Operators' Life with no-Sampling Flow Data
～通信品質，通信関係の計測～

2007/7/12

NTT 情報流通プラットフォーム研究所

石橋　圭介

ishibashi.keisuke@lab.ntt.co.jp

はじめに

- 攻撃の検出，特定はできた
- もう少しディープに調べたい

- 攻撃の品質への影響度合い/範囲は？
  - そもそも影響がなければ攻撃対処する必要もない？

⇒品質計測

- 攻撃の指示元は？
  - 根っこから対処することは可能か？

⇒ホスト間通信関係構造観測

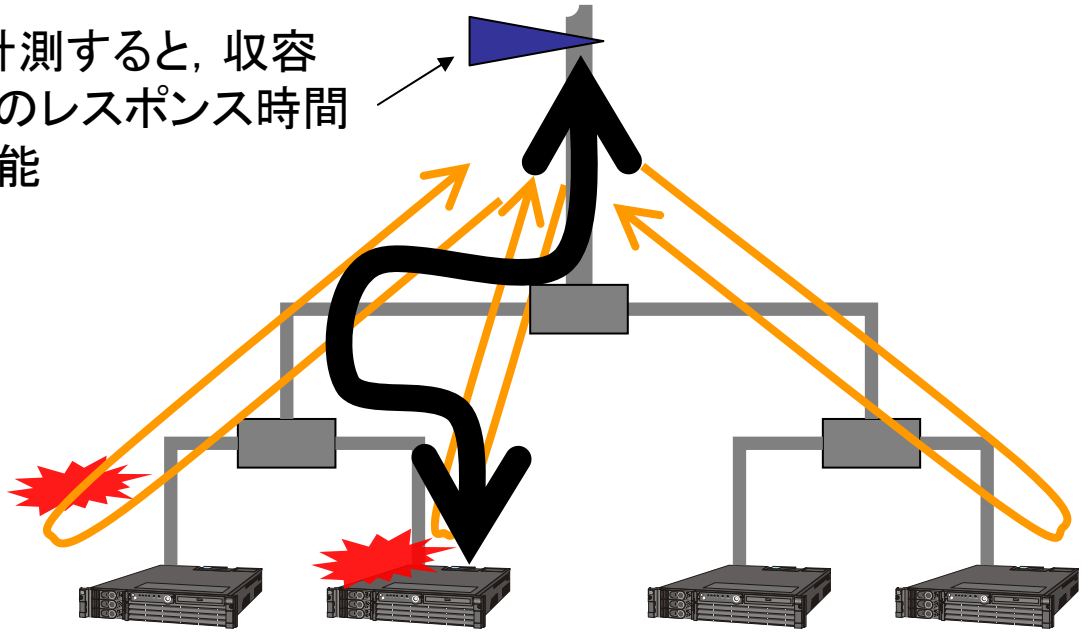
- いよいよサンプルでは無理な領域

通信品質計測

サーバレスポンス時間計測

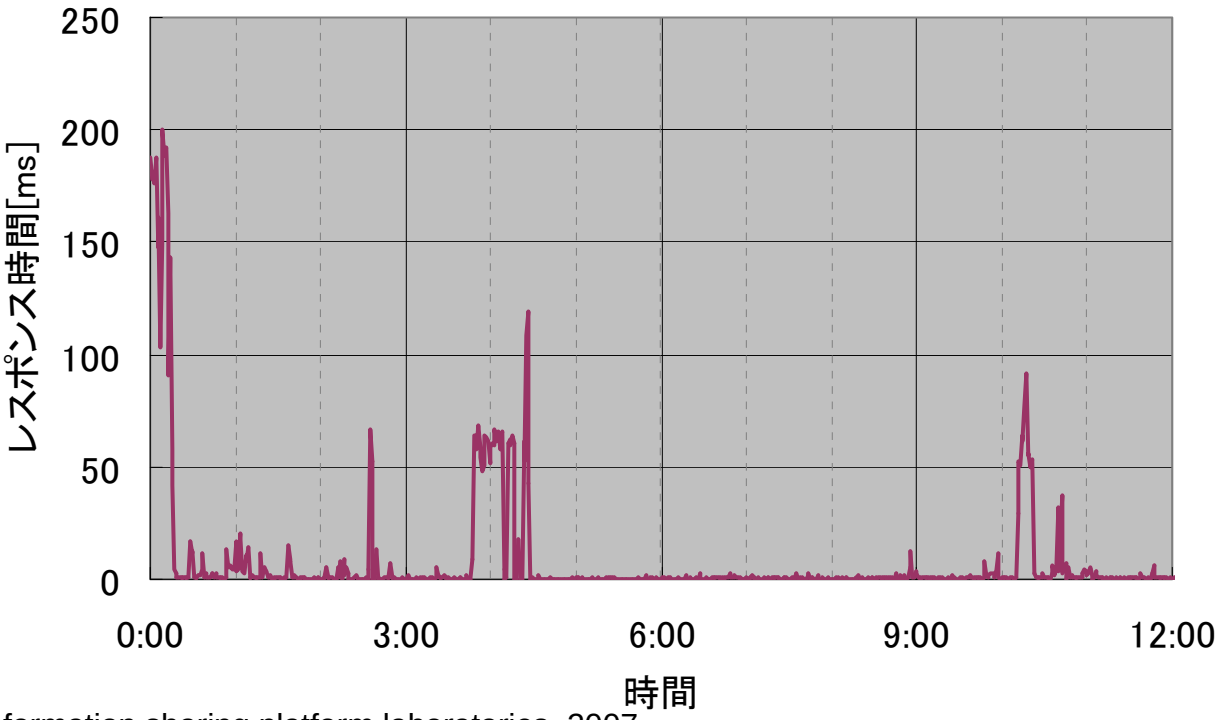
品質劣化とトラフィック増大の関係を見たい

- 大量トラフィックの影響
  - 当該サーバに影響があるか，近隣サーバにも影響は出ていないか？
- 影響⇒レスポンス時間
  - サーバ上位リンクでパケットをサンプル無しでモニタし，上り下りパケットタイムスタンプのマッチングで計測可能
    - ランダムサンプリングすると計測困難
    - 特殊なサンプリング（ハッシュサンプリング）が必要



