

Layer2 ネットワーク構成情報表示システムにおける 構成推測条件の適用順序の影響について

藤田 俊輔[†] 吉田 和幸^{††}

[†] 大分大学工学研究科知能情報システム工学専攻

^{††} 大分大学学術情報拠点情報基盤センター

E-mail: †{v08e3024,yoshida}@csis.oita-u.ac.jp

あらまし 我々は LAN スイッチを主として構成されるイーサネットワークのトポロジを推測するアルゴリズムを提案し、機器間の接続関係を推測・表示するシステムを作成し利用してきた。本システムではトポロジとは無関係に、LAN スイッチに割り当てられた IP アドレスの順序に従って物理的な接続の推測を行う。しかし設定された IP アドレスは必ずしもトポロジを反映しない。そのためアルゴリズムの適用順序によっては推測を誤る可能性がある。そこで本論文では現在使用していないネットワーク構造の情報として、ネットワークの末端に存在する機器を FDB (Forwarding Database) から特定し、実際に LAN スイッチとの接続に用いるインタフェースの数に応じて機器分類し、各区分ごとにアルゴリズムを適用することで推測結果にどのような影響を及ぼすかについて述べる。

キーワード ネットワーク構成, LAN スイッチ, ルータ, SNMP, MIB

Effect of application order of topology inference rule for Layer2 Network Topology Visualizing System

Shunsuke FUJITA[†] and Kazuyuki YOSHIDA^{††}

[†] Dept. of Computer Science and Intelligent Systems, Oita University

^{††} Center for Academic Information and Library Services, Oita University

E-mail: †{v08e3024,yoshida}@csis.oita-u.ac.jp

Abstract We propose the algorithm that infers the Ethernet Network Topology, and implemented the system which infers and displays network topology. In this system, physical connections are inferred in order of IP address assigned to LAN-Switches without any relations to the topology. However, assigned IP addresses of LAN switches do not necessarily reflect the physical topology. Therefore, according to application order of inference rules, the system might lead wrong network topology. In this paper, we describe how to specify the equipment that exists in the end of the network using FDB (Forwarding Database), and describe to classify LAN-Switches according to the number of interfaces connected to the other LAN-Switches, and describe effect of the application order of inference rules on the result.

Key words Network Topology, LAN-Switches, Router, SNMP, MIB

1. はじめに

近年、肥大化と複雑化が顕著なネットワークの管理をする上で、ネットワーク構成を把握することは重要である。ネットワークの構成を正確に把握することにより、ネットワークの障害時には問題に的確に対処し、あるいは防止することができる。ネットワークは物理層・データリンク層・ネットワーク層などの複数の層に分けて管理されるが、その中でも物理層でのネットワークトポロジを反映するデータリンク層のネットワーク

トポロジを把握することが、ネットワークの障害管理において重要である。

しかしながら、データリンク層のネットワークトポロジを把握することは困難である。我々はデータリンク層であるイーサネットワークの物理的構成を把握するために、LAN スイッチをはじめとするネットワーク機器から FDB (Forwarding Database) 等の情報を収集し、機器間の接続関係を推測・表示するシステムを作成し利用してきた [1]~[9]。このシステムでは機器間の物理的な接続関係の推測はイーサネットワークの構造に関わらず、

機器に設定された IP アドレスの順序に応じて行われる。しかし、ネットワークの構造は様々であり、現在運用しているトポロジ以外では推測結果に影響を及ぼすことがある。

そこで、本論文では FDB と、推測された間接接続を含めた接続関係からその機器が LAN スイッチであるのか、ルータやサーバであるのかを判断し、LAN スイッチである場合にはインターフェース数に応じた区分の方法について述べる。またこの区分を用いた推測手順の変更により、推測結果にどのような影響を与えるかについて述べるに伴うアルゴリズム、まず、2 章では関連研究について述べ、3 章でネットワーク構成情報表示システム (以下、本システム) の概要について述べる。4 章では大分大学学内 LAN のネットワーク構成と、それらのネットワーク機器から収集する情報について詳しく述べ、5 章では収集した情報を基に LAN スイッチ間の物理的な接続関係を推測するアルゴリズムについて述べる。6 章では機器の区分の方法と、区分ごとの推測アルゴリズムの適用による影響について述べる。7 章では結論としてまとめと、本システムの今後の課題について述べる。

