BGPの更新情報を用いた障害検出について

九州工業大学 情報工学研究科
尾家研究室 寺沢 一平
ippei@infonet.cse.kyutech.ac.jp

Oie Lab

目次

- 背景、目的
- 提案手法
- 結果
- まとめ

背景

- インターネットの普及と共に品質に対する要求の向上
  - 接続性の紛失
  - 頻繁な経路変動

→ ネットワーク品質の低下

- BGP (Border Gateway Protocol) における AS (Autonoumus System)間の経路情報の監視
  - 接続状態の把握
  - 経路更新の影響範囲の把握

→ 障害の早期発見、対策、事前対策

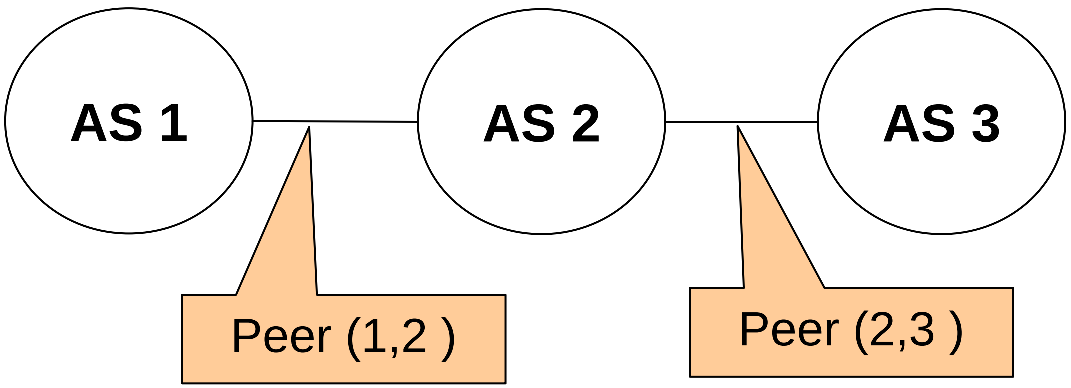
BGP

- BGP (Border Gateway Protocol)
  - ネットワークをAS (Autonoumus System)で分割
  - 隣接するAS間で経路情報の交換
- 経路更新
  - セッションを確立時、互いの経路情報の交換
  - 経路情報交換後は差分情報のアップデート
  - 経路更新情報の通知は1日およそ20万回程

大量の経路更新情報の中から影響の大きいイベントに対して実際に経路変動が発生している
箇所や状況の把握は困難である。

用語の定義

- Peerとは
  - 本研究ではBGP情報を直接交換するAS同士の接続をPeerと呼ぶ。



研究の目的 (1)

- BGP 計測の目的
  - 経路更新メッセージや計測用パケットの遅延情報から Peer に着目したネットワークの状態を把握する
  - 計測対象のBGPルータから得られる情報のみを使って運用上有用な情報を抽出する
- 項目
  - 障害情報やその発生源となる Peer の推定
  - パケットの遅延変動と経路変動の関係の調査
  - リアルタイム化及び可視化
  - 障害時に対するアラート機能

研究の目的 (2)

- 本年度の目標と課題
  - 経路更新情報の格納
    - 情報の収集
    - データベースへ保存等
  - 運用上大きな問題となる障害Peerの推定
    - 更新メッセージをクラスタ化し
      前後関係を比較して発生源の範囲を絞り込み
    - 経路変動による影響度の数値化

目次

- 背景、目的
- 提案手法
- 解析結果
- まとめ

環境、前提条件

- 環境
  - 使用ツール
    - BGPView(http://www.bgpview.org/software/bgpview/)
    - 本研究ではインテック･ネットコア様にご協力頂きBGPルータからの更新情報を取得
  - 取得情報
    - 経路更新情報
      - 更新のタイムスタンプ
      - 更新の種類(UPDATE or WITHDRAW)
      - Prefix
      - ルータのアドレス
      - AS Path

手法手順

- BGPの経路情報やAS Pathの扱い方
  - AS隣接関係の2次元テーブル化

- 各Peerの稼働率 (Availability)
  - BGPのPeerが経路上で使用されている時間と使用されていない時間の傾向
- ネットワークへの影響(Impact)
  - BGPのPeerの変化がネットワーク与える影響

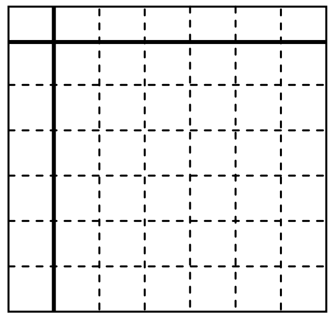
AS隣接関係のテーブル化

- AS Pathより見えるAS同士の接続を二次元テーブルで表現
  - トポロジの表現の簡単化
    - AS隣接関係の簡単化

経路情報

```
2005/07/16 08:55:47: WITHDRAW: 65.198.40.0/22
2005/07/16 08:55:47: WITHDRAW: 65.214.156.0/22
2005/07/16 08:55:47: WITHDRAW: 194.85.84.0/22
2005/07/16 08:55:47: WITHDRAW: 65.244.246.0/23
2005/07/16 08:55:48: UPDATE  : 65.198.40.0/22     192.41.187.26   7682 2497 3549 32723
2005/07/16 08:55:48: UPDATE  : 65.214.156.0/22    192.41.187.26   7682 2497 3549 32723
2005/07/16 08:55:48: UPDATE  : 216.55.240.0/20    192.41.187.26   7682 2497 3549 32723
```

二次元テーブル

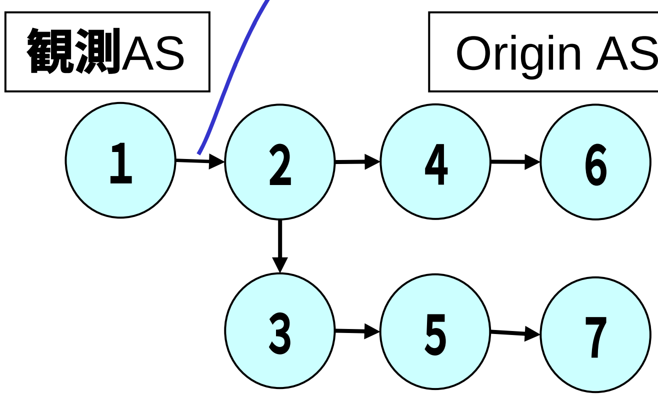


経路－２次元テーブル化の例(1)

経路状態1

P6：１２４６
P7：１２３５７

仮定：この2経路をフルルート情報とする



テーブル化

2次元テーブル

source \ destination	1	2	3	4	5	6	7
1		○					
2			○	○			
3					○		
4						○	
5							○
6							
7							

- ASの経路の２次元テーブル化
  - 経路状態1をテーブル化すると右の２次元テーブルのようになる。

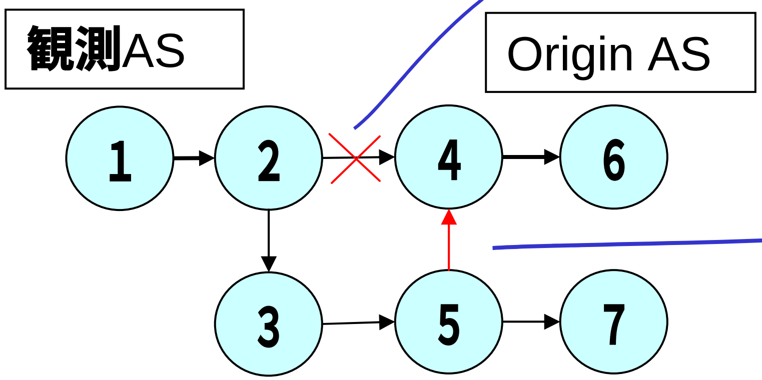
経路－2次元テーブル化の例(2)

更新

P6 : 1 2 3 5 4 6

経路状態2

P6 : 1 2 3 5 4 6
P7 : 1 2 3 5 7



テーブル化

2次元テーブル

source \ destination	1	2	3	4	5	6	7
1		○					
2			○	×			
3					○		
4						○	
5				○			○
6							
7							