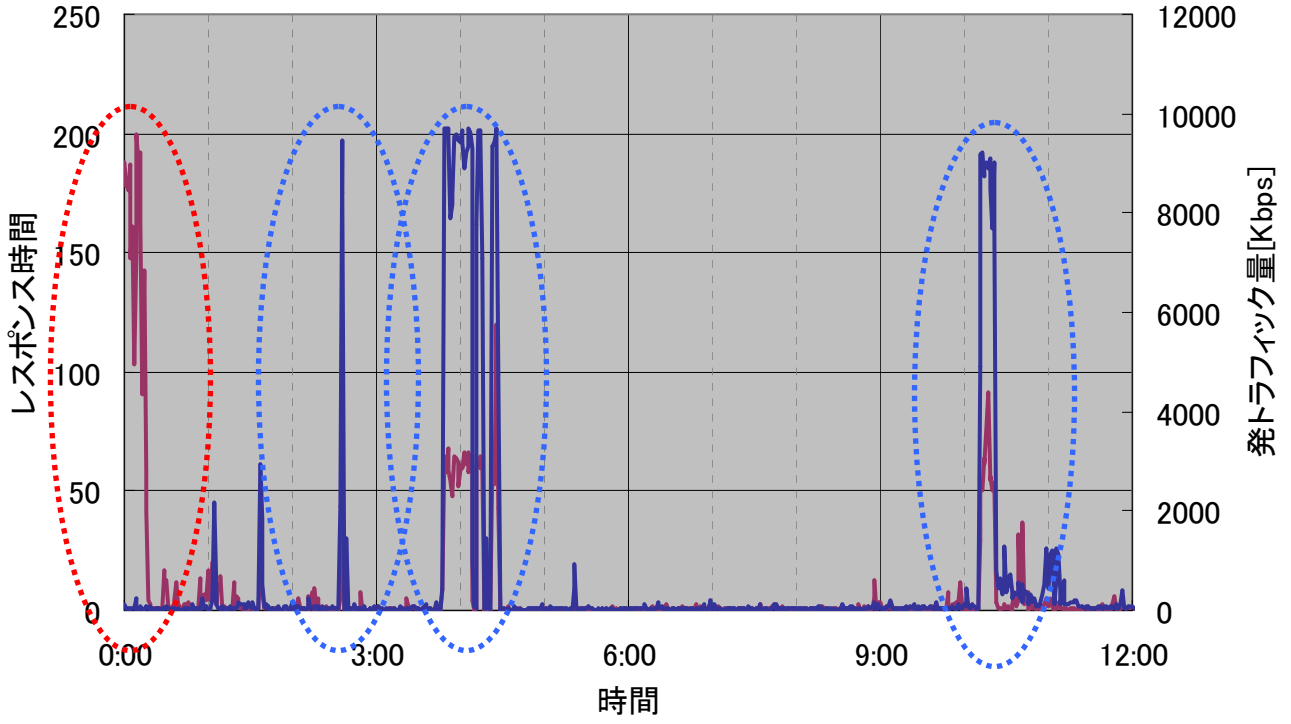
サーバレスポンス時間計測例

- サーバ行きSYNパケットとサーバ発Ackパケットのペアから往復遅延を計測
  ≒サーバレスポンス時間の計測
  - 実際のユーザパケット（≠ping）の品質計測が可能.
- 時系列でみると，ところどころ劣化している
  - 地球の裏側まで133msで往復できるはずなのに…



