

研究計画書

物理トポロジを考慮したネットワークマネジメント手法の研究

慶應義塾大学 環境情報学部

自署: _____

希望プログラム:サイバーインフォマティクス(CI)

平成 21 年 3 月 25 日

概要

インターネットに代表される広域ネットワークの高速化が進む中で、P2P ネットワークというインターネットとは違ったコミュニケーションモデルを利用する手法が研究されている。このようなネットワークは、インターネットの新しい利用方法を提案する反面、そのマネジメントが難しいことから様々な問題を抱えている。本研究では、そのようなマネジメントの難しさを克服し、P2P ネットワークの利点を最大化できるフレームワークを提案する。その結果、P2P ネットワークが抱える最大の問題であるインターネットバックボーンへの負荷が解消され、ユーザは今まで以上に便利なアプリケーションやサービスを利用可能となることが期待できる。

1 はじめに

近年、インターネットなどの計算機を用いたネットワークの高速化が進んでいる。特に、企業や大学、家庭などで利用されるネットワークの高速化はめざましく、以前では実現の難しかった新しいアプリケーションやサービスが次々にインターネット上に登場し、様々な場面で様々な人々に利用され、利用形態の多様化が進んでいる。

また、そのような中で、計算機ネットワークの上のコミュニケーションモデルにおいて、オーバーレイネットワークというインターネットとは違ったネットワーク構築手法を採用したサービスが展開されるようになってきている。オーバーレイネットワークとは、平面的に接続性を提供しているインターネット上に、自由に仮想網を構築することによって成立するネットワークであり、下位層トポロジとは別に独自の論理トポロジによって構成されている。

オーバーレイネットワークを基盤とした新しいネットワークとして P2P ネットワークが注目されている。P2P ネットワークとは、オーバーレイネットワークによって構成される論理トポロジにおいて、互いの計算機が対等な役割を担うことにより成立するネットワークである。P2P ネットワークでは大容量のデータを高速かつ低コストで転送することが可能であり、インターネットの可能性を大きく発展させている。例えば、インターネットを介した音声・映像通話を実現している Skype[1] では、インターネット上に P2P ネットワークを構築することでサービスを行っている。

しかし、P2P ネットワークには、ネットワーク自体の論理トポロジやネットワーク上で流通しているコンテンツの管理・制御、構成の最適化などマネジメントが難しいという問題がある。そのため、P2P ネットワークが発生させる膨大なトラフィックがインターネットバックボーンに必要以上の負荷を掛けたり、コンピュータウイルスなどの影響により流通すべきでない情報が交換されるなどの様々な問題が起き、特に、トラフィックに関しては、単一の ISP や国だけでなく、世界レベルの問題となっている。

したがって、本研究では P2P ネットワークにおいて、マネジメントを可能にする機構を取り入れる。この機構を取り入れることで、冗長なネットワークの最適化・ネットワーク中のデータ管理が可能となる。そして、社会問題となっている P2P のデメリットを克服し、インターネットにおけるポリシー・制御などへの影響を最小限にする。

2 コミュニケーションモデル

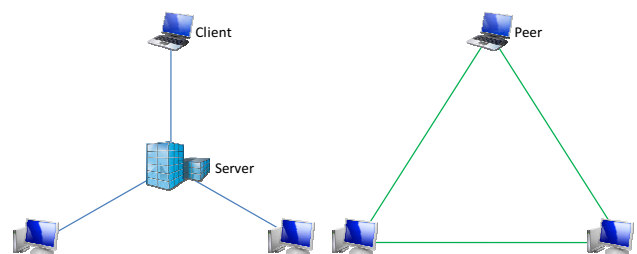
P2P ネットワークは既存のインターネットとは異なるネットワーク構築手法によって構成されている。本章では、コミュニケーションという観点からそれぞれのネットワークをモデル化し、その性質や特徴について説明する。

