ないことと、BA トポロジにおけるハブが情報の送信回数の削減に大きく寄与しているためであると考えられる。



(a) ルータ数 171

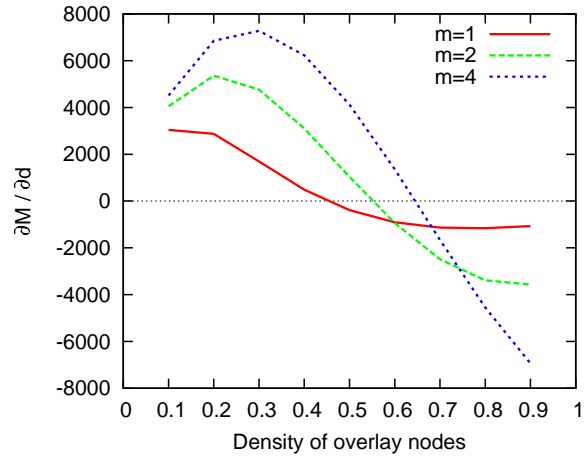


(b) ルータ数 1000

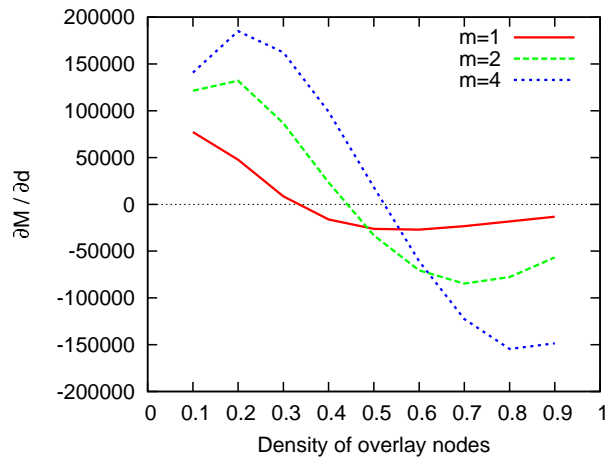


(c) ルータ数 10000

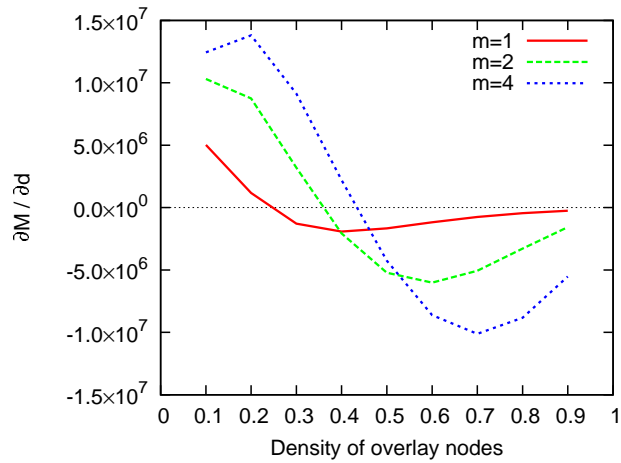
図 4: BA トポロジにおける計測パス数



(a) ルータ数 171

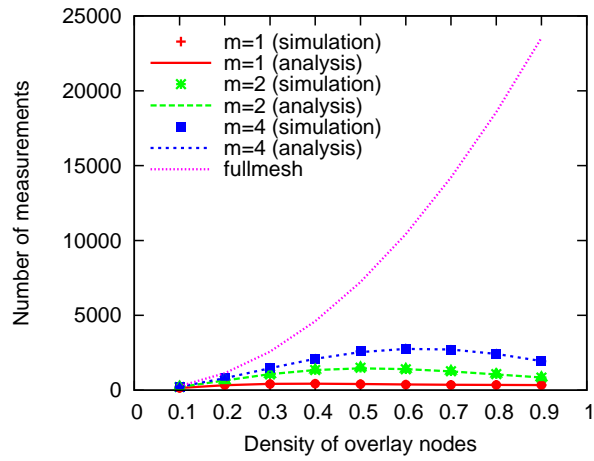


(b) ルータ数 1000

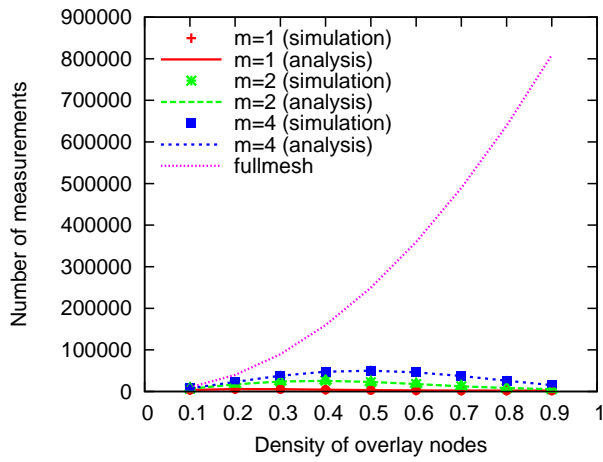


(c) ルータ数 10000

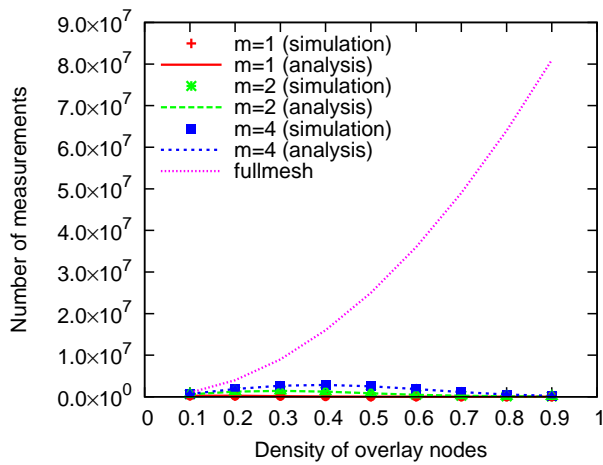