2. 関連研究

関連研究として、Lowekamp 氏らによる "Topology Discovery for Large Ethernet Networks" [10] と、Breitbart 氏らによる "Topology discovery in heterogeneous IP networks" [11] がある。彼らのアルゴリズムでは対象となる機器すべての間での接続関係が判明した上で、物理的な接続の推測を行う。一方、我々の推測アルゴリズムでは機器間の接続は、必ずしもすべてが判明する必要がなく、また、Lowekamp 氏らのアルゴリズムでは対応できない VLAN により作られるループ構造も不完全ではあるが検出可能である。その他に Son 氏らによる "Topology Discovery in Large Ethernet Mesh Networks" [12] では、STP のみによるネットワークトポロジの発見を行っている。しかし、すべての機器が STP を実装する必要があり、VLAN によるループのない環境でなければならない。様々な機器の混在するネットワークでは、我々のシステムほど有効ではない。

3. イーサネットワークトポロジ表示システムの概要

本システムはネットワークの構成情報を収集する「構成情報収集部 (以下、収集部)」, 収集した情報を基に接続関係を推測する「構成情報推測部 (以下、推測部)」, 推測結果を見やすく表示する「構成情報表示部 (以下、表示部)」の 3 つからなる。図 1 に本システムの概要図を示す。収集部と推測部については、以降の章で述べる。

3.1 管理対象ネットワークの構成について

本システムが管理対象とするネットワークは LAN スイッチ・ルータ・サーバ等がイーサネットにより接続されたネットワークであり、これらのネットワーク機器について、ネットワーク

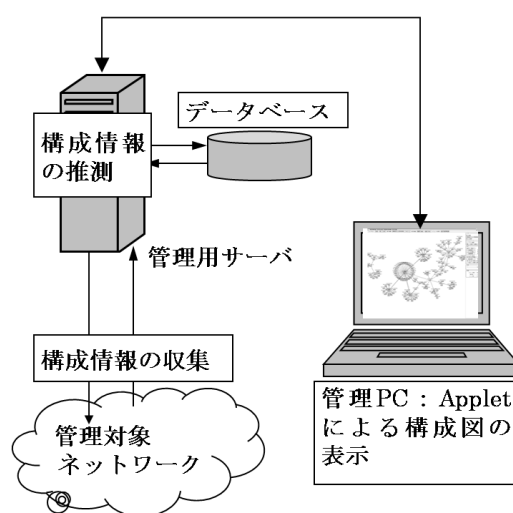


図 1 イーサネットワークトポロジ表示システムの概要図

構成に関する情報を収集し、トポロジを推測する。収集方法としては、telnet などを用いることも考えられるが、この場合は各機器ごとに必要なコマンドと、収集した情報の形式が異なるため構成情報の収集と処理に非常に手間がかかり、有効ではない。そこで本システムではベンダに依存しないネットワーク管理プロトコルである SNMP (Simple Network Management Protocol) を用いる。SNMP を用いることでベンダだけではなく、LAN スイッチやサーバといった機器の種類も意識する必要がなく、得られる情報も統一された形式となる。そのため、本システムでは SNMP を実装している機器を管理対象機器とする。本システムでは多少の例外は許すが大部分の LAN スイッチは、1 つの管理用 VLAN に接続されていることが必要となる。

3.2 構成情報表示部

ネットワークトポロジ表示部は、JavaApplet で作成している。ネットワーク機器を現すノードをランダムに配置し、スプリング手法により徐々に整った配置に移動させる。図 2 はトポロジ情報を読み込んだ時点、図 3 は自動配置を行った時点の実効結果である。

自動配置終了後、各ノードの座標値をファイルに格納することができる。このことにより、別の日に収集したデータから得られたネットワークトポロジとの比較が容易になる。

4. 構成情報収集部

本章では本システムの運用環境である大分大学旦野原キャンパスの LAN (以下、学内 LAN) のネットワーク環境と収集する情報について詳しく述べる。

4.1 収集対象

学内 LAN では、ネットワーク管理用 VLAN を設定して、大部分のネットワーク機器をこの VLAN に参加させている。本システムでは LAN スイッチ以外にもルータや、ファイルサー