2.1 クライアント・サーバモデル

現在のインターネット上の通信の基本であるコミュニケーションモデルを図 1(a) に示す。コミュニケーションはサーバを中心に行われるため、クライアント・サーバモデルと呼ばれており、各ユーザの計算機はサーバとのコミュニケーションのみによりサービスの提供を受ける。

クライアント・サーバモデルのメリットは、通信の対象が決まっているということである。ユーザは必ずサーバへ接続し、他の計算機との接続は行わないため、ネットワークは自動的に最適化され、サーバや管理者がマネジメントを行う必要はない。

一方で、デメリットはサーバが単一障害点になりやすいということである。サーバはユーザと常にコミュニケーションする計算機であるため、一度に大量の接続や要求を受ける可能性があり、過負荷による故障や不具合が起こる可能性がある。特に、ユーザからの大量の接続を受ける場合にはサーバだけでなく、(ユーザからサーバまでの)ネットワーク全体に負荷が掛かる可能性があり、影響範囲がサーバ周辺で収束しない場合もある。



(a) クライアント・サーバ (b) P2P
図 1: コミュニケーションモデルの比較

2.2 Peer-to-Peer (P2P) モデル

本研究で対象とするコミュニケーションモデルを図 1(b) に示す。コミュニケーションにおいて中心は存在せず、ユーザの計算機は互いに対等な役割で互いにサービスの提供・享受を行う。

P2P モデルのメリットは、通信の対象が複数存在することである。ユーザの計算機はネットワーク内の複数の計算機とコミュニケーションするため、単一の計算機に接続や処理が集中することはない。そして、一部の計算機に故障などが起こった場合にも、他の計算機が役割を代替することでネットワークを継続して運用できるため、ネットワーク全体としての耐故障性は高い。

デメリットは、マネージメントが難しいという点である。P2P モデルにおいて、ユーザが必ずコミュニケーションする計算機は存在しないため、各計算機の挙動を全体で把握することはできない。特に、P2P ネットワークではサービスの提供以外に、データを転送する際の経路や配送方法についても各計算機が判断を行うため、管理機構などを備えていない場合にはそのマネージメントを行うことができない。

また、P2P コミュニケーションモデル特有のデメリットとして、コミュニケーション手法であるプロトコルの違いが挙げられる。P2P ネットワークにはインターネットのような共通のプロトコルがなく、様々なネットワークがそれぞれ独自のプロトコルで運用されている。そのため、ネットワーク上で複数のサービスを連帯、コラボレーションさせて提供することが難しく、その可能性を限定してしまっている。

したがって、本研究では P2P コミュニケーションモデルにおけるマネージメント機構の提案を行い、P2P ネットワークのメリットを最大限活用し、様々なサービスを提供することが可能なフレームワークの構築を目指す。

3 P2P ネットワークの問題

P2P ネットワークをより効率的に運用するためには、そのマネージメントが不可欠である。本章では、マネージメントを行う前提となる P2P ネットワークモデルについて考察し、その問題点について分析を行う。

既存の P2P ネットワークモデルにおける、各計算機間の接続例を図 2 に示す。

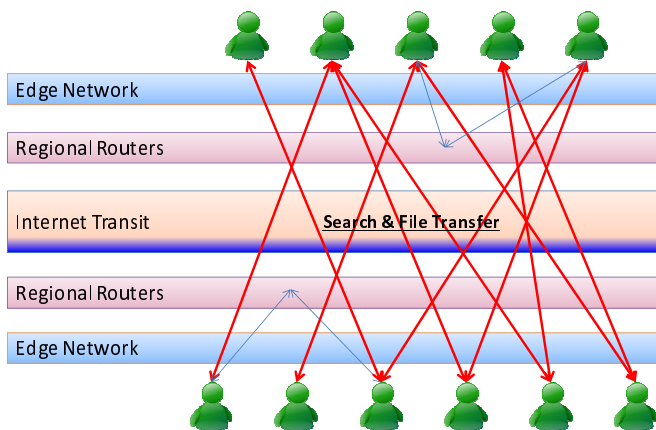


図 2: P2P ネットワークの接続例

図 2 において、各計算機は自律的にネットワーク内の計算機と通信を行い、インターネットを介して大量のデータ転送を行う。データ転送の際の接続方法は P2P ネットワーク上の論理トポロジに依存しており、下位層トポロジには関知しない。