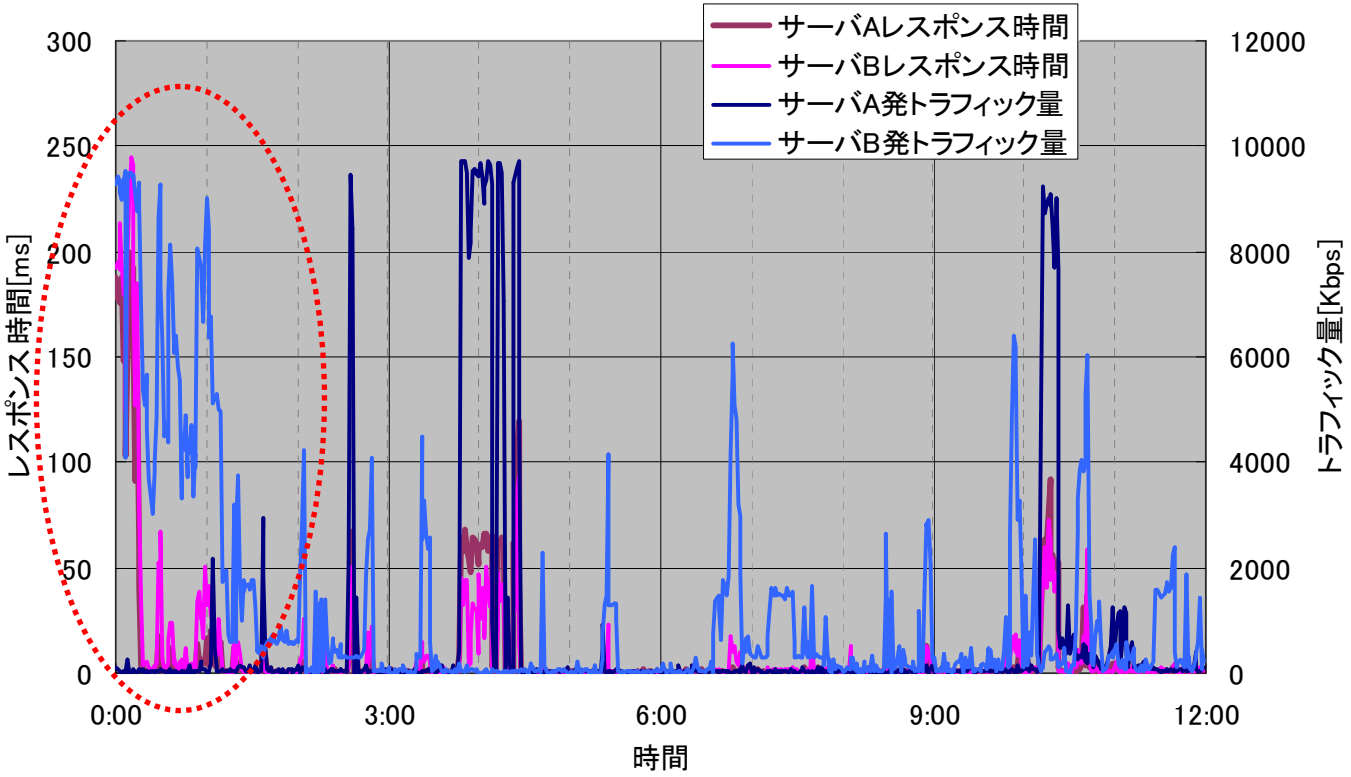
トラフィック量と付き合わせる

- 発トラフィック量とつき合わせてみると．．．．
  ⇒発トラフィック量10Mbps近くになるとレスポンス時間劣化していた．
- 一方，発トラフィック量がほとんどないのに劣化しているところもある．