図 5: BA トポロジにおける式 (2) の評価結果



(a) ルータ数 171

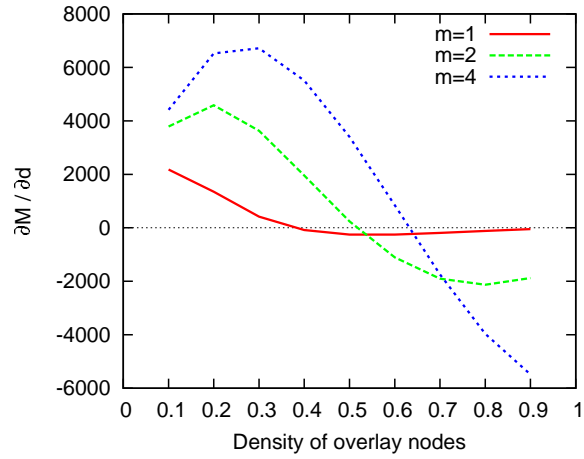


(b) ルータ数 1000

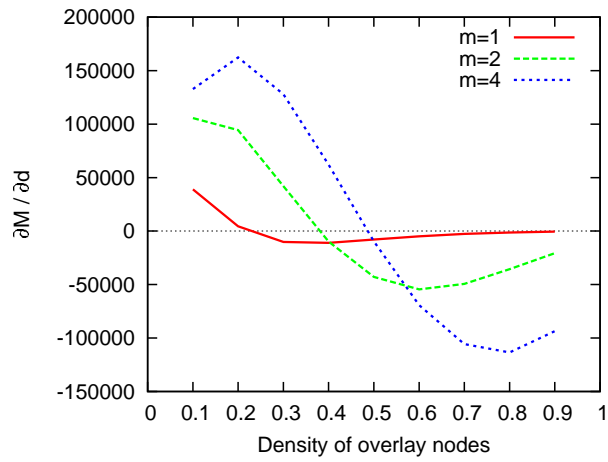


(c) ルータ数 10000

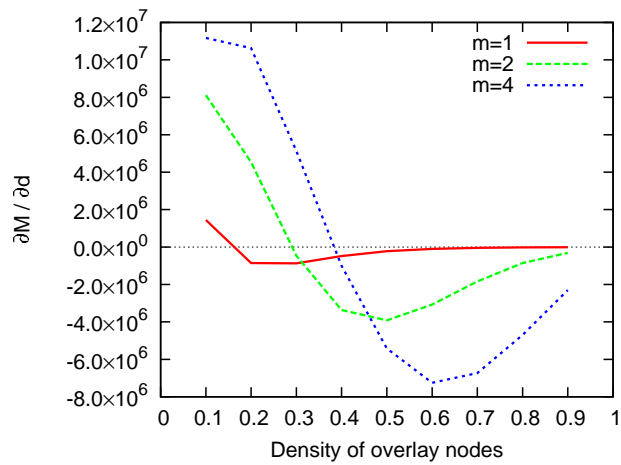
図 6: ランダムトポロジにおける計測パス数



(a) ルータ数 171



(b) ルータ数 1000



(c) ルータ数 10000

図 7: ランダムトポロジにおける式 (2) の評価結果

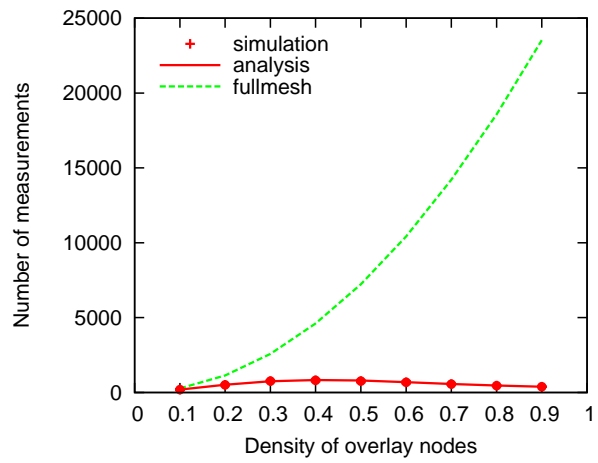


図 8: Abilene-inspired トポロジにおける計測パス数

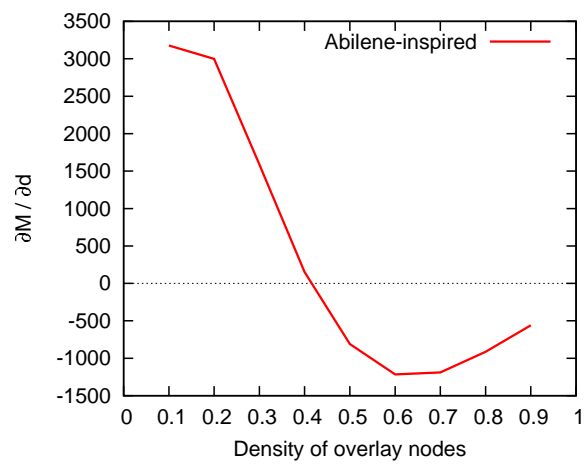