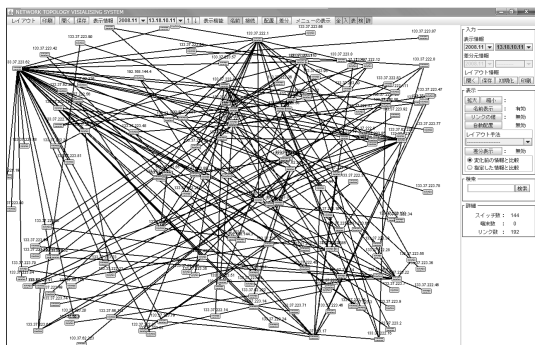


図 2 自動レイアウト前

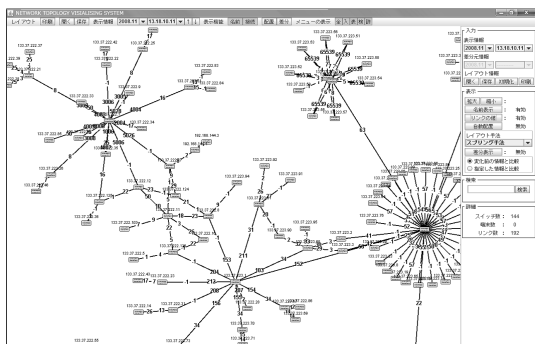


図 3 自動レイアウト後

バ, DNS サーバ, Web サーバなどもネットワーク運用を円滑に行うための重要な要素である。特にルータと DNS サーバはネットワークの運用において重要な役割を持つ。ルータは外部ネットワーク (WAN) と内部ネットワーク (LAN) の境界であり, ルータに異常が発生すると外部ネットワークとの通信が不可能となる。DNS サーバはドメイン名を IP に変換するために必要であり, DNS サーバが利用できない場合には WWW やメールなど, ドメイン名によりアクセス先を管理するシステムが利用できなくなる。そのためこれらのルータ, サーバからも情報を収集している。

本システムでは, 情報の収集はすべて SNMP (Simple Network Management Protocol) を用いている。そのため収集対象となる機器も SNMP が利用可能である必要がある。そのため, サーバには個別に SNMP エージェントをインストールする必要がある。

4.2 収集情報

上記ネットワークについて, 実際に LAN スイッチ及びルータ・サーバから収集する MIB [13]~[17] について表 1 に示す。SNMP では, インタフェースに関わる MIB のグループである Interfaces グループが実装されていなければならない。そのため SNMP が利用可能な機器からはこれら Interfaces グループの情報を取得することができる。

基本的には, 表 1 の MIB から情報を収集するのだが, いくつか例外がある。Cisco 社の LAN スイッチでは FDB を VLAN ごとに管理しており, LAN スイッチに設定しているコミュニティ名で dot1dTpFdbPort を収集しても, すべて

表 1 主に収集する MIB

名称	MIB の意味
ifType	インタフェースの種類
ifPhysAddress	インタフェースの MAC アドレス
dot1dBase BridgeAddress	LAN スイッチのブリッジ としての MAC アドレス
dot1dBase PortIfIndex	ポート番号とインタフェース 番号の対応表
dot1dTpFdbPort	FDB 情報
dot1qTpFdbPort	VLAN ごとの FDB 情報

の FDB を収集できない。また, VLAN ごとの FDB である dot1qTpFdbPort を実装していないため, VLAN ごとの FDB を得るためには SNMP のコミュニティ名に「@VLAN-ID」を付加して, dot1dTpFdbPort を取得し, VLAN ごとの FDB 情報として用いる必要がある。また, Alaxala 社の LAN スイッチでは, dot1dBasePortIfIndex が VLAN に対応していないため, privateMIB の axsVBasePortIfIndex を収集し, dot1dBasePortIfIndex の代わりとしている。NetGear 社の LAN スイッチのように, 仕様には BridgeMIB をサポートしているとありながら, SNMP による FDB の取得ができない機器も存在する。