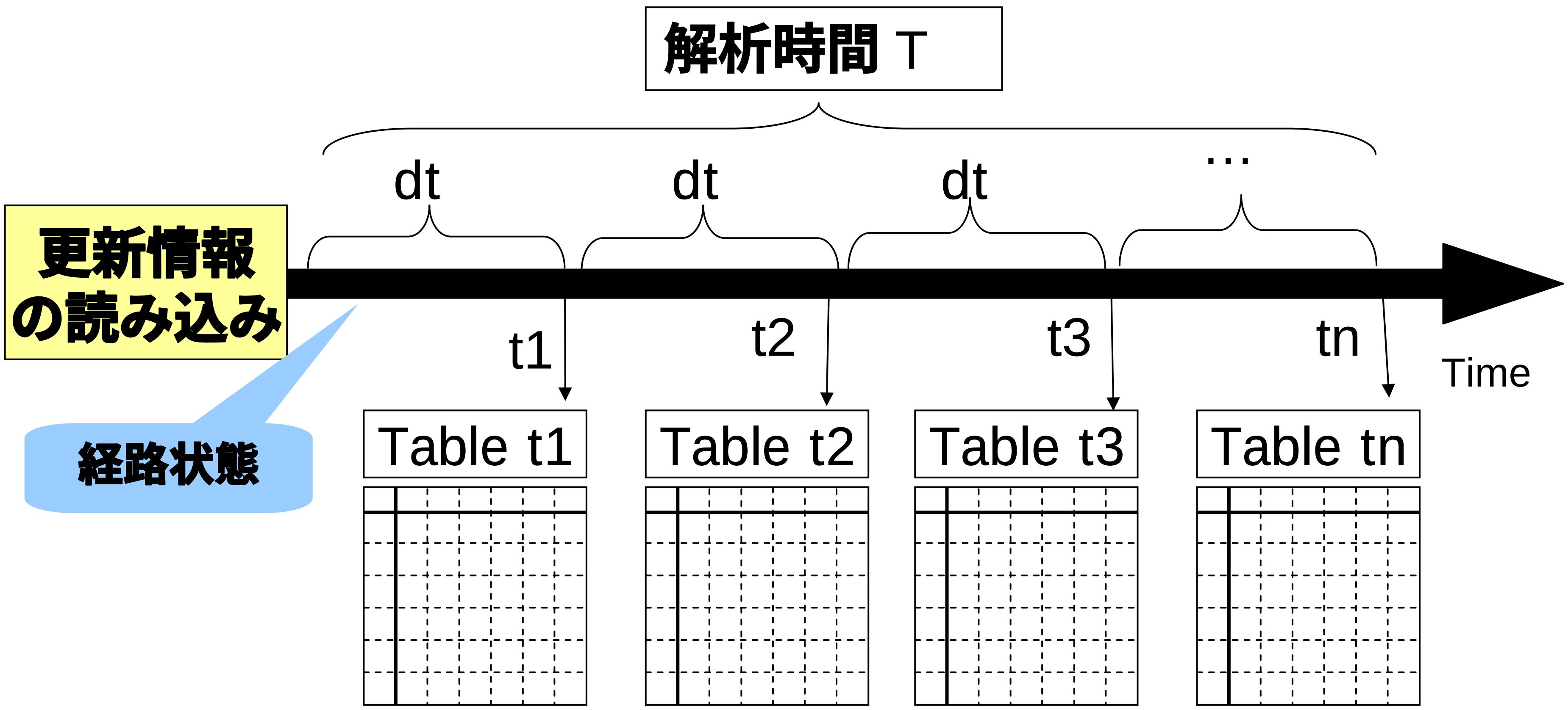
消失となったPeer

発生したPeer

- **更新により経路状態が変化**
  - **変化したテーブルと変化前のテーブルの比較**

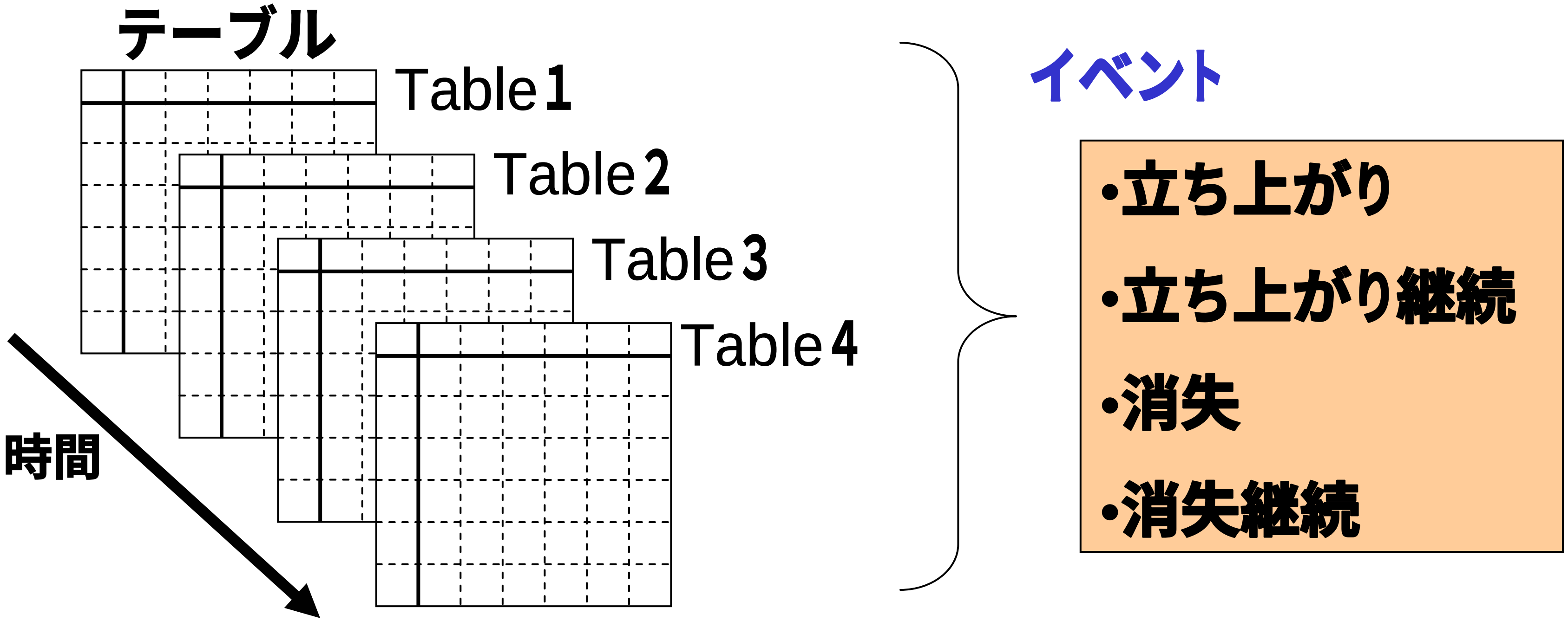
差分情報から変化箇所の特定

2次元テーブル作成のタイミング



- 更新情報を時間順に読み込み経路テーブルを更新
- 任意で決定した一定時間dt毎にテーブルを作成

Peerの消失と立ち上がり(1)



- 作成した複数のテーブル情報の前後関係を比較
  - 「立ち上がり」、「消失」、「立ち上がり継続」、「消失継続」
    4つの状態が抽出される

⇒ これらをPeer上のイベントと定義

Peerの消失と立ち上がり(2)

- Peerのイベントから得られる情報
  - Peer毎のMTBF (Mean Time Between Failure) と MTTR (Mean Time To Repair)、稼働率(Availability)
    - Peerの接続時間と切断時間から稼働率を導出
  - イベントのネットワークに対する影響
    - 各イベントが管理対象ネットワークに対して どれだけの影響を与えるかを把握