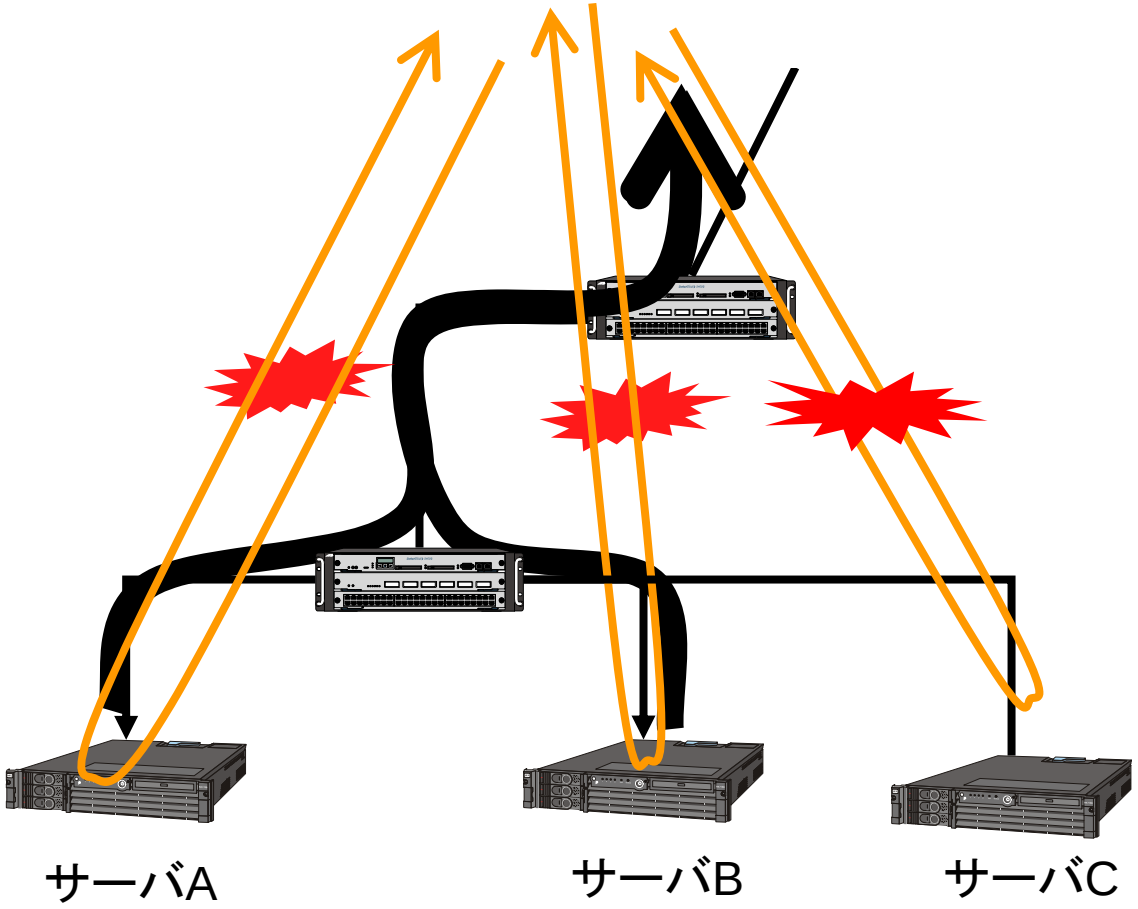
近隣サーバトラフィック量とも付き合わせる

- 当該時間は近隣サーバの発トラフィック量が増加していた.
  - 巻き添えを食らっている.