図 9: Abilene-inspired トポロジにおける式 (2) の評価結果

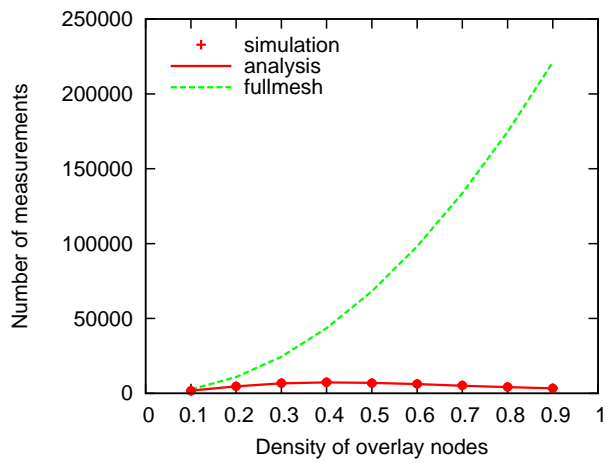


図 10: AT&T トポロジにおける計測パス数

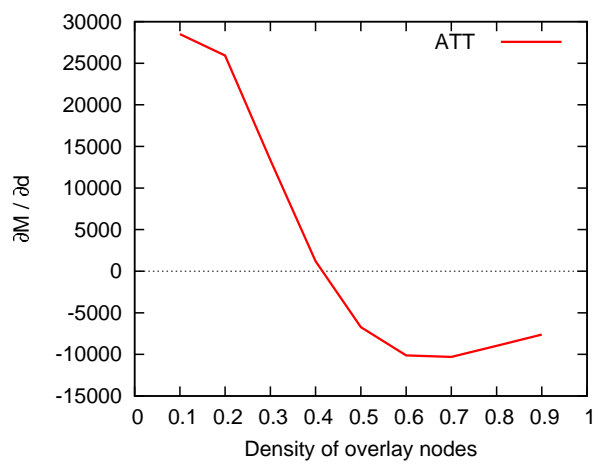


図 11: AT&T トポロジにおける式 (2) の評価結果

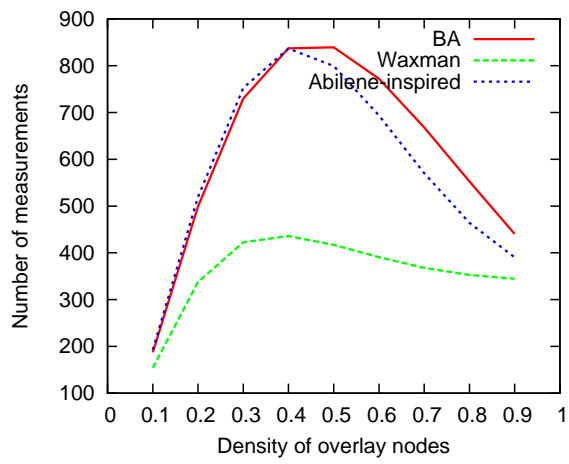


図 12: トポロジの違いが与える影響

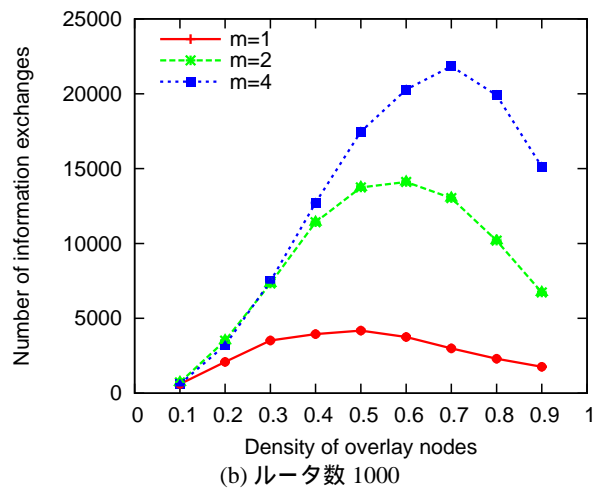
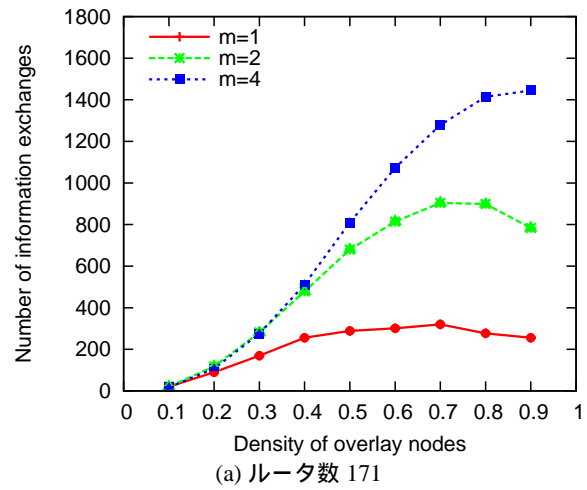
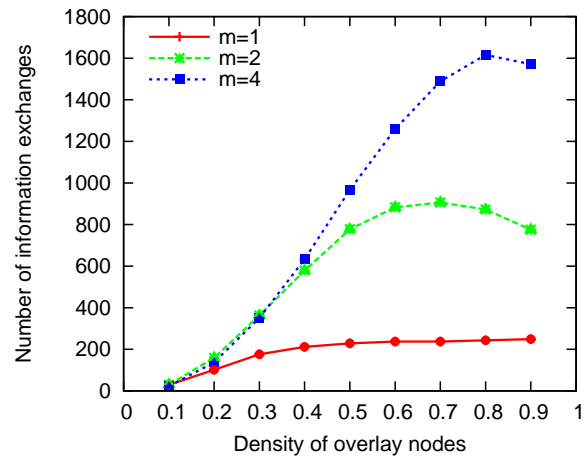
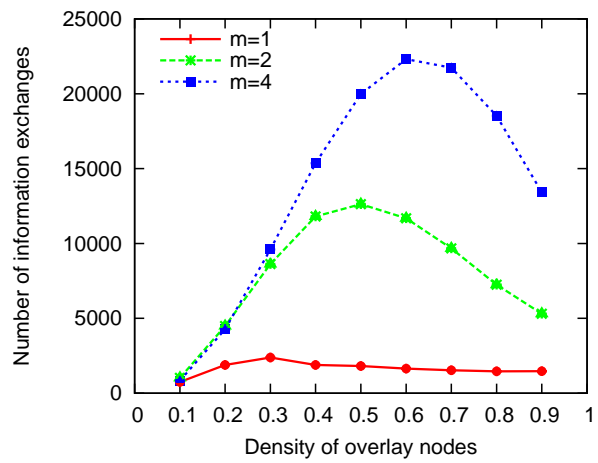