これらの情報を数値化し、
Peerの不安定指標とする

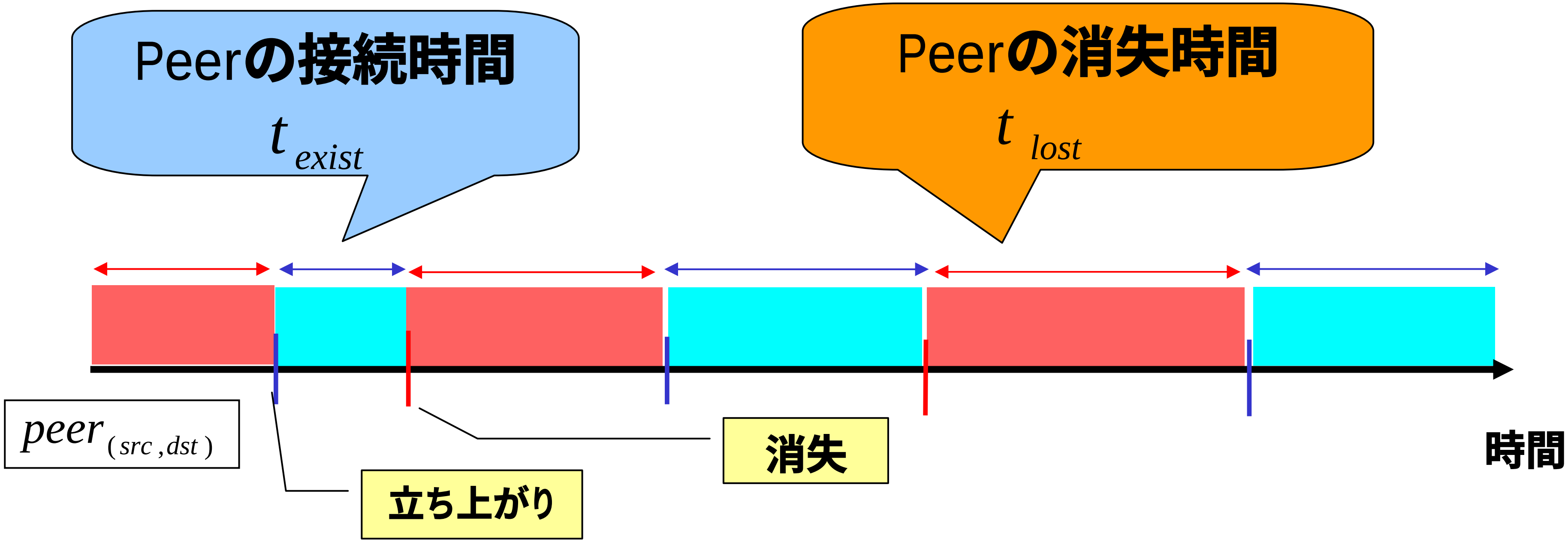
手法手順

- BGPの経路情報やAS Pathの扱い方
  - AS隣接関係の2次元テーブル化

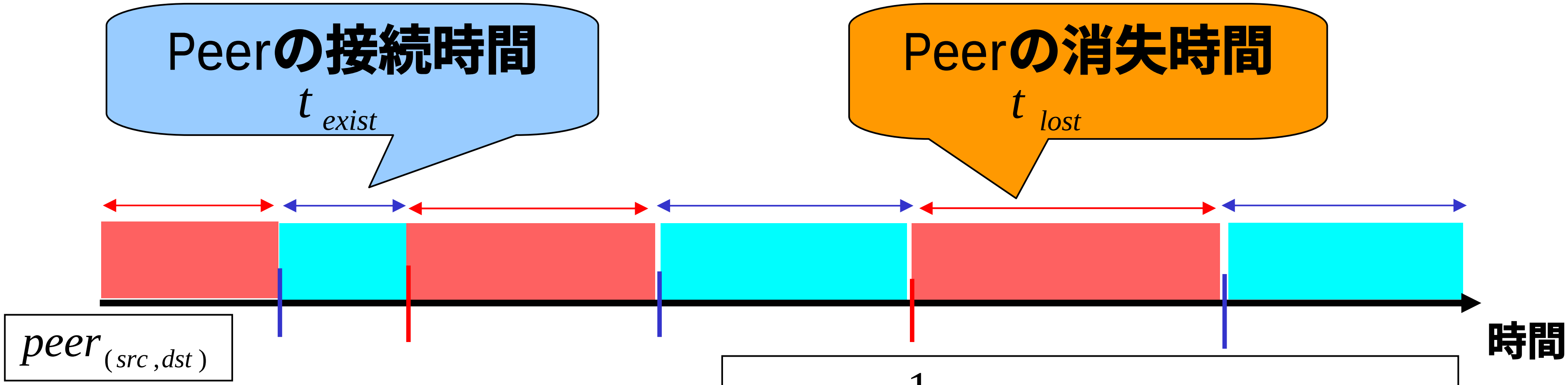
- 各Peerの稼働率 (Availability)
  - BGPのPeerが経路上で使用されている時間と使用されていない時間の傾向
- ネットワークへの影響(Impact)
  - BGPのPeerの変化がネットワーク与える影響

PeerのMTBFとMTTR

- MTBFとMTTRの導出の為に
  - 立ち上がりのイベントと消失のイベントの組み合わせから接続を継続している時間と消失している時間が求められる



PeerのMTBFとMTTR



消失がm回
観測された場合。

$$MTBF = \frac{1}{m}\sum t_{exist}$$

$$MTTR = \frac{1}{m}\sum t_{lost}$$

$$稼働率(Availability) = \frac{MTBF}{MTBF + MTTR}$$

- 図のような接続時間 t_{exist} と消失時間t_{lost}により MTBFとMTTRの値を導出
  - MTBFとMTTRの値からPeerの稼働率が求められる。

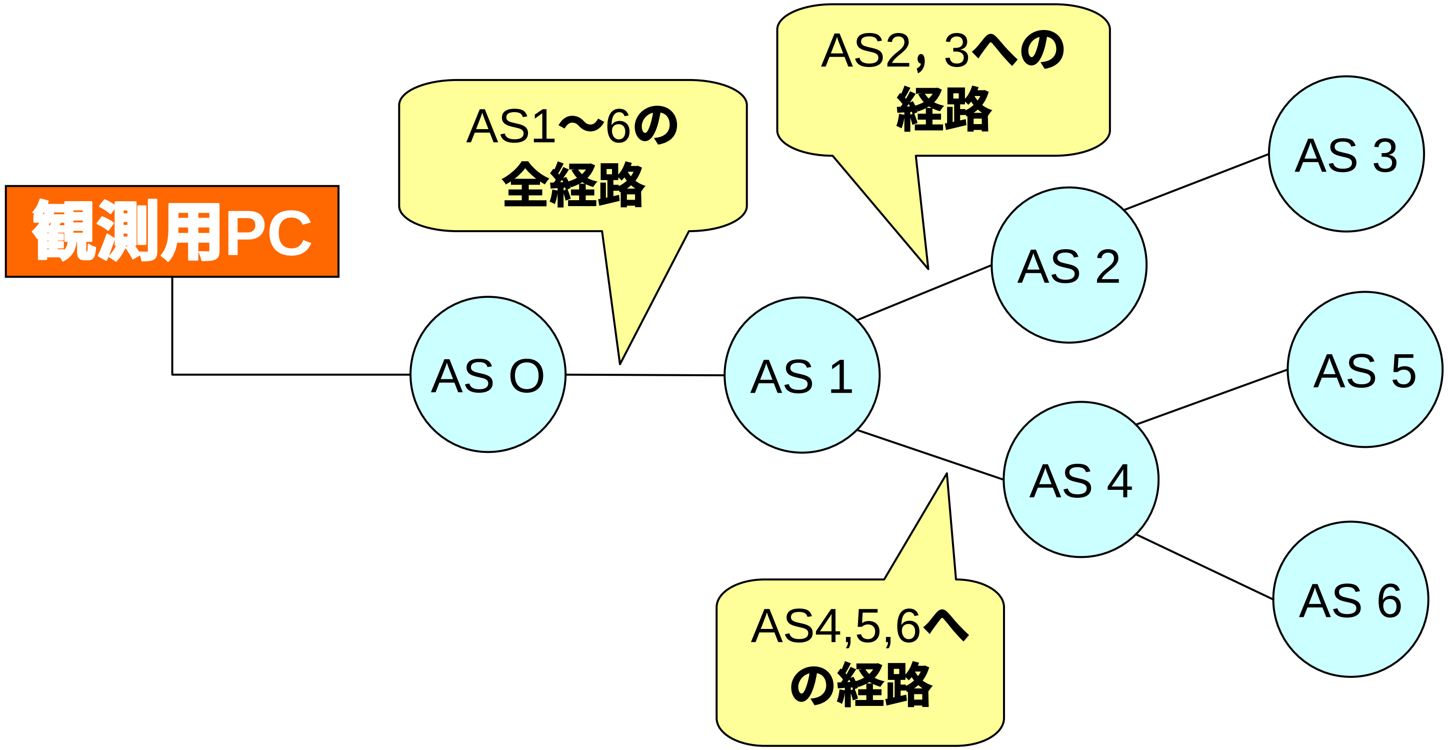
手法手順

- BGPの経路情報やAS Pathの扱い方
  - AS隣接関係の2次元テーブル化

- 各Peerの稼働率 (Availability)
  - BGPのPeerが経路上で使用されている時間と使用されていない時間の傾向
- ネットワークへの影響(Impact)
  - BGPのPeerの変化がネットワーク与える影響