- 近隣サーバのレスポンス時間も同期して劣化



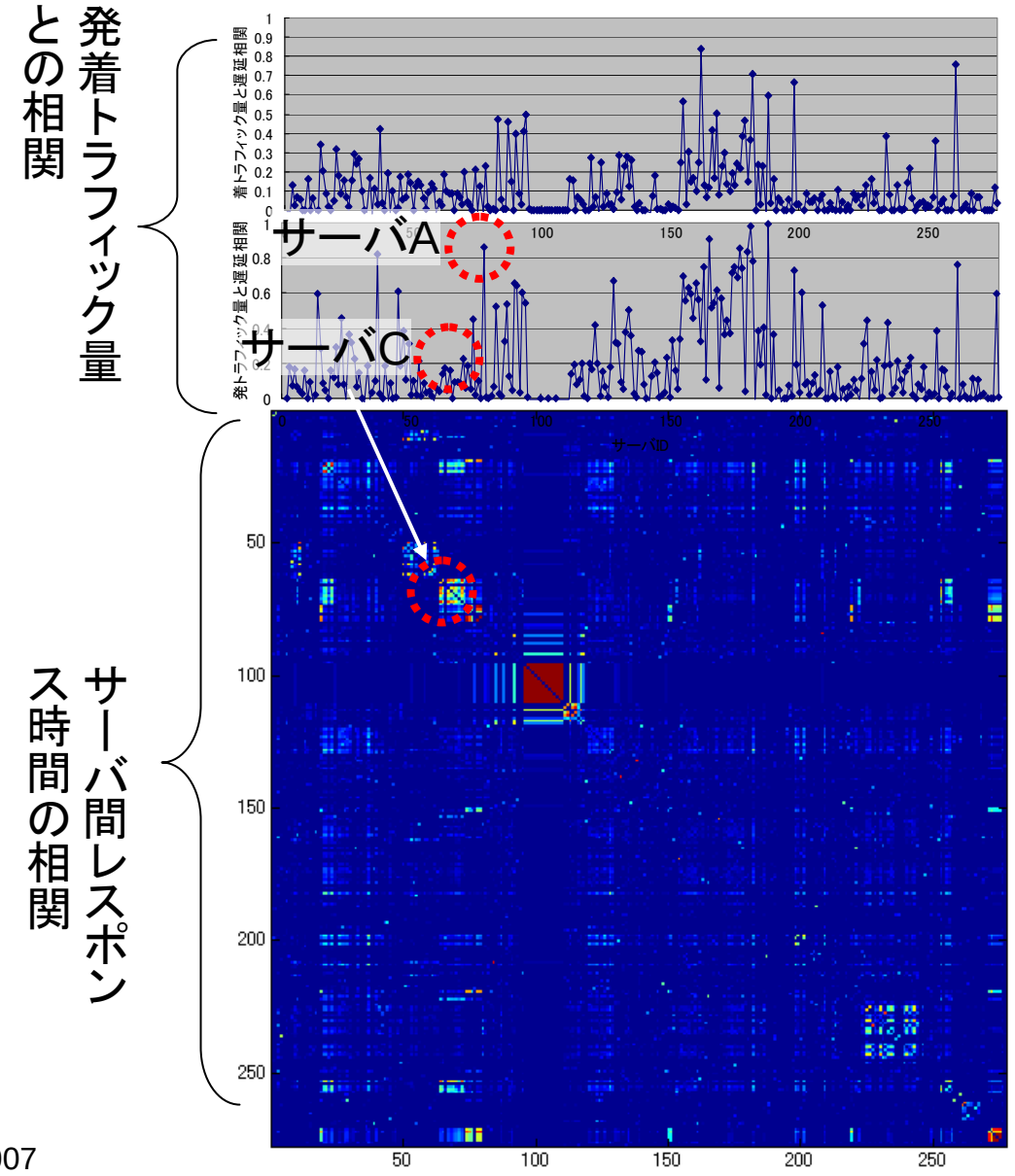
ようするに. . . .

- お互い様
- でも一方的に巻き添え食らっている人もいる.