そして、図 2 のような接続体系の場合、同一のデータを転送する場合にも、P2P ネットワークで発生するトラフィックの大部分はインターネットに流れる。そのためにデータ転送時の高負荷かつ大量のトラフィックが直接インターネットに流入し、インターネットバックボーンを圧迫する最大の原因となっている。

特に、P2P ネットワークの論理トポロジは実際のネットワークの物理トポロジと異なる可能性が高く、物理トポロジ上で比較的近いところにデータ転送を行う計算機がいる場合にも、P2P ネットワーク上のトポロジによりインターネットに余分な負荷をかける可能性がある。

したがって、本研究では実際のネットワークの物理トポロジを考慮した P2P ネットワークの構成手法を提案する。この手法を用いることで、P2P ネットワーク上に物理トポロジ情報を反映することが可能となる。そのため、図 2 で問題となっているインターネットを介した P2P ネットワークによるトラフィックを図 3 のように集約し、同一データ転送時の影響を最小化することが可能となる。

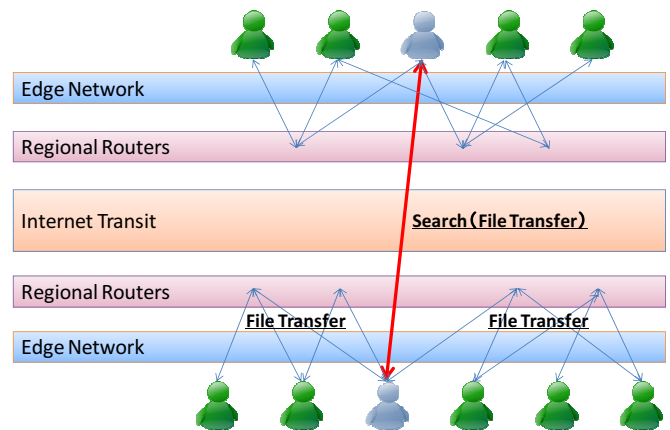


図 3: マネージメントされた P2P ネットワークの接続例

4 既存技術

P2P ネットワークマネージメントについては、いくつかの製品、プロジェクトによって研究されている。

4.1 Skype

Skype[1] は、P2P ネットワークを使用したインターネット電話アプリケーションである。ネットワーク中ですべての計算機（ノード）を対等に扱うのではなく、スーパーノードとリーフノードに役割を分担することでマネージメントを行っている。このようなネットワーク形態を特にマルチレイヤモデルと呼ぶ。

マルチレイヤモデルでは役割分担によりネットワークの負荷分散を簡単に行うことができるが、論理トポロジの最適化に関しては従来の P2P と同様の問題を抱えている。

4.2 Azureus

Azureus[2] は、P2P ネットワークを使ったファイル転送プロトコルである BitTorrent[3] を使用したファイル共有アプリケーションである。Vivaldi[4] という Round Trip Time (RTT) を指標にしたクラスタリングアルゴリズムによりマネージメントを行っている。

クラスタリングはトラフィックを P2P ネットワーク上の狭い範囲で収束できるが、物理トポロジを評価する指標が RTT であるため、必ずしもインターネットを介したトラフィックを減らすことができるわけではない。

4.3 ONO

ONO[5] は、ダウンロードを高速化するための Azureus のプラグインである。ネットワーク的に近いノードの発見に RTT と Contents Delivery Network (CDN) を利用している。CDN を指標として用いることで、物理的に近いノードに接続することが可能となっている。

ただ、ONO の CDN の使い方は本来の使い方ではないため、プラグインを導入したノードが CDN へ影響を与える可能性や CDN がそれほど地理的に分散して配置されていない場合などを考慮しなければならず、CDN のポリシー的・技術的な課題がある。