5. 構成情報推測部

LAN スイッチ間の接続関係を決定するアルゴリズムについて説明する。接続の推測にあたって, 次のように定義する。

[定義 1] 「接続」

2つの LAN スイッチが接続関係にあるとき, これを「接続」という。

[定義 2] 「直接接続」

2つのインタフェースが1本の物理線によってのみ接続されている場合, これを「直接接続」という。直接接続でない接続を「非直接接続」という。

本アルゴリズムでは, まず機器間の「接続」関係を推測し, その後「直接接続」を絞り込む。

5.1 接続の推測

LAN スイッチ間の接続関係は, 推測を必要としない方法と FDB の状態から推測する方法の2種類がある。前者を「機器自身の MAC アドレスに基づく接続」と呼ぶ。後者は更に推測に基づく接続として, 「一方向からの接続」と「逆方向からの接続」に分けられる。

手順 1 「機器自身の MAC アドレスに基づく接続」

FDB はどのインタフェースの先にどのような機器が存在するかを記憶するデータベースであるが, FDB に保存される MAC

送信元MACアドレス	送信先MACアドレス
機器X (送信元IPアドレス)	機器Y (送信先IPアドレス)

↓

00:00:00:00:00:00	送信先MACアドレス
機器Z (送信元IPアドレス)	機器Y (送信先IPアドレス)

図4 MACアドレスとIPアドレスの変更

アドレスの中には本システムで管理対象とするルータやサーバ、更にLANスイッチ自身のMACアドレスが含まれる。しかしルータやサーバと異なり、LANスイッチは基本的には自発的な通信を行わないため、LANスイッチのMACアドレスが学習されることは少ない。そこで、本システムでは機器間で通信を行わせるためTCPの3-way Hand Shakeを利用している。TCPのイーサヘッダとIPヘッダには送信元MACアドレス、送信先MACアドレス、送信元IPアドレス、送信先IPアドレスが記述されている。このヘッダを図4のように変更する。送信元IPアドレスを機器Yと通信させたい機器Zに偽装することで、機器Yから機器Zに向かう(SYN+ACK)パケットが生成される。機器Zでは予期しないパケットであるため、機器Yに向けてRSTパケットを送信する。このやりとりにより、本来通信を行わないLANスイッチ動詞を通信させることが可能となり、FDBにLANスイッチのMACアドレス情報を蓄えることが可能となる。

このやりとりにより得られたFDBから、機器間の接続を直接取得することが可能となる。

手順2 「一方向からの接続」

手順1で多くの接続を取得可能であるが、VLANを用いてトポロジの制御をおこなっている場合、手順1で示したFDBの学習方法ではデータフレームの通過しないパスが存在する。そこでこのような接続を取得するため、機器iのインタフェースxに関連付けられFDBに注目する。図5のように機器iのインタフェースxが参照しているMACアドレスAとBが、機器jのインタフェースyとzに分散されている場合、機器iのインタフェースxから機器jへの一方向の接続があると推測される。

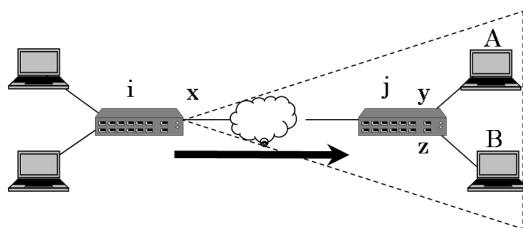


図5 一方向の接続

手順3 「逆方向からの接続」

手順1ではFDBにより参照される機器自身(以下、機器j)

のMACアドレスを基に接続を取得するが、このときMACアドレスが機器jのインタフェースと1:1対応している場合には2つの機器で互いのインタフェースが判明する。しかし、MACアドレスが機器j全体を表わしていたり、複数のインタフェースと対応していたりすると、機器j側のインタフェースが判明しない。また、手順2でも機器j側のインタフェースは考慮されていない。そこでこれらの場合に機器j側のインタフェースを推測する必要がある。これは機器jのインタフェースyのFDB情報と、機器iのインタフェースr($r \neq x$)のFDB情報に共通のMACアドレスが存在する場合に決定できる。図6はこの場合の接続例である。