トラフィック／品質，近隣品質との相関分析

- ちまちま分析するのは大変.
  - もうすこしシステマティックに分析できないか.
- レスポンス時間が劣化したサーバそれぞれに対して下記を分析
- 自身のトラフィック量との相関
  - 自身とのトラフィック量との相関が高ければ自分のせい
- 他サーバのレスポンス時間との相関
  - トラフィック量と相関がないのに他のサーバとのレスポンス時間との相関が高ければそっちのせい

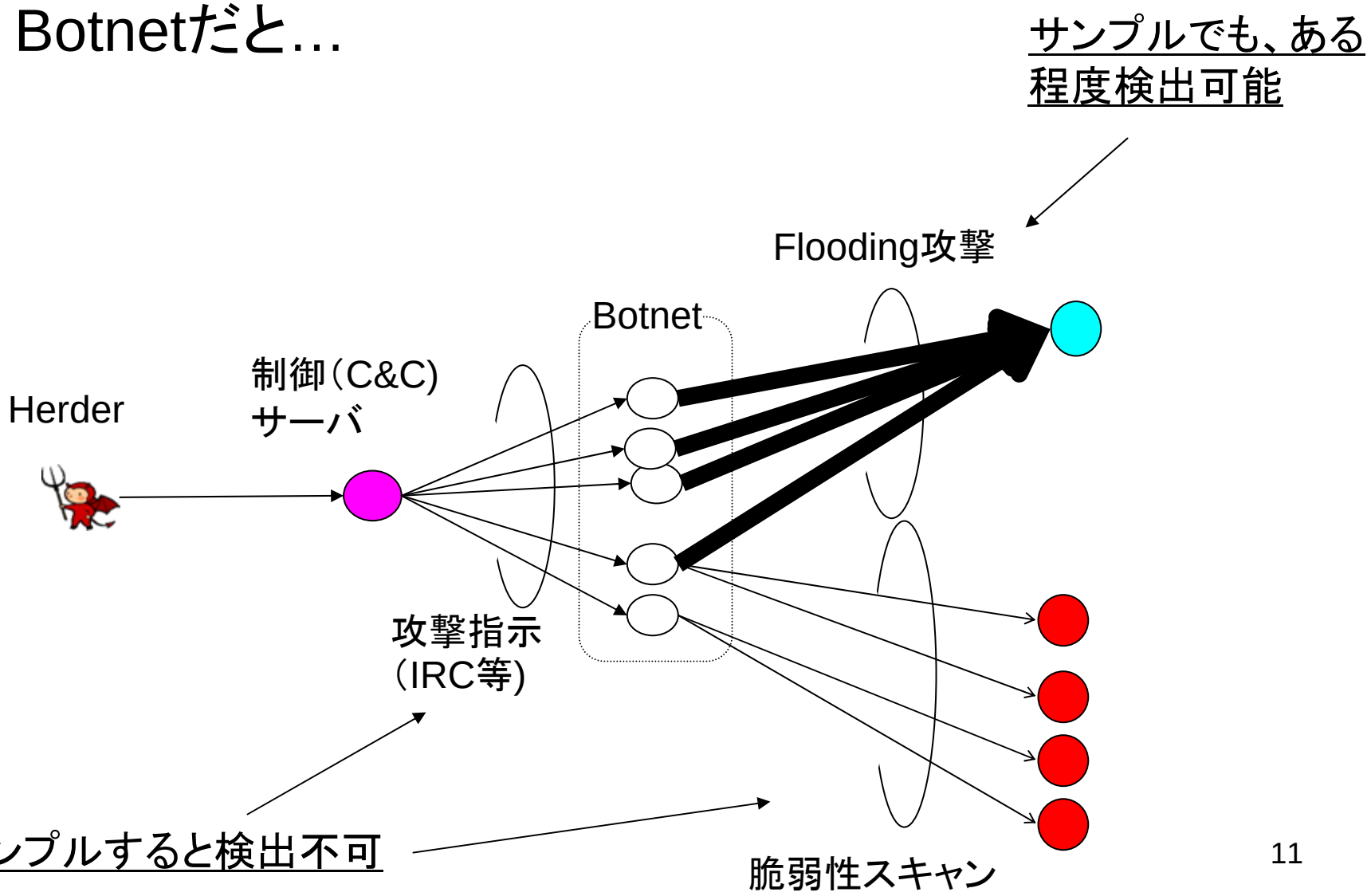


ホスト間通信関係の計測

Botnetの検出

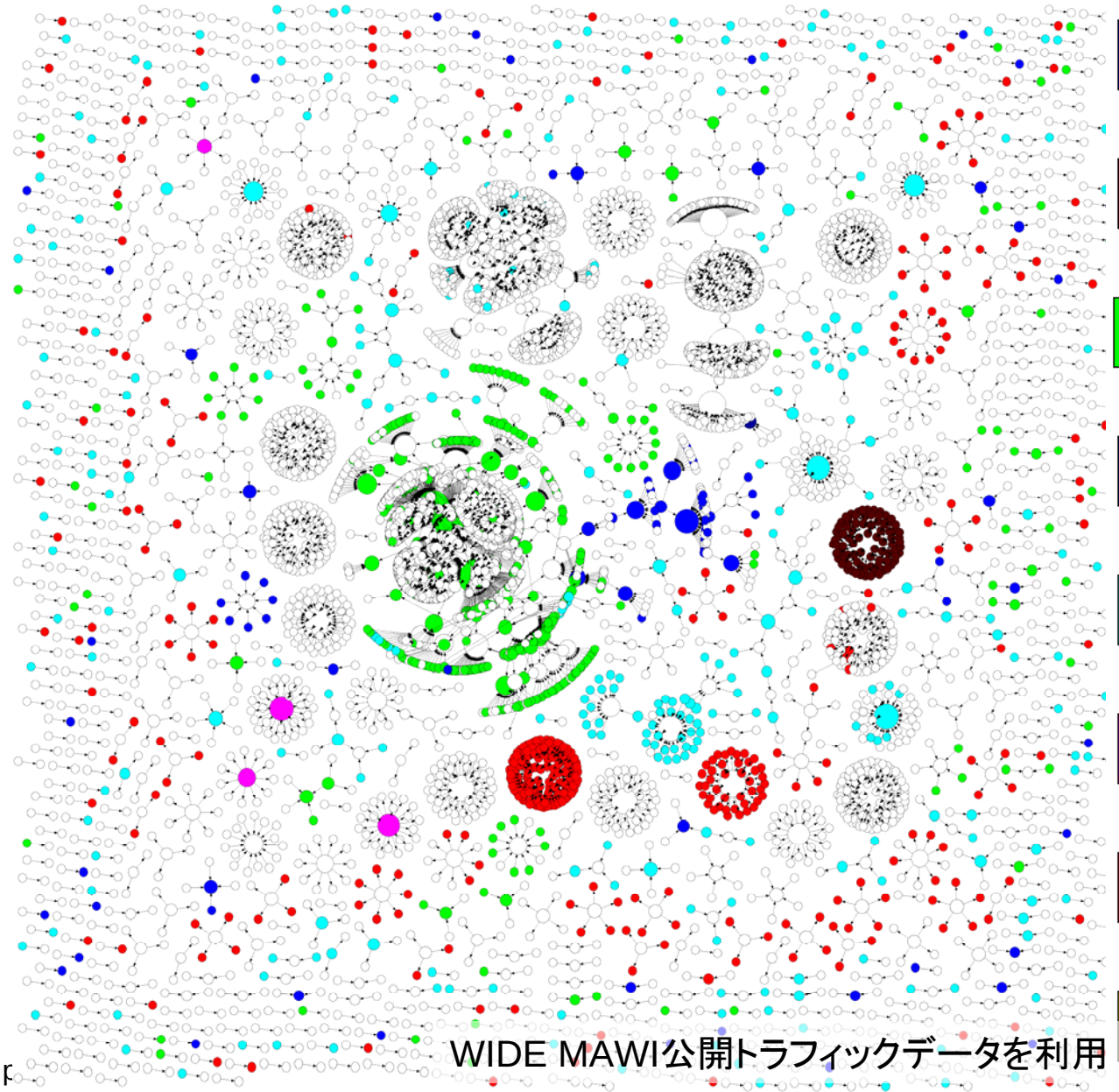
ホスト間通信関係

- Botnetだと…



ホスト間通信関係その1:サンプル無し