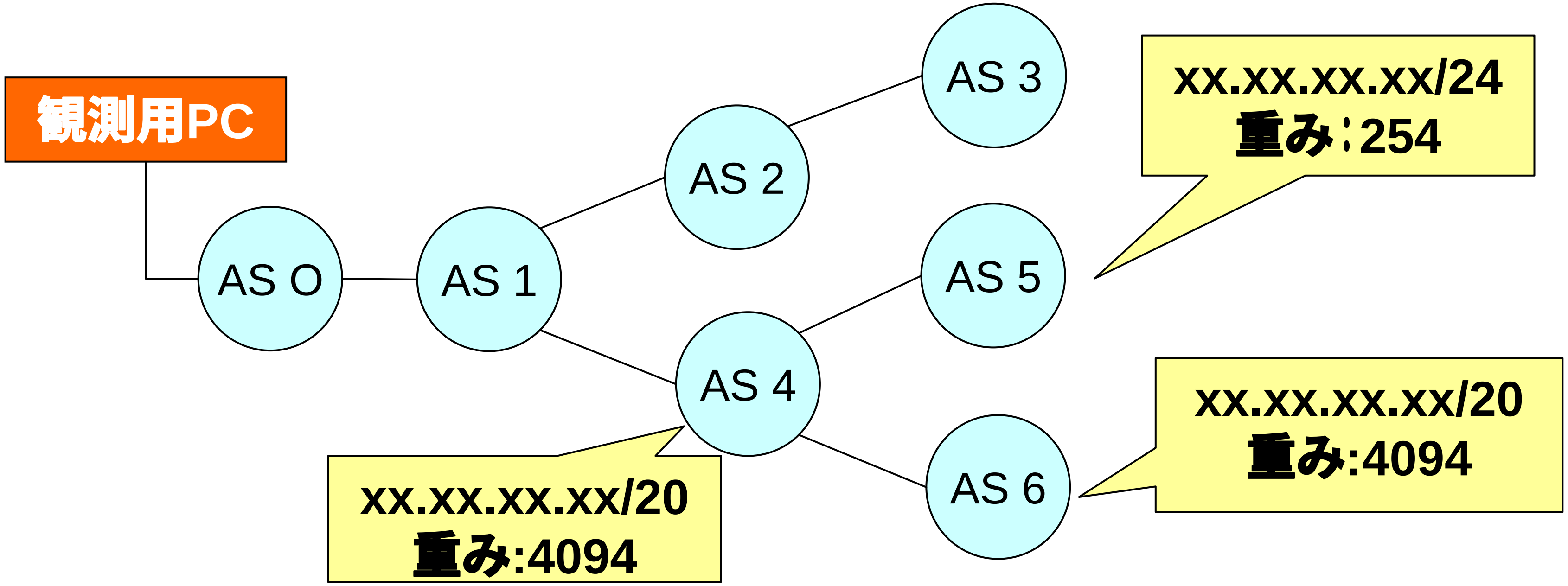
イベントのネットワークに対する影響

- 観測点に対するインパクト
  - 『AS O』を観測ASとする。
  - 各Peerを経由する経路は下のようになる。



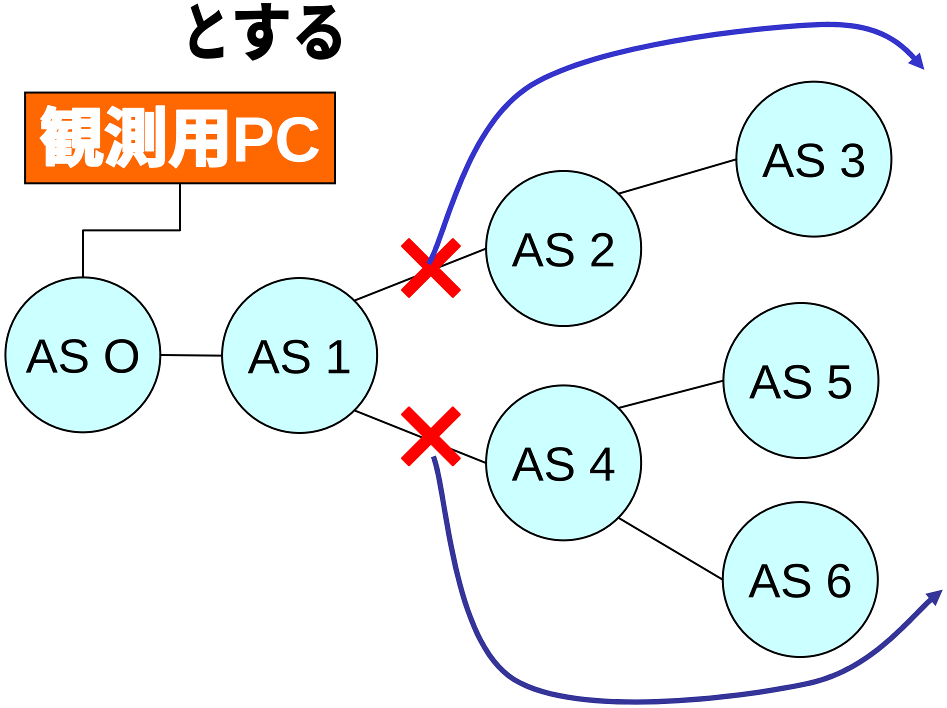
イベントのネットワークに対する影響

- アドレス数による経路の重み
  - 広告された経路のPrefixによって規模がそれぞれ違う
  - そこでPrefixのアドレス数から重み付けを行う



イベントのネットワークに対する影響

- Impact (影響)の値の導出例
  - 『AS O』が持っているアドレスの総数を母数とし、Peerが切断された時などに観測点から見たネットワークに対するインパクト(影響) を導出
  - AS1が/16、AS2,AS3がそれぞれ/24のネットワークを持っているとする



$$impact_{(1,2)} = \frac{weight_{(1,2)}}{\sum weight_{o}}$$

$$= \frac{254 + 254}{74484}$$

$$= 0.007$$

目次

- 背景、目的
- 提案手法
- **解析結果**
- まとめ

解析結果

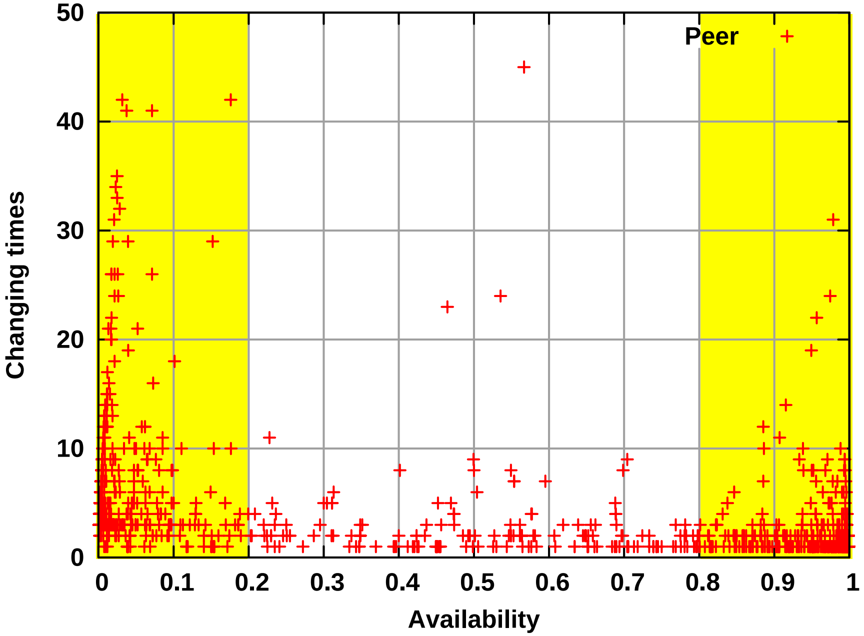
- 大手町停電による大規模通信障害
  - 2004年5月31日 15時　～
    2004年6月1日 3時
    (http://internet.watch.impress.co.jp/cda/news/2004/06/01/3314.html)
    - 関東一円の企業向け通信サービス約2万回線
    - 関東甲信越地域のOCNサービス約15万回線
    - 「HOTSPOT」約1,000拠点に関しては全国規模でのサービス停止
- 解析ルータ
  - ルータ1:エッジＡＳ
  - ルータ2：トランジットＡＳ
- 解析粒度(dt=) 1分