4.4 P4P

P4P[6] は、インターネット接続業者 (ISP) と連帯してネットワークマネジメントを行う手法である。通常では手に入れることができない P2P ネットワークの下位層トポロジやポリシーに関して正確な情報が得られるため、アプリケーション側でそれらの推測や構成を行う必要が無く、効率的な網の構成を可能としている。

しかし、ISP は P2P ネットワーク事業者へ情報を提供せねばならず、また、すべての ISP が統一的に情報を提供する仕組みや取り決めがないため、今後の動向が注目されている。

4.5 Landmark Routing

Landmark Routing[7] は、Landmark をネットワークの適所に配置することによってマネジメントを行う経路制御手法である。Landmark を中心とした階層的なネットワーク構成を取ることで、構造化 P2P ネットワークにおける経路を集約、マネジメントを行っている。

ただ、Landmark Routing は通信を行う際に必ず Landmark を利用するため、Landmark が過負荷になった場合などにパフォーマンスが著しく低下する可能性がある。また、構造化 P2P ネットワークを基盤としているため、そのオーバーヘッドや最適化について考慮する必要があり、マネジメントが煩雑になる。

5 物理トポロジによるマネジメント

本研究では、実際のネットワークの物理トポロジに着目した P2P ネットワークマネジメント機構の導入を提案する。

これは、P2P ネットワークを構成する計算機の物理トポロジにおける近傍性を考慮してマネジメントを行うことで、P2P ネットワークで発生するトラフィックのインターネットバックボーンへの流入をできるだけ減らすという手法である。

物理トポロジにおける近傍性を考慮する際の指標に関しては、ユーザ (エンドポイント) の環境や、階層などによ

り様々な要因が存在するため、一概に決定することは難しく、P2P ネットワークの柔軟性を失わせることにもなりかねない。したがって本研究計画では、エンドポイントの影響を受けることなく近傍性を表現する手法として、計算機が接続している ISP に着目する。

5.1 ISP を指標とした物理トポロジ

本研究計画の立案にあたり、日本国内で最大規模のユーザ数を誇る P2P ネットワークである Winny[8] について、利用者の ISP 分布を調査した。結果を図 4 に示す。

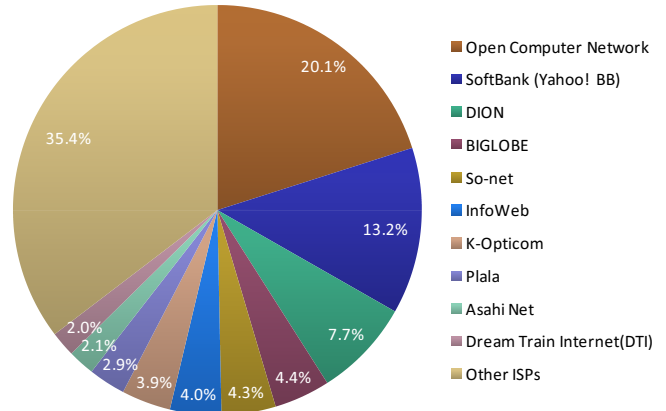


図 4: Winny 利用者の ISP 分布

詳細を以下に示す。(データ引用基: WHOIS[9])

- 環境
一台の計算機に Winny をインストール、初期設定を行った後に、起動・監視 (48 時間)
- 総 ISP 数
(海外なども含めて) 430 ISPs
- 接続ノード数
約 10000 ノード
- 国別分布
日本:98%以上、アジア:1%、北米・欧米:1%未満

図 4 より、Winny ネットワークにおいては上位 5 個の ISP で約半分のノードをカバーしていることがわかる。

したがって、このような分布において、現状では図 2 のような構造のために、P2P ネットワークによるトラフィックがインターネットバックボーンに多大な負荷をかけてしまっている。しかし、本研究で提案する近傍性を考慮したネットワークマネジメント手法により、図 3 のような P2P ネットワークを実現することができれば、トラフィックの大部分を ISP 内で完結させ、インターネットバックボーンを多大な負荷から保護することが可能となる。