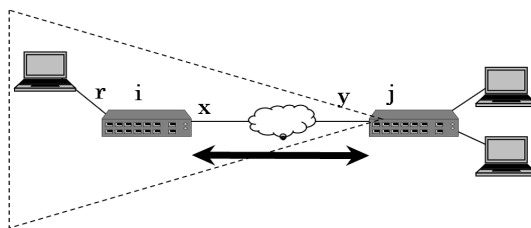


図6 逆方向からの接続

手順1から手順3を順にすべての機器の組で行う。これにより直接接続の候補となる接続をすべて取得する。次節で現在の直接接続の絞り込みの方法について述べる。

5.2 直接接続の推測

接続の推測で得られた接続関係の大部分は非直接接続である。直接接続の推測は、まず手順4で非直接接続を推測、除外し、手順5で直接接続を絞り込む。この手順4と手順5を非直接接続の除外、もしくは直接接続の決定がなくなるまで繰り返す。

手順4 「非直接接続の推測」

相互に接続関係にあるインタフェースが互いに同じ機器と接続している場合、インタフェース間に何らかの機器が存在することになる。このような接続を非直接接続とする。また、すでに直接接続が決定しているインタフェースがある場合、そのインタフェースとの直接接続以外の接続も、非直接接続とする。

図7は非直接接続となるトポロジの例である。機器iのインタフェースxと機器jのインタフェースyは接続関係にあるが、それぞれ同一の機器に接続しており、実際には直接接続でないと推測される。

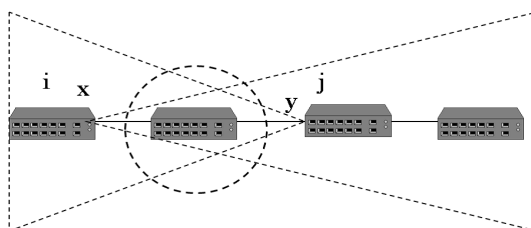


図7 非直接接続の例

手順5 「直接接続の推測」

手順4により機器iのインタフェースxが接続している機器が一つに絞られた時、この機器をインタフェースxが直接接続している機器(機器j)とする。このとき機器j側のインタフェースが判明していれば手順4に戻り、機器j側のインタフェースで直接接続でない接続をすべて除外する。

6. FDBを基にした機器の区分と推測条件の適用順序による推測結果への影響

以上の推測手順のうち機器iと機器jについて、手順1, 手順2, 手順3による接続の推測では他の機器との接続に影響を受けないため、どのような順序で推測を行ったとしても得られる接続は変化しない。しかし、手順4, 手順5における直接接続の推測では、機器i・機器jと機器kの接続関係が直接接続であるかどうか重要となる。特にそれぞれの機器の属するトポロジがメッシュ構造とスター構造に分かれているとき、非直接接続の推測を誤ることがある。そのため、メッシュ構造とスター構造を分けて推測を行い、スター構造の部分の直接接続を早期に判明させることで、メッシュ構造とスター構造が入り混じるネットワークにおいて推測の誤りを最小限に抑えることが可能となる。メッシュ構造の機器とスター構造の機器を正確に区分することは難しいため、今回はまずイーサネットワークの終端に位置するルータやサーバとLANスイッチを区別し、LANスイッチを他の管理対象機器との接続に使用しているインタフェース数に応じて区分する。これらの区分ごとに推測を行い、推測結果にどのような影響があるか調べた。

6.1 FDBを基にした機器の区分

ルータやサーバはFDBを持たない。また、STP(Spanning Tree Protocol)による隣接機器の情報の持たない。よって、これらの情報を持たないものはルータ・サーバとみなす。ルータに関しては収集するMIBにOSPF(Open Shortest Path First)などを用い、LANスイッチと区別することも考えられるが、現在のネットワークではLayer3で動作するLANスイッチも存在するため、決定的な違いにはできない。