解析結果

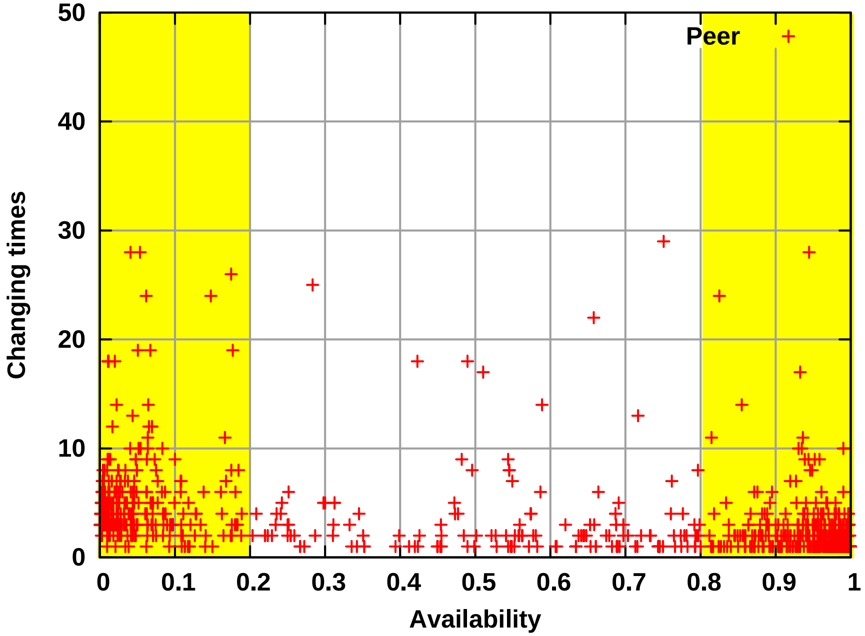
- BGPの経路情報やAS Pathの扱い方
  - AS隣接関係のテーブル化

- 各Peerの稼働率 (Availability)
  - BGPのPeerが経路上で使用されている時間と使用されていない時間の傾向
- ネットワークへの影響(Impact)
  - BGPのPeerの変化がネットワーク与える影響

Availabilityーイベント検知回数 (2004/5/30)



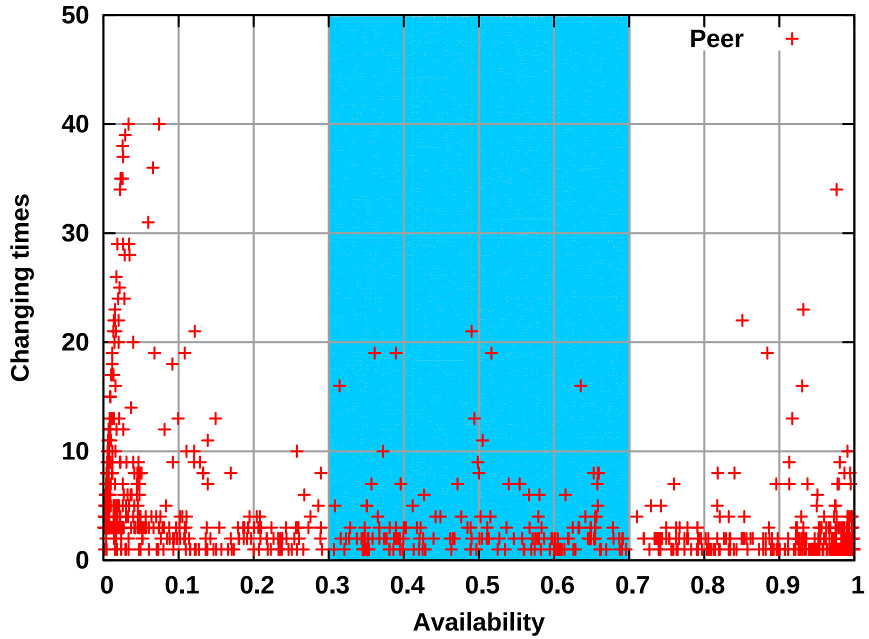
ルータ1 (2004/5/30)



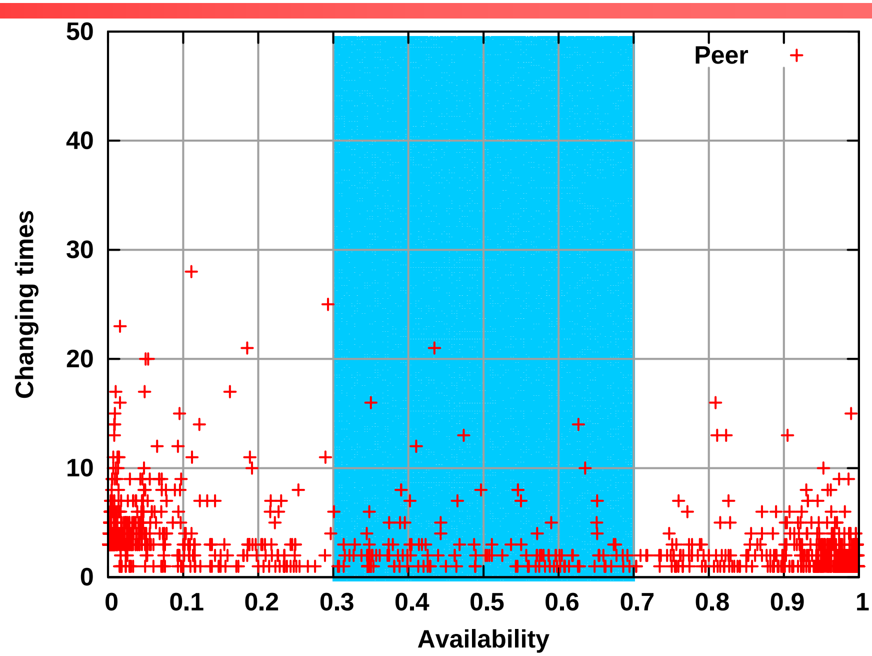
ルータ2 (2004/5/30)

- 稼働率(Availability)とイベント検知回数
  (Y軸 消失と立ち上がりイベント検知回数：X軸 Availability)
  - 2004年5月30日で平常運用された日の分布
  - 稼働率が0や1の近傍にPeerが集中

Availabilityーイベント検知回数(2004/5/31)



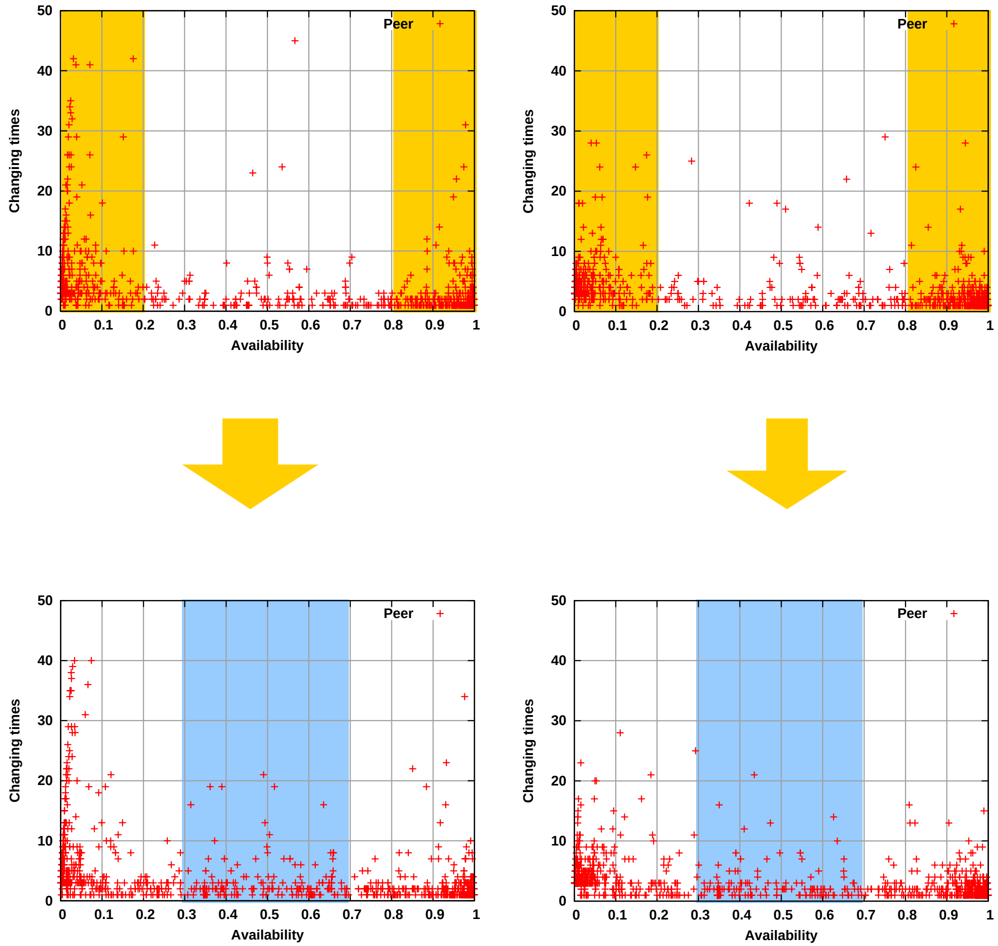
ルータ1 (2004/5/31)



ルータ2 (2004/5/31)

- 稼働率(Availability)とイベント検知回数の図
  (Y軸 消失と立ち上がりイベント検知回数：X軸 Availability)
  - 障害発生日の2004年5月31日のデータ
  - 30日と比較すると中央の印の箇所に
    イベント回数の増加傾向がある