図 13: BA トポロジにおける資源情報の送信数



(a) ルータ数 171



(b) ルータ数 1000

図 14: ランダムトポロジにおける資源情報の送信数

5 計測パス数を削減するためのオーバレイノードの設置戦略

これまでの結果から、ノード次数分布がべき乗則に従うようなネットワークにおいては、ノード次数の大きな少数のルータが多くのオーバレイパス上に存在する可能性が高く、そのことが計測パス数に大きな影響を与えることがわかった。そこで本章では、ノード次数が高いルータにオーバレイノードを優先的に設置し、残りのオーバレイノードは他のルータからランダムに選択することで、計測パス数を削減できることを示す。本報告ではこの手法をハブ優先設置法と呼ぶものとする。

5.1 ハブ優先設置法の解析的評価

まず、4.2節と同様にハブ優先設置法における計測パス数を数学的に解析することによって、各種パラメータが計測パス数へ与える影響の評価を行う。 M は計測パス数、 n はルータ数、 d はオーバレイノード密度、 h は優先的にハブに設置するオーバレイノード数のルータ数に対する割合、 R_{max} はオーバレイパスの最大経路長、 $P(r, d, h)$ はノード密度 d 、ハブに設置するオーバレイノードのルータ数に対する割合が h であるの場合の、全経路に対する経路長が r のオーバレイパスの割合、 $l(i)$ はノード次数が上位 i 番目のルータが持つノード次数、 E はネットワーク中のリンク数である。

経路長が r のオーバレイパス数は $P(r, d, h) \times nd(nd-1)$ で表される。経路上のルータがハブ(ノード次数が上位 nh 番目以内のルータ)に接続しておらず、かつオーバレイノードである確率は、各ルータに対して、((送信ノード、受信ノード以外のオーバレイノード) - (ハブのオーバレイノード数)) / (送信ノード、受信ノード、ハブおよびこれまでに通過したオーバレイパス上のルータでオーバレイノードではないことが確認されたルータ以外のネットワーク内の全ルータ数) で表される。また、送信ノード、受信ノードのうち、ハブに設置されているノード数が 2、1、0 である確率はそれぞれ $\frac{nh}{nd} \frac{nh-1}{nd-1}$ 、 $2 \frac{nh}{nd} \left(1 - \frac{nh-1}{nd-1}\right)$ 、 $\left(1 - \frac{nh}{nd}\right) \left(1 - \frac{nh}{nd-1}\right)$ である。したがって、オーバレイパス上の送信ノードから i ホップ目のルータがハブではなく、かつオーバレイノードではない確率 $NH(n, d, h, i)$ は

$$NH(n, d, h, i) = 1 - \left(\frac{nh}{nd} \frac{nh-1}{nd-1} \frac{nd-nh}{n-nh-i+1} + 2 \frac{nh}{nd} \left(1 - \frac{nh-1}{nd-1}\right) \frac{nd-nh-1}{n-nh-i} + \left(1 - \frac{nh}{nd}\right) \left(1 - \frac{nh}{nd-1}\right) \frac{nd-nh-2}{n-nh-i-1} \right) \quad (3)$$

となる。

一方、ランダムに選択したリンクがハブに接続している確率は、(ハブの持つノード次数

表 1: α の変化

	$m = 1$	$m = 2$	$m = 4$
$n = 171$	2.3	2.2	2.1
$n = 1000$	3.0	2.8	2.7

の和) / (2 × (ネットワーク全体のリンク数)) で表される。ただし、オーバーレイパスの経路上のリンクがハブに接続している確率は、ランダムに選択したリンクがハブに接続している確率より大きくなると考えられるので、このときの確率の増加割合を α とする。このとき、オーバーレイパス上の隣接するルータがハブではない確率 $H(n, h, E)$ は、

$$H(n, h, E) = 1 - \alpha \frac{\sum_{j=1}^{nh} l(j)}{2E}$$

となる。したがって、計測されるオーバーレイパス数 M は以下ようになる。

$$M = \sum_{r=1}^{R_{max}} \left(P(r, d, h) nd(nd-1) \prod_{i=1}^{r-1} (NH(n, d, h, i) H(n, h, E)) \right) \quad (4)$$