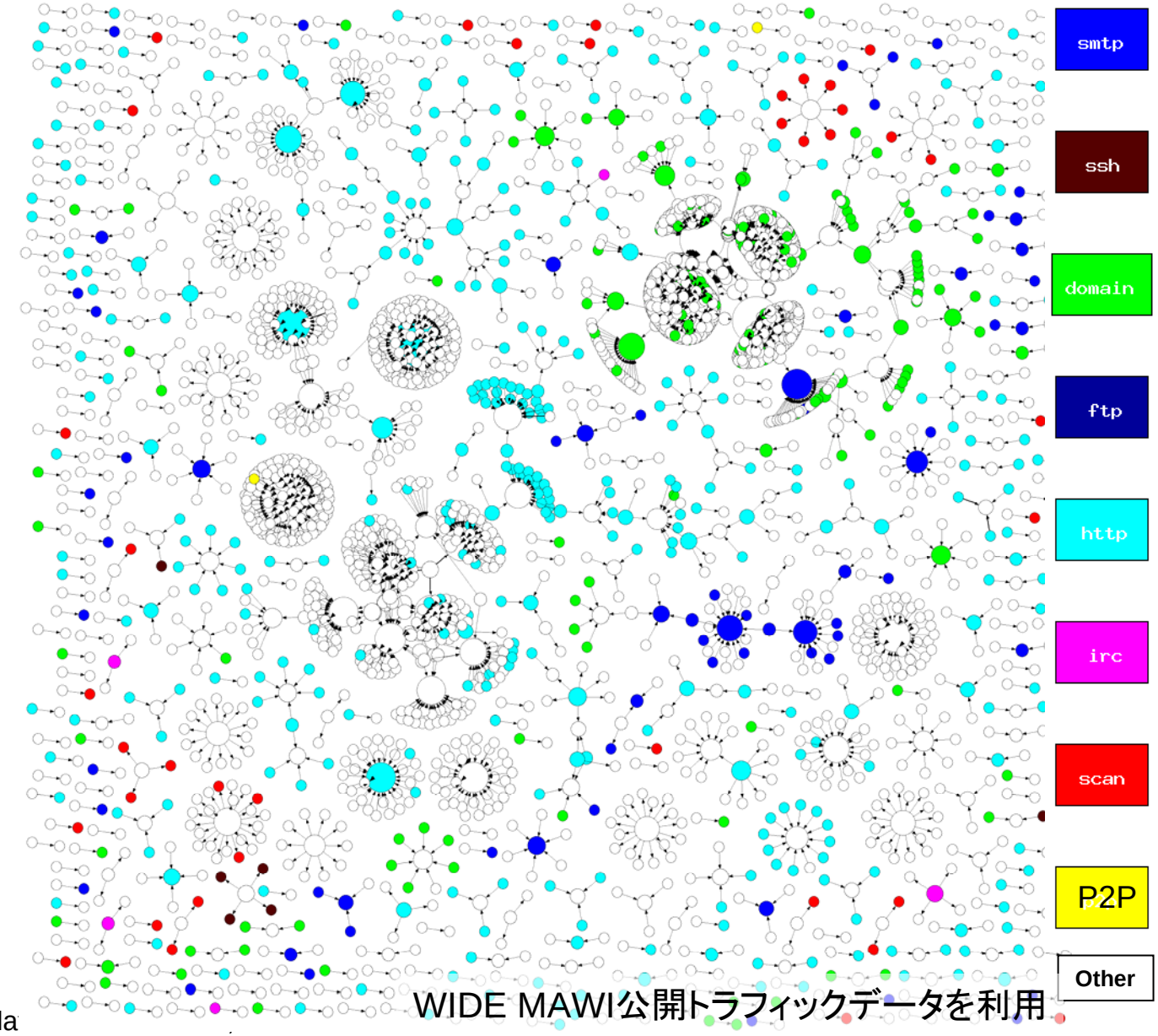
- graphvizを使って通信関係構造を可視化
- 見えるもの
  - –DNSのクラスタ
  - –ポート番号135/445のスキャンクラスタ多数
  - –ssh:brute force attackのためのスキャン？



WIDE MAWI公開トラフィックデータを利用

ホスト間通信関係その1:サンプル有り