5.2 マネージメント概要

本手法によるマネジメントの概要を図 5 に示す。

本研究では、ISP 毎に複数のスーパーノードを設置し、各 ISP をマネジメントするスーパーノードネットワークと、各スーパーノードの下位につながるリーフノードネットワークに P2P ネットワークを大別する。このようなマルチレイヤモデルを採用することで、マネジメント・検索などのネットワーク全体のデータを把握する必要がある部分

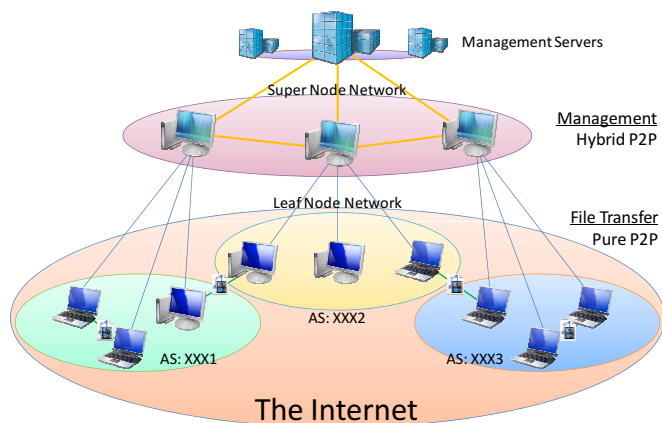


図 5: ネットワークモデル

をスーパーノードが、データ転送などの高負荷な処理部分を（スーパーノードによって管理された）リーフノードが行うことが可能となり、役割分担により通信効率が向上する。スーパーノードはリーフノードネットワークから以下の条件で選ばれ、負荷によってその数は増減する。

- 起動時間
長時間リーフノードネットワークに滞在している。
- 回線速度
FTTH など高速な回線を使って接続している。
- マシンスペック
CPU、メモリなどの性能が高い。

また、P2P ネットワーク全体を統括する管理サーバを設置する。サーバは以下の機能を提供し、P2P ネットワークの利便性を向上させる。

- ランデブーポイント
P2P ネットワークに参加するために、ノードが最初にアクセスするポイントとなる。
- ネットワーク管理
P2P ネットワークの状態を監視し、ネットワークの組み替えや構成を行う。
- コンテンツ管理
ネットワーク内のコンテンツの情報を管理し、サービス管理やユーザ認証などを行う。また、不適切な情報が流通した場合には、データの削除などを指示する。

ただ、サーバを P2P ネットワークに設置する場合には、そのメリットである自律性や分散性に影響が出ないよう、十分に考慮する必要がある。そのため、本研究ではサーバの機能について、利便性を向上させる範囲内での提供を想定している。管理サーバはインターネット上に複数設置することが可能であり、その際、ISP や物理トポロジを考慮する必要は無い。