なお、 α の値はトポロジによって異なる。BA トポロジに対する評価結果を表 1 に示す。ここで、式 (4) は $P(r, d, h)$ 、 $NH(n, d, h, i)$ 、および $H(n, h, E)$ という項を含むため、オーバーレイノード密度 d が計測パス数へ与える影響を考察することは困難である。そこで、 $h \ll d$ 、すなわち、ハブとなるルータ数はオーバーレイノード数に比べて十分小さいという条件の下に近似を行う。

$h \ll d$ である場合、オーバーレイパスの分布はオーバーレイノードを全てランダムに選択した場合に近くなる。そこで、ハブ優先設置法における経路長分布を $P_{hub}(r, d, h)$ 、全てのオーバーレイノードをランダムに設置した場合の経路長分布を $P_{random}(r)$ とすると、

$$P_{hub}(r, d, h) \simeq P_{random}(r) \quad (5)$$

である。また、 $h \ll d$ ならば、ほとんどのオーバーレイノードはハブには設置されておらず、また、 $n > nd \gg nh$ より $n \gg nh$ かつ $nd \gg nh$ であるため、式 (3) は、

$$NH(n, d, h, i) \simeq 1 - d \quad (6)$$

と近似することができる。これらより、計測パス数 M は以下のように近似できる。

$$M = \sum_{r=1}^{R_{max}} \left(P_{random}(r) n d (n d - 1) (1 - d)^{(r-1)} \left(1 - \alpha \frac{\sum_{j=1}^{nh} l(j)}{2E} \right)^{(r-1)} \right) \quad (7)$$

式 (7) を d に対して偏微分を行うと以下ようになる。

$$\frac{\partial M}{\partial d} = \sum_{r=1}^{R_{max}} \left(P_{random}(r) n (2nd - (r+1)nd^2 - 1 + rd) (1-d)^{r-2} \left(1 - \alpha \frac{\sum_{j=1}^{nh} l(j)}{2E} \right)^{(r-1)} \right) \quad (8)$$

式 (7) より、 $\sum_{i=1}^{nh} l(i)$ が大きいほど計測パス数が削減されることが読み取れる。

さらに、適切な h の値、すなわち、ハブに優先的に設置するオーバーレイノード数の適正値を導出するために、BA トポロジにおける h の値がノード次数上位 nh 個のルータのノード次数の和 ($\sum_{i=1}^{nh} l(i)$) に与える影響について考察する。

BA トポロジにおいてノードの平均次数が $2m$ であるとき、ルータ i のノード次数 k_i がノード次数 k 未満である確率は、

$$P(k_i < k) = 1 - \frac{m^2}{k^2} \quad (9)$$

である。なお、 n はルータ数である。したがって、ノード次数 k のルータの割合 $P(k)$ は

$$P(k) = P(k_i < k+1) - P(k_i < k) \quad (10)$$

$$= \frac{m^2(2k+1)}{k^2(k+1)^2} \quad (11)$$

となる。よって、ノード次数のランキングが下位から数えて $n(1-h)$ 番目のルータのノード次数 k' は

$$1 - h = P(k_i < k') \quad (12)$$

$$1 - h = 1 - \frac{m^2}{k'^2} \quad (13)$$

$$k' = \frac{m}{\sqrt{h}} \quad (14)$$

と求めることができる。ここで、ノード次数が高い上位 nh 個のルータのノード次数の和 ($\sum_{i=1}^{nh} l(i)$) はノード次数が低い下位 $n(1-h)$ 個のノード次数の和をノード次数の総数 (リ