Availabilityーイベント検知回数



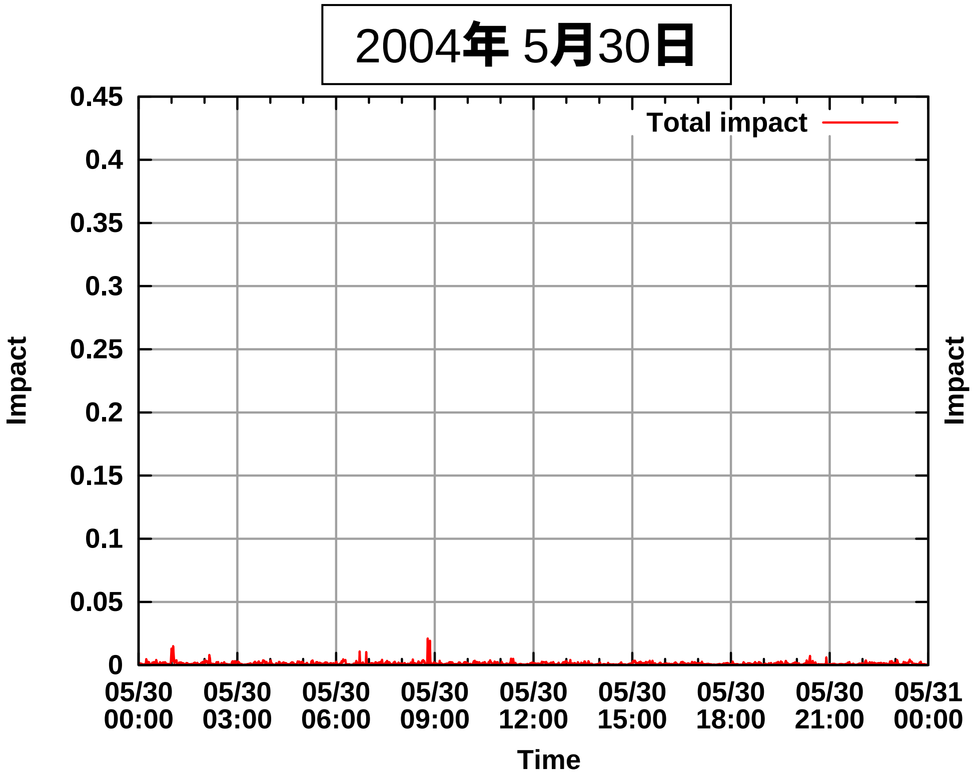
- 30日と31日の比較
  - 稼働率 0.5 近辺の値が増加傾向
  - 指標となる明確な変化は見られず

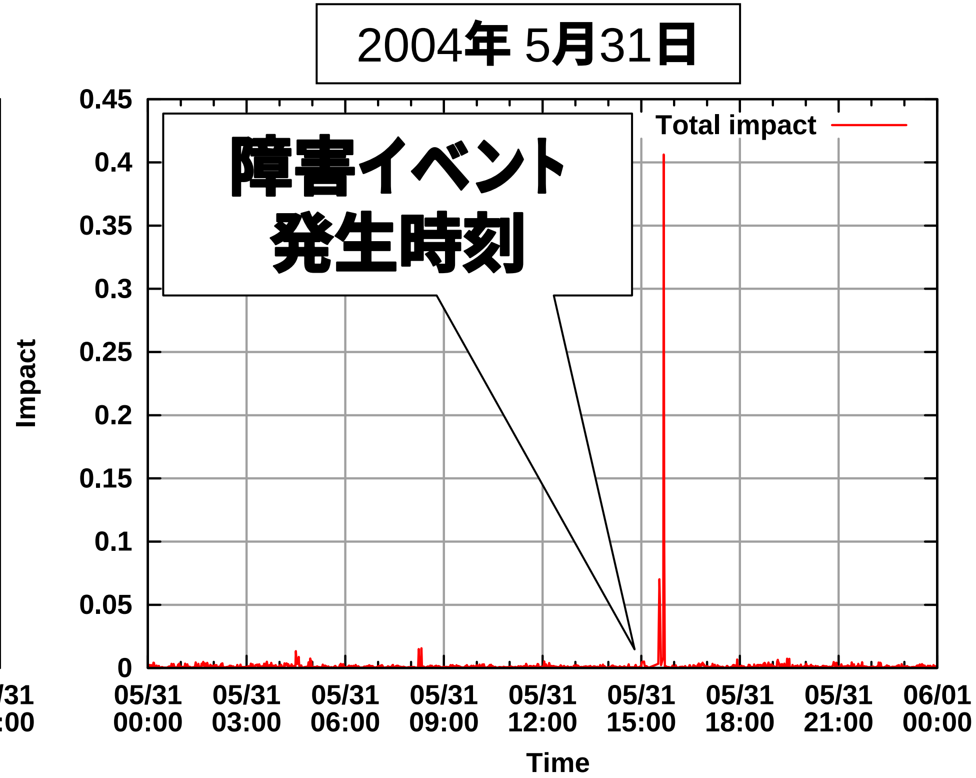
指標として扱う為には更に検討が必要

解析結果

- BGPの経路情報やAS Pathの扱い方
  - AS隣接関係のテーブル化
- 各Peerの稼働率 (Availability)
  - BGPのPeerが経路上で使用されている時間と使用されていない時間の傾向
- ネットワークへの影響(Impact)
  - BGPのPeerの変化がネットワーク与える影響

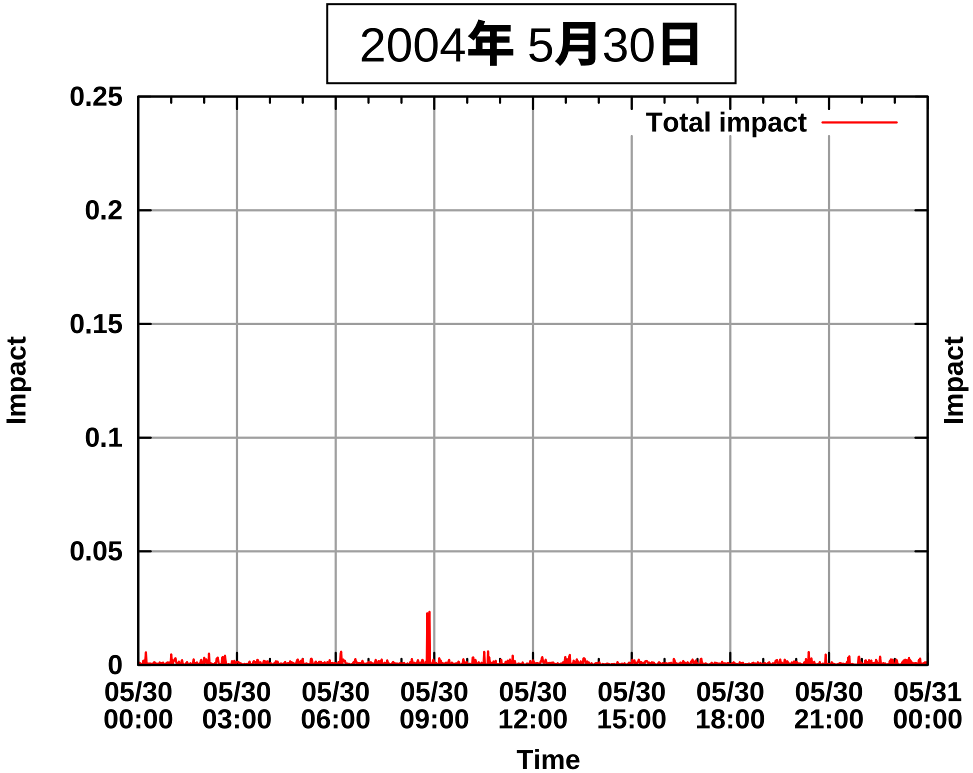
インパクトの変動特性 (ルータ1)