6 マネージメント動作

ネットワーク管理はノードの参加時に行われる。以下にその動作を示す。

1. サーバへの接続

ノードは起動時にランデブーポイントであるサーバへ接続、認証処理などを行う。

2. 接続先ノードの決定

サーバは接続してきたノードのネットワークポロジと P2P ネットワークモデルを考慮し、P2P ネットワークへのエン트리ポイントを選択する。

3. ネットワークへの参加

サーバより受け取った情報を基にスーパーノードへ接続することで、P2P ネットワークへの参加が完了する。

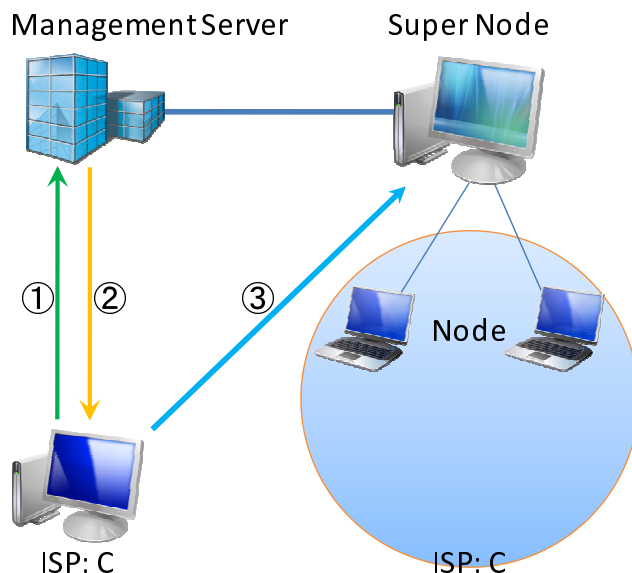


図 6: ネットワーク管理動作

6.1 サーバへの接続

本機構で、サーバはノードが最初に P2P ネットワークに接続する際に接続する必要がある計算機である。ノードがサーバに接続する際、サーバは接続してきたノードのネットワーク情報を得ることが可能である。この情報はリモートホストと呼ばれ、例えば以下のように表される。

p1234-ipbf567marunouchi.tokyo.ocn.ne.jp

この情報は、インターネット利用者の ISP の情報を示したもので、接続してきたノードがどの ISP から接続してきたかを知ることが可能である。

リモートホストの特徴として、その情報の中にリモートホストの設置場所や所有者名が含まれていることがある。上の例では「marunouchi.tokyo」という部分から東京にある計算機が、「ocn.ne.jp」という部分から Open Computer Network (OCN) という ISP 経由で接続してきたということがわかる。

したがって、この方法を用いることでノードが接続する ISP・物理トポロジを調べることが可能となる。本研究ではこの情報を基本にネットワーク管理を行う。

6.2 接続先ノードの決定

サーバは前述した方法で、P2P ネットワークに参加したノードの情報を取得した後、以下の優先順位によりノードのエン트리ポイントを決出し、ノードに情報を送信する。

1. 相対トポロジ情報

ISP と IP アドレスより、ネットワークトポロジ的に最も近いと思われるスーパーノードを選択する。

2. 絶対トポロジ情報

該当する ISP にスーパーノードが存在しない場合、上位プロバイダ情報、トランジットサービス情報などを基に、ネットワークトポロジ的に最も近いと思われるスーパーノードを選択する。

3. 地理情報

リモートホストからの情報や、あらかじめ用意したデータベースなどにより地理情報が利用できる場合、地理的に近いスーパーノードを選択する。

6.3 ネットワークへの参加

ノードは送信されてきたエントリポイント情報を基に、スーパーノードへ接続を行う。スーパーノードはノードの接続を受け入れるが、その際自ノードに掛かる負荷を考慮し、高負荷になると判断した場合はリーフノードネットワークの中から新たにスーパーノードを選出し、以降に接続を要求された場合は新しいスーパーノードへの誘導を行う。

また、ノードは同一 ISP 内に存在する複数のスーパーノードとコネクションを張る。これは、自分の接続しているスーパーノードがネットワークから離脱した場合にも、他のスーパーノードによって P2P ネットワークへの接続性を提供するためである。

7 フレームワークの構築

本研究では、先述した P2P ネットワークモデルにおけるマネージメント機構を実装した、P2P フレームワークの構築を行う。P2P フレームワークとして枠組みを作ること、独自のプロトコルで運用されている様々な P2P ネットワークを容易に連帯、コラボレーションさせ、統一的にマネージメントすることが可能となる。

そして、このようなフレームワークの実現により、社会的な問題となっている P2P ネットワークによるインターネットバックボーンの圧迫やネットワーク上のコンテンツ管理ができないという P2P のデメリットは解消され、サービス事業者・ユーザは今まで以上に便利なアプリケーションやサービスを利用可能となることが期待できる。