ンク数の2倍)から引いたものに等しいので、

$$\sum_{i=1}^{nh} l(i) = 2E - \sum_{j=m}^{k'} nP(j)j \quad (15)$$

$$= 2E - nm^2 \sum_{j=m}^{k'} \frac{2j+1}{j(j+1)^2} \quad (16)$$

となる。

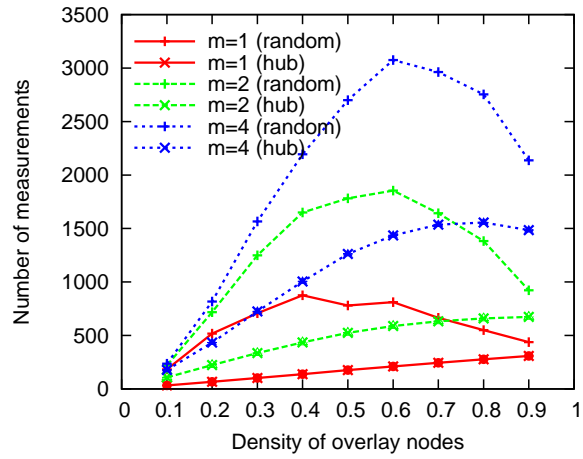
5.2 シミュレーション評価

図 15(a) に、BA トポロジにおいてルータ数が 171 および $h=0.1$ の場合における、オーバーレイノード密度と計測パス数の関係を示している。図から、オーバーレイノードをすべてランダムに選択した場合と比較して計測パス数を最大で約 1/5 程度に削減できることがわかる。また、同じリンク数を持つネットワークにおける最大計測パス数は 0.3-0.5 倍となっている。また、計測パス数が最大となるオーバーレイノード密度が、オーバーレイノードをランダムに選択した場合と比較して大幅に大きくなることがわかった。これは、ハブに優先的にオーバーレイノードを配置するため、オーバーレイノード密度が増加すると、増加したオーバーレイノードは他のオーバーレイパス上に存在する可能性が低い位置に配置されることが多くなるためであると考えられる。図 15(b) はルータ数が 1000 および $h=0.05$ の場合の結果を示している。全体的な傾向は図 15(a) の場合と同様であるが、 h が小さいにもかかわらずオーバーレイ密度が等しい場合におけるの計測パス数が最大で約 1/30、同じリンク数を持つネットワークにおける最大計測パス数が 0.1-0.3 倍となり、計測パス数の削減効果が高い。これは、ノード次数分布がべき乗則に従うようなネットワークでは、ルータ数が増加するほどにより多くのルータがハブに接続されているためであると考えられる。

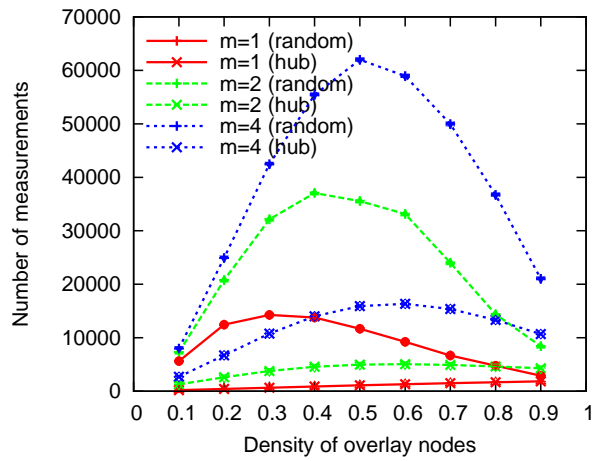
また、式 (4) の評価結果とシミュレーション結果の比較を行ったものを図 16 に示す。図より、ハブ優先設置法における計測パス数の解析結果が、シミュレーション結果とほぼ一致しており、解析結果の妥当性が確認できる。さらに、式 (7) とシミュレーション結果の比較を行った結果を図 17 に示す。 d と比較して h が大きい場合にはシミュレーション結果との差が大きくなっているが、オーバーレイノード密度の変化に対する増減の傾向は等しいことがわかる。そこで、オーバーレイノード密度に対する式 (8) の評価結果を図 18 に示す。図より、オーバーレイノードを完全にランダムに設置した場合と比較して、計測パス数が最大となるオーバーレイノード密度が大きいこと、特に $m = 1$ の場合においてはオーバーレイノード密度の増加にあわせて計測パス数が増加し続けることが確認できる。

図 19 は、横軸を h とし、縦軸を式 (16) の計算結果としたものである。なお、グラフが滑

らかでないのは数式(15)において総和を用いているため、 h の増加に対し、 k' が1以上増加しないと値が変化しないためである。図より、ハブへのオーバレイノードの設置コストと計測パス数の削減効果を考慮すると h は0.15-0.2程度がよいのではないかと考えられる。