LANスイッチの区分であるが、手順1, 手順2, 手順3による接続の推測により、他の機器との接続に使用されるインタフェースの数は判明している。この接続に使用されるインタフェース数をLANスイッチの「ランク」に設定しランクの低い順に手順4, 手順5を実行する。さらにランクはあるインタフェースの直接接続が決定するたびに減少する。これによりスター構造の末端からの推測に近い手順を行うことができる。ルータ・サーバはイーサネットワークの末端であり、つまりスター構造の末端であるためランクを'0'とし、推測の最も初期にこれらの機器の関わる直接接続を決定する。

6.2 推測条件の適用順序による推測結果への影響

図8はランクごとに直接接続を推測するときのフローチャートである。今回はランクiの接続をすべて決定してから次のラ

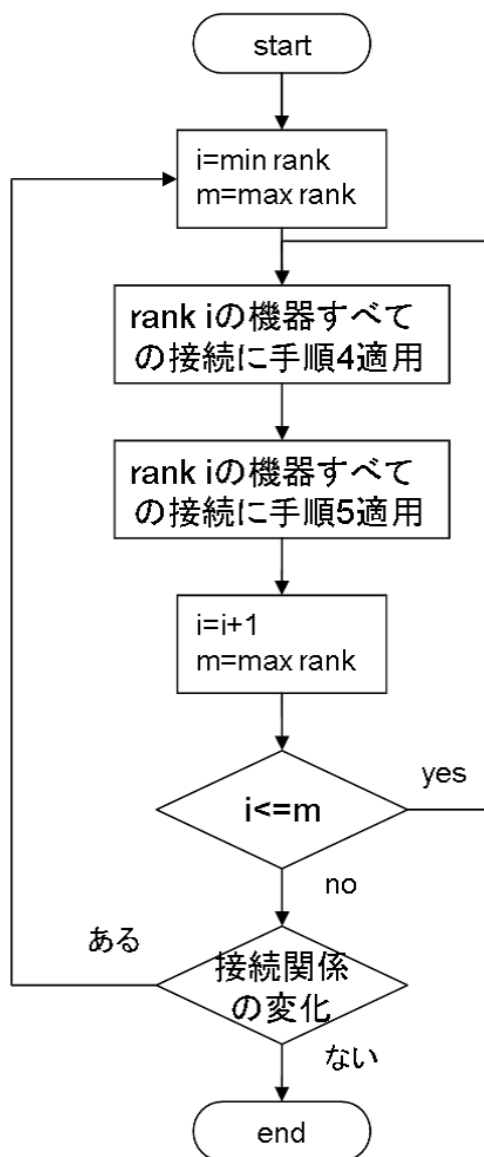


図8 ランクごとの推測手順

ンクの手順のするのではなく、ランクiの非直接接続の推測と直接接続の推測を行った後、即座に次のランクに移ることにしている。この手順で推測を行った際の実行結果であるが、従来のIPアドレスの順の推測とまったく同じ結果を得られた。一方、IPアドレスの順序を逆にして従来の推測を行った結果、多くの誤りが見られた。これは本学学内LANのIPアドレスが機器のトポロジ上の位置や隣接関係に近い形で設定されていたため、逆順の推測では推測を多く誤ることになった。ランクごとの推測で従来の推測手順と同一の結果が得られたということは、トポロジを考慮した推測がなされているといえる。

また、2次的な効果ではあるが、ランクごとの推測では推測にかかる時間を短縮できることがわかる。従来の推測では一回のループですべての接続関係について手順4, 手順5を実行していた。一方ランクごとの推測では繰り返し回数は増えたが、一度に推測する接続関係はランクが等しいもののみであるため少なく、より高速な推測が可能となった。

7. おわりに

イーサネットワークのトポロジを収集、推測し表示するシステムについて述べ、推測順序を入れ替えるための LAN スイッチの区分と、推測順序を変更した際の推測結果について述べた。

トポロジの推測アルゴリズムでは、直接接続の推測も重要であるが、接続の推測の方法も重要となる。手順 1, 手順 2, 手順 3 で推測される接続のうち、どの手順で推測されたものが対応しているかを知ることが重要となる。また、直接接続の推測を誤ることに限っては、現状での改善策とし、VLAN の情報を用い、同じ VLAN に属するインタフェース同士の接続を重視するなどが挙げられる。しかし VLAN の情報は動的には取得できず、静的に取得するにしても telnet などを用いてペンダごと収集方法を変える必要がある。VLAN の情報を効果的に収集することが今後の課題である。