例えば、地上デジタル放送に代表される大容量・高品質の情報配信を行う場合に、この P2P フレームワークを使うことで、インターネットバックボーンの負荷を最小限に抑えることが可能となり、同時に、サービスやユーザの管理も容易に行うことができる。また、ユーザ側も情報の取得から展開までをストレス無く行うことが可能となる。

現在でも P2P ネットワークを使った情報配信は BitTorrentDNA[10] や SkeeCast[11] など一部のサービスで利用されているが、ユーザはそれぞれ異なる手順で情報にアクセスしなければならない。しかし、この P2P フレームワークでは同一の P2P プラットフォームを提供するため、ユーザはサービス毎に手順を変える必要は無く、どのような情報に対しても同じようにアクセスすることが可能となる。

8 これまでの活動

私は学部 2 年次より、村井研究室に所属している。研究室では本研究の基礎となる知識や技術について学び、同時に、技術的な問題や仕組みを理解するためにプログラムの作成などを行った。

3 年次に行った活動では、学期前半で実際に動いている P2P ネットワークとして Winny[8] を取り上げ、そのプロトコル動作などを考察し、学期後半でそれらを応用した仮想ストレージを構築するためのソフトウェアを開発した。

私はこのような研究活動を通して、P2P ネットワークの自律性、分散性、スケーラビリティなどの技術を学び、さらには幅広く人々に利用されている要因などについて学んできた。

9 志望理由

政策・メディア研究科では、インターネットに関係する基盤技術だけでなく、社会にどのように研究成果を還元するかといった様々な応用研究が行われている。そして、技術的な側面ばかりに着目するのではなく、多角的かつ総合的な視点で研究を捉えるフィールドがあると考えている。そのため、研究成果が一面的に留まるのではなく、新しい社会を創造・創発するという意味で研究を進めることが可能である。特に、本研究の目的である近傍性の追求は、インターネットなどの情報通信分野だけでなく、現代のネットワーク社会のあらゆる部分で必要となる要素である。したがって、フレームワークとして P2P ネットワークの枠組みを策定するだけでなく、近傍性という視点で、これからのユビキタス社会やモビリティ社会などの発展の可能性を多角的に分析・模索することで、豊かな社会の創造・創発に貢献することが可能となると考えている。

また、研究を進める上で重要なネットワークや計算機資源などの環境があるだけでなく、様々な知見や経験を豊富に持つ指導者の方々が多数在籍している。これは、本研究を実現・進展させ、認知してもらう上で、大変重要な環境である。

以上の理由から私は政策・メディア研究科への進学を強く希望する。

参考文献

- [1] Skype communications s.a. <http://www.skype.com>.
- [2] Azureus. <http://azureus.sourceforge.net/>.
- [3] Bittorrent. <http://www.bittorrent.com/>.
- [4] Frank Dabek, Russ Cox, Frans Kaashoek, and Robert Morris. Vivaldi: A decentralized network coordinate system. In *Proceedings of the ACM SIGCOMM '04 Conference*, Portland, Oregon, August 2004.
- [5] David R. Choffnes and Fabian E. Bustamante. Taming the Torrent: A practical approach to reducing cross-ISP traffic in P2P systems. In *Proceedings of the ACM SIGCOMM '08 Conference*, August 2008.
- [6] P4P: Explicit Communications for Cooperative Control Between P2P and Network Providers. http://www.dcia.info/documents/P4P_Overview.pdf.
- [7] Ben Y. Zhao, Yitao Duan, Ling Huang, Anthony D. Joseph, and John D. Kubiatowicz. Brocade: Landmark routing on overlay networks, March 2002.
- [8] Isamu Kaneko. *The Technology of Winny*. ASCII, October 2005.
- [9] L. Daigle. WHOIS Protocol Specification. RFC 3912 (Draft Standard), September 2004.
- [10] BitTorrentDNA. <http://www.bittorrent.co.jp/>.
- [11] SkeeCast. <http://www.dreamboat.co.jp/>.