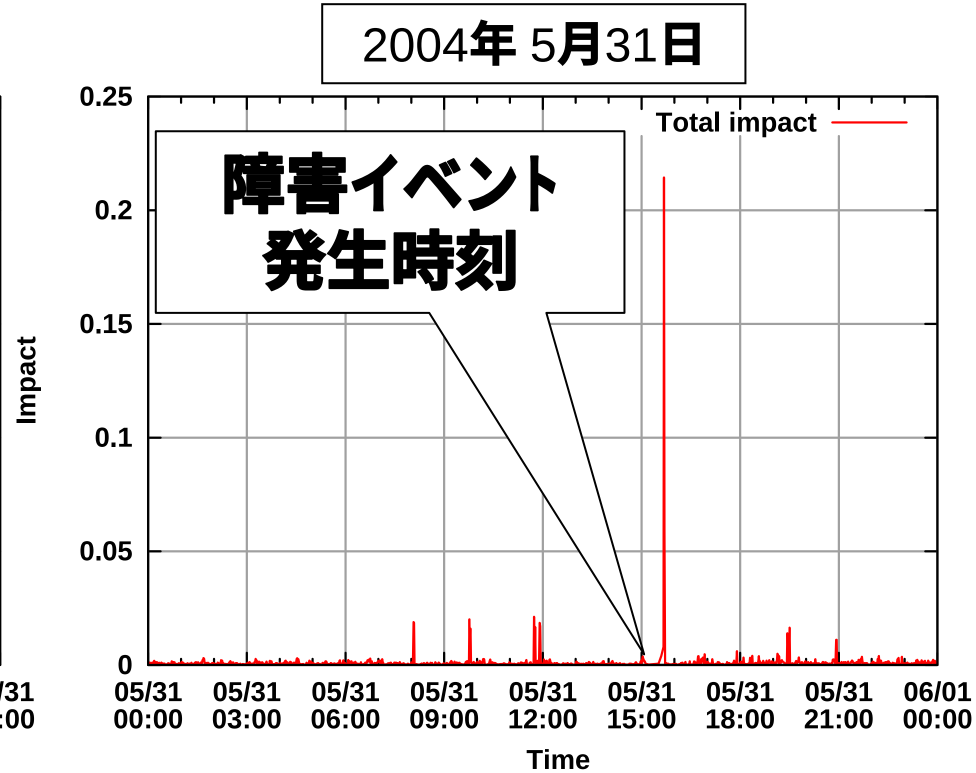




- **ルータ1のインパクト変動特性**
  (Y軸 影響(インパクト) ：X軸 時間(タイムスタンプ))
  - ルータ全体で検知されたPeerのインパクトの数値の総和をY軸に取っている。
  - 停電の障害が起きた時間帯(31日 15:30以降)に大きなインパクトを検知

インパクトの変動特性 (ルータ2)

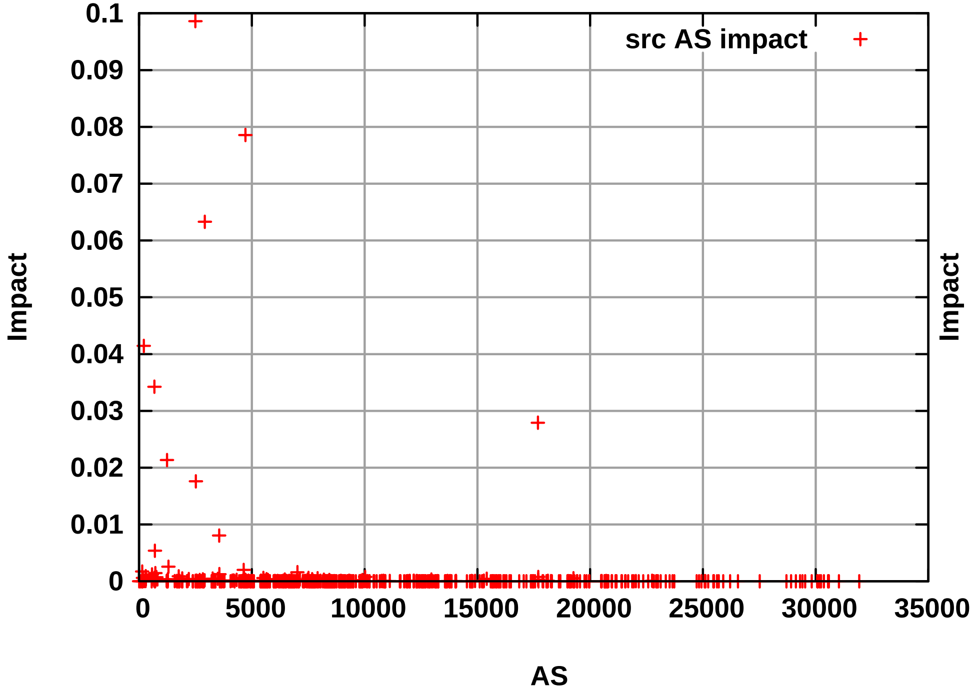




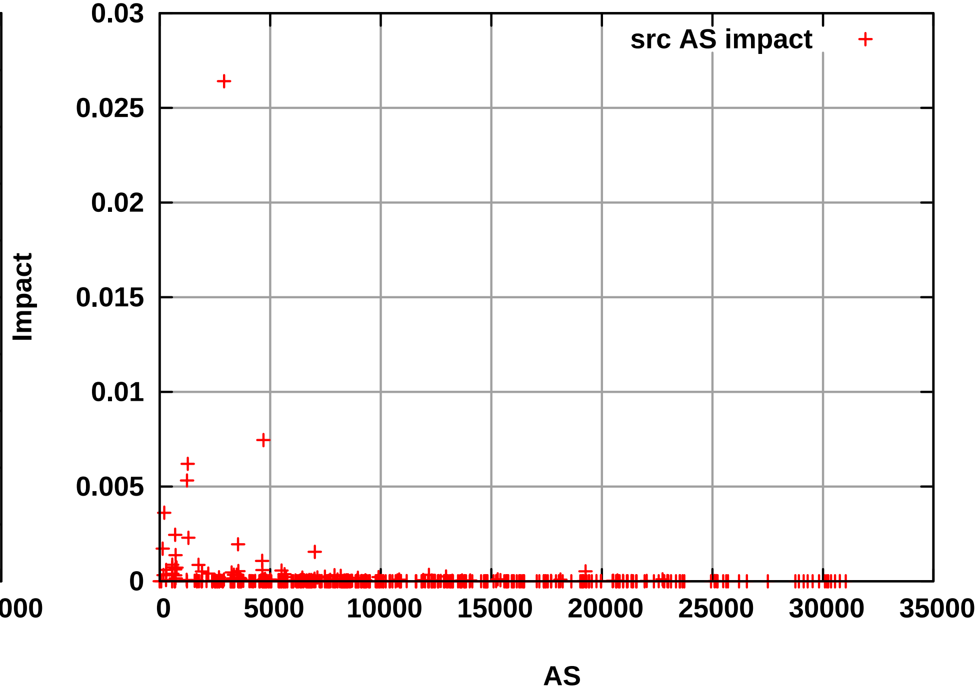
- **ルータ2のインパクト特性**
  (Y軸 影響(インパクト) : X軸 時間(タイムスタンプ))
  - ルータ1と同様に障害時に大きなインパクトを検知
  - エッジASのルータ1に比べインパクトの絶対値が低い
    - ルータ2はルータ1に比べて大きな影響を受けないBest経路を使っていたと思われる

AS毎インパクト

ルータ1のAS別インパクト



ルータ2のAS別インパクト



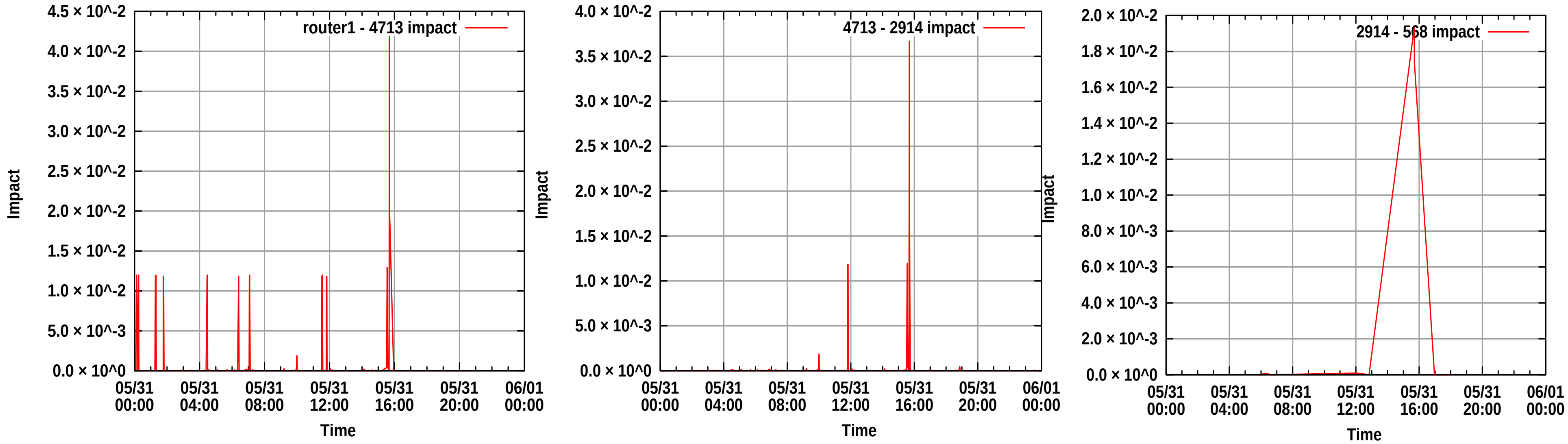
- 2004/5/31 15:00 〜16:00 間のAS毎のインパクト
  (Y軸 影響(Impact)：X軸 AS番号)
  - PeerでSRC_ASとなっているASのインパクトの総和
  - いくつか特定のASにインパクトが集中している事が判る