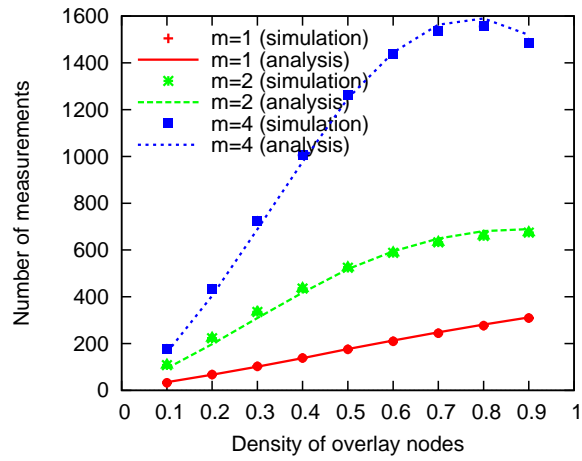


(a) $n = 171, h = 0.1$

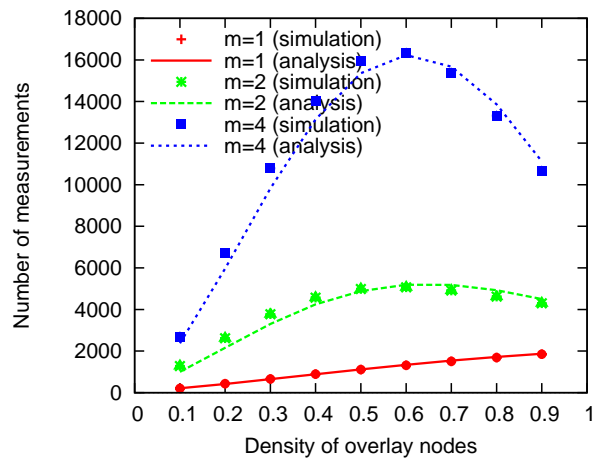


(b) $n = 1000, h = 0.05$

図 15: ハブ優先設置法の性能評価

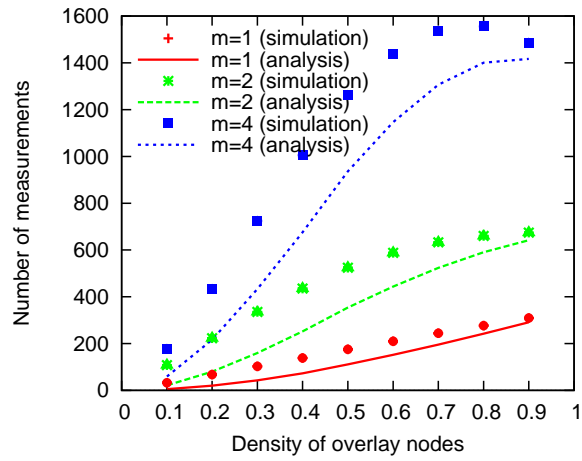


(a) $n = 171$, $h = 0.1$

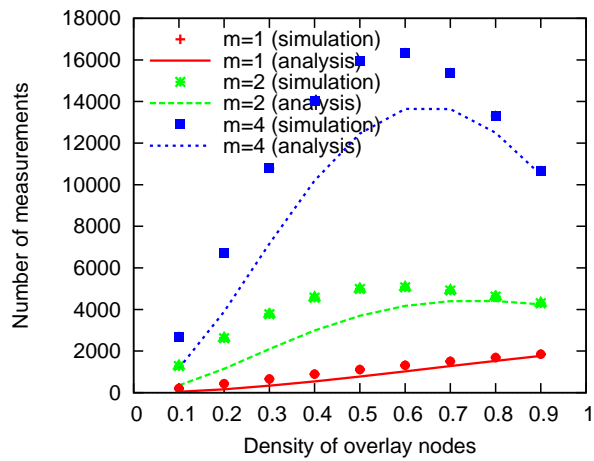


(b) $n = 1000$, $h = 0.05$

図 16: シミュレーション結果と式 (4) との比較



(a) $n = 171, h = 0.1$



(b) $n = 1000, h = 0.05$

図 17: シミュレーション結果と式 (7) との比較

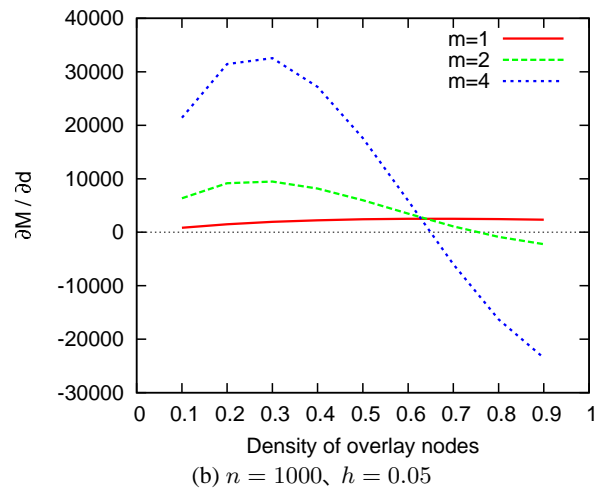
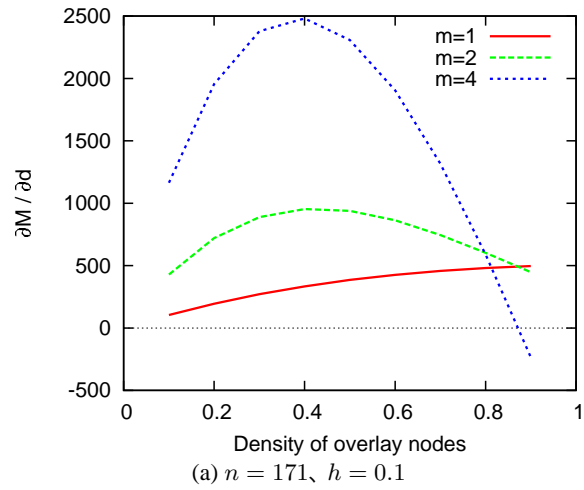
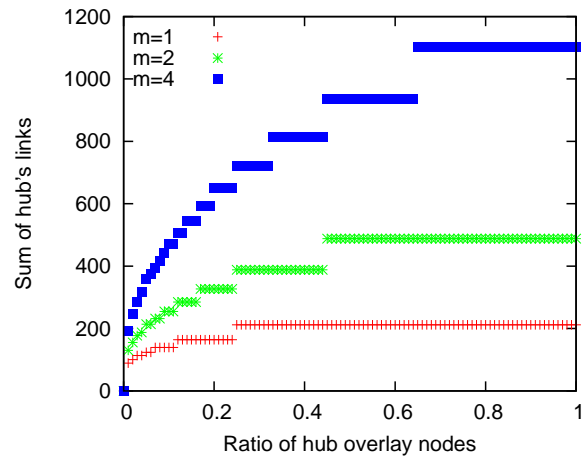
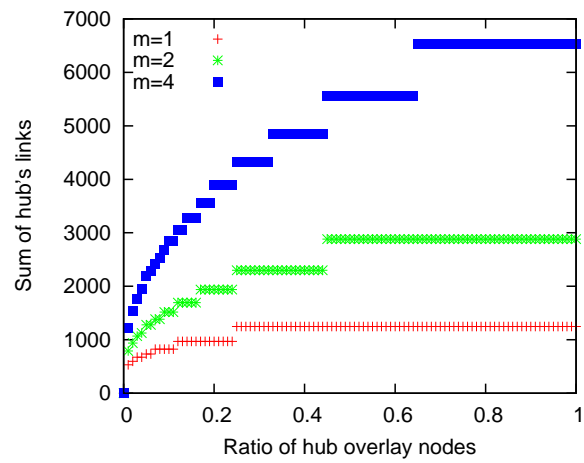