文 献

- [1] 藤田俊輔, 吉田和幸, Layer2 ネットワーク構成情報表示システムにおける STP および隣接機器探索プロトコルの活用について, 情報処理学会第 1 回インターネットと運用技術シンポジウム 2008 論文集, pp.113-pp.120, 2008
- [2] 藤田俊輔, 吉田和幸, Layer2 ネットワーク構成情報推測・表示システムのための構成情報の収集-構成情報のペンダ依存性について-, 情報処理学会第 1 回 IOT 研究会 2008 論文集, pp.41-46, 2008
- [3] 藤田俊輔, 吉田和幸, Layer 2 ネットワーク構成情報推測システムの改良について, 情報処理学会火の国情報シンポジウム 2008 論文集, C-2-5, pp.1-7(published by CD-ROM), 2008
- [4] 飯田隆義, 兒玉清幸, 吉田和幸, ネットワーク構成情報表示システムのための自動配置アルゴリズムの改良について, 情報処理学会火の国情報シンポジウム 2008 論文集, C-6-1, pp.1-8(published by CD-ROM), 2008
- [5] 吉田和幸, ネットワークトポロジの発見とその表示について, 大学情報システム環境研究, VOL.10, pp.68-74, 2007.
- [6] 兒玉清幸, 釜崎正吾, 吉田和幸, ネットワーク構成情報表示システムのための自動配置アルゴリズムの評価, 情報処理学会マルチメディア, 分散, 協調とモバイル (DICOMO2007) 論文集, pp.1754-1761, 2007.
- [7] 兒玉清幸, 加来徹, 釜崎正吾, 吉田和幸, ネットワーク構成情報表示システム-自動配置アルゴリズムの改善-, 情報処理学会火の国情報シンポジウム 2007 論文集, C-8-3, pp.1-8(published by CD-ROM), 2007.
- [8] 河野優, 釜崎正吾, 大浦昇, 吉田和幸, ループを考慮した Layer2 ネットワーク構成情報の推測アルゴリズムについて, 分散システム/インターネット運用技術シンポジウム 2006 論文集, pp.7-12, 2006.
- [9] 大浦昇, 河野優, 釜崎正吾, 吉田和幸, VLAN を考慮した Layer2 ネットワーク構成情報推測アルゴリズムについて, 情報処理学会マルチメディア, 分散, 協調とモバイル (DICOMO2006) 論文集, pp.629-632, 2006.
- [10] Lowekamp, B., O'Hallaron, D.R. and Gross, T.R., Topology Discovery for Large Ethernet Networks, Procs. of Applications, Technologies, Architectures, and Protocols for Computer Communications, pp.27-31, 2001.
- [11] Breitbart, Y. Garofalakis, M. Martin, C. et al, Topology discovery in heterogeneous IP networks, Procs. of INFOCOM2000, 2000.
- [12] Son, M., Kim, B., Lee, J, Topology Discovery in Large Ethernet Mesh Networks, IEICE TRANSACTIONS on Communications, Vol.E89-B No.1, pp.66-75, 2006
- [13] Case, J., Fedor, M. Schoffstall, M. et al., A Simple Network Management Protocol, RFC 1157, <http://www.ietf.org>, 1990.
- [14] McCloghrie, K., Management Information Base for Network Management of TCP/IP-based internets, RFC 1156, <http://www.ietf.org>, 1990.
- [15] McCloghrie, K., Management Information Base for Network Management of TCP/IP-based internets: MIB-2, RFC 1213, 1991.
- [16] Decker, E., Langille, P., Rijssinghani, A. et al., Definitions of Managed Objects for Bridges, RFC 1493, <http://www.ietf.org>, 1991.
- [17] Levi, D., Harrington, D., Definitions of Managed Objects for Bridges with Traffic Classes, Multicast Filtering and Virtual LAN Extensions, RFC 4363, <http://www.ietf.org>, 2006.