AS毎の影響(31日15:00～16:00)

ルータ1	
AS	Impact
2497	0.0986
4713	0.0785
2914	0.0632
209	0.0414
17685	0.0279
1239	0.0213
2516	0.0176
3549	0.00806
701	0.00537

ルータ2	
AS	Impact
2914	0.0264
4694	0.00745
1273	0.00620
1239	0.00532
209	0.00362
701	0.00245
1299	0.00229
3549	0.00195
7018	0.00155

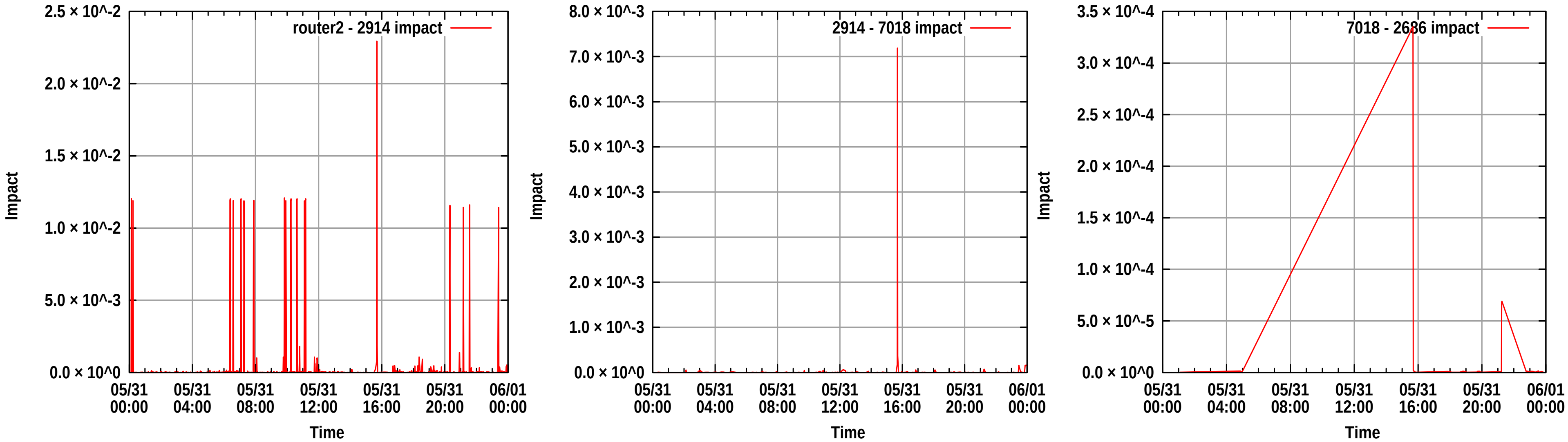
ルータ1
Peerのインパクト変動特性



- ルータ1の各Peerにおけるインパクトの変動
  (Y軸 impact(影響)ー X軸 時間)
  - 観測ASに繋がるPeerから障害時のインパクトを追う

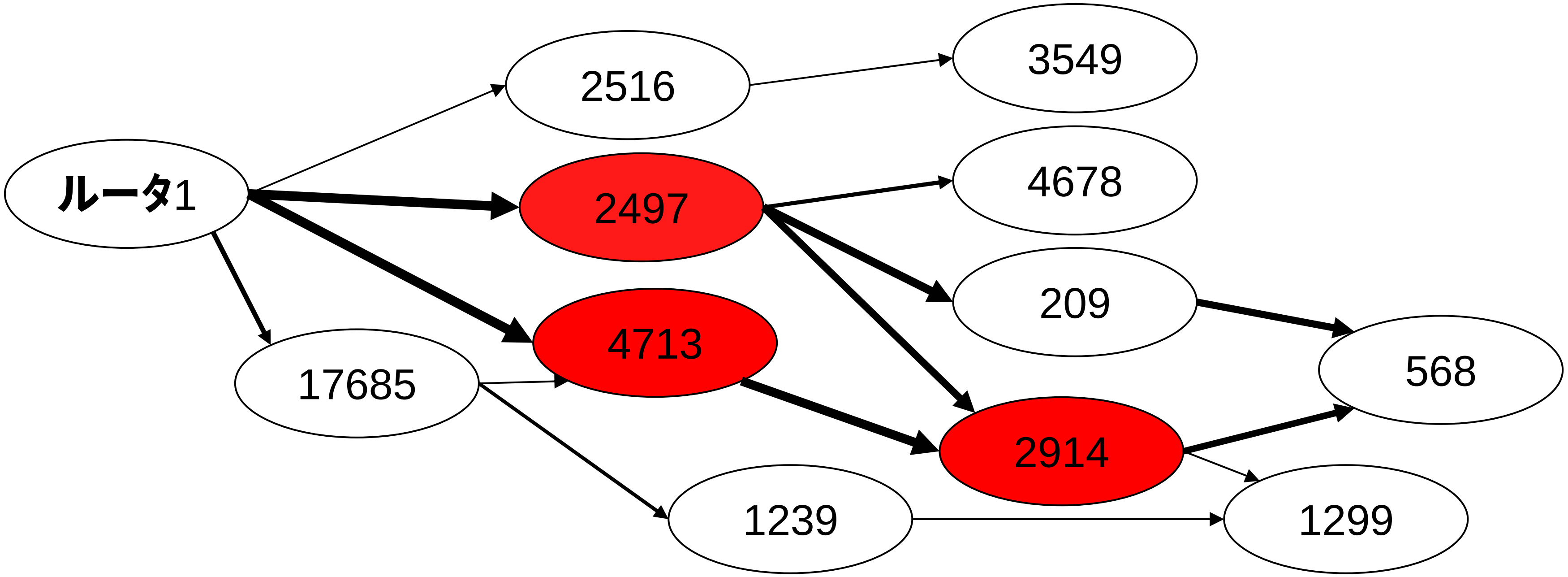
各Peerのインパクト特性を追跡し
影響範囲を絞る

ルータ2
Peerのインパクト変動特性



- Peerにおけるインパクトの変動
  (Y軸 impact(影響)– X軸 時間)
  - 同様にルータ2のPeerのグラフを示す

インパクト値とBGP経路



- 31日15:00 頃のインパクト値の高い AS や Peer から大きい影響を受けたBGPの経路図が作られる
  - 上図はルータ1からの情報を追跡した

➡ 障害等の影響範囲やおおよその箇所を推測し回避可能

目次

- 背景、目的
- 提案手法
- 結果
- まとめ

まとめ

- MTTRとMTBFによる稼働率(Availability)とイベント検知回数
  - 障害発生日(31日)に稼働率 0.5前後に消失と立ち上がりのイベント検知回数の増加傾向が見られたが、明確な変化は検知されず

  ➡ **今後主成分分析をかけるなどの検討が必要**

- 影響(Impact)値
  - ルータにとって影響の大きいイベントの有無と時間帯が判明
  - Peerの変動とルータ全体の変動を比較により影響を受けているPeerやASを区別する事が可能

  ➡ **ネットワークの状態把握や障害等のイベント検知に利用可能**

問題点、課題点

- 更に多くの具体例
  - 障害の危険性を表せる指標の決定

→ 更に解析を行い
障害時の数値的特性を分析

- リアルタイム性
  - 現在は過去の情報を解析する機能のみ

→ bgpviewのようなBGPと直接通信出来る
ツールに組み込みリアルタイム情報を取り扱う

Next Step
PingのRTT情報との比較

- 本研究ではBGPの経路情報から障害情報の検出を行った
  - パケット遅延情報を指標に加える
    - BGPの経路変動と遅延情報の関係
    - インパクト値の信頼性評価

⇒ 次のステップとして今回解析されたデータと遅延情報の関係を調べる。

最後に

- オペレータの目から見て有用な情報の有無
  - 本研究の中でオペレーティング上で有効と思える点
  - 同様に解析に改善を加えた方がよいと思える点

- 更に解析を
  - 本件では大手町の大停電について述べましたが
    他に解析を試すべき日など有れば

皆さんのご意見を待っています