図 18: ハブ優先設置法における式 (8) の評価



(a) ノード数 171



(b) ノード数 1000

図 19: BA トポロジにおけるハブのノード次数の総和

6 まとめと今後の課題

本報告では、オーバーレイネットワークにおける資源情報の計測の際に、ホップ数の大きなオーバーレイパスがホップ数の小さな複数のオーバーレイパスから構成されている場合に、ホップ数の大きなオーバーレイパスの資源情報を計測ではなくホップ数の小さなオーバーレイパスの計測結果から推測によって取得することによって、計測パス数を削減する手法の提案を行った。また、提案手法を数学的解析手法、および様々なアンダーレイネットワークトポロジを用いたシミュレーションによって評価し、フルメッシュ計測と比較して大幅に計測パス数が削減できることを示した。さらに、ノード次数の分布を利用したオーバーレイノードの設置戦略を提案し、提案手法よりもさらに計測パス数を削減できることを示した。

今後の課題として、オーバーレイノード密度の増加とともに問題となる、オーバーレイパスの重複が引き起こす他の問題に対する改善手法の提案が挙げられる。また、経路の重複の増加によるネットワーク計測の衝突の回避法の提案も重要な課題として挙げられる。

謝辞

謝辞本報告を終えるにあたり、御多忙の中、時間を割いていただき、多大なる御指導、御教授を頂きました中野博隆教授、村田正幸教授に深く感謝いたします。また、本報告において終始御指導頂きました長谷川剛准教授には、多くの的確な助言を頂きましたことを心よりお礼申し上げます。的確な助言をいただきました谷口義明助教および大阪大学大学院工学科の笹部昌弘助教に深く感謝致します。最後に日頃から様々な相談にのって頂いた児玉瑞穂氏、橋本匡史氏、堀江拓郎氏、松田一仁氏、および野田健氏をはじめとする中野研究室の皆様に心よりお礼申し上げます。

参考文献

- [1] C. Tang and P. MacKinly, “On the cost-quality tradeoff in topology-aware overlay path probing,” in *Proceedings of ICNP 2003*, Nov. 2003.
- [2] Y. Chen, D. Bindel, H. Song, and R. Katz, “An algebraic approach to practical and scalable overlay network monitoring,” in *Proceedings of SIGCOMM 2004*, Aug. 2004.
- [3] C. Tang and P. McKinley, “A distributed approach to topology-aware overlay path monitoring,” in *Proceedings of ICDCS 2004*, Mar. 2004.
- [4] D. Andersen, H. Balakrishnan, F. Kaashoek, and R. Morris, “Resilient overlay networks,” in *Proceedings of 18th ACM symposium on operating systems principles*, Oct. 2001.
- [5] A. Nakano, L. Peterson, and A. Baver, “Scalable routing overlay networks,” *ACM SIGOPS Operating Systems Review*, vol. 40, pp. 49–61, 2006.
- [6] S. Banerjee, B. Bhattacharjee, and C. Kommareddy, “Scalable application layer multicast,” in *Proceedings of SIGCOMM 2004*, Aug. 2004.
- [7] B. Zhang, S. Jamin, and L. Zhang, “Host multicast: A framework for delivering multicast to end users,” in *Proceedings of INFOCOM 2002*, June 2002.
- [8] T. S. E. Ng and H. Zhang, “Predict Internet network distance with coordinates-based approaches,” in *Proceedings of INFOCOM 2002*, June 2002.
- [9] S. Ratnasamy and et al., “Topologically-aware overlay construction and server selection,” in *Proceedings of INFOCOM 2002*, June 2002.
- [10] P. Francis, S. Jamin, C. Jin, Y. Jin, D. Raz, Y. Shavitt, and L. Zhang, “IDMaps: a global Internet host distance estimation service,” *IEEE/ACM Transactions on Networking*, vol. 9, no. 5, pp. 525–540, 2001.
- [11] Y. Chen, K. H. Lim, R. H. Katz, and C. Overton, “On the stability of network distance estimation,” *SIGMETRICS Performance Evaluation Review*, vol. 30, no. 2, pp. 21–30, 2002.
- [12] Skype Web Page. available at <http://www.skype.com/>.
- [13] BitTorrent Web Page. available at <http://www.bittorrent.com/>.

- [14] Veoh Web Page. available at <http://www.veoh.com/>.
- [15] M. Fraiwan and G. Manimaran, "On the schedulability of measurement conflict in overlay networks," in *Proceedings of Networking 2007*, Nov. 2007.
- [16] A. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, Oct. 1999.
- [17] BRITE: Boston university Representative Internet Topology gEnerator. available at <http://www.cs.bu.edu/brite/>.
- [18] B. M. Waxman, "Routing of multipoint connections," *IEEE journal on selected areas in communications*, vol. 6, pp. 1617–1622, Dec. 1988.
- [19] L. Li, D. Alderson, W. Willinger, and J. Doyle, "A first-principles approach to understanding the Internet's router level topology," in *Proceedings of INFOCOM 2004*, Aug. 2004.
- [20] N. Spring, R. Mahajan, and D. Wetherall, "Measuring ISP topologies with rocketfuel," in *Proceedings of SIGCOMM 2002*, Jan. 2002.
- [21] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, pp. 269–271, 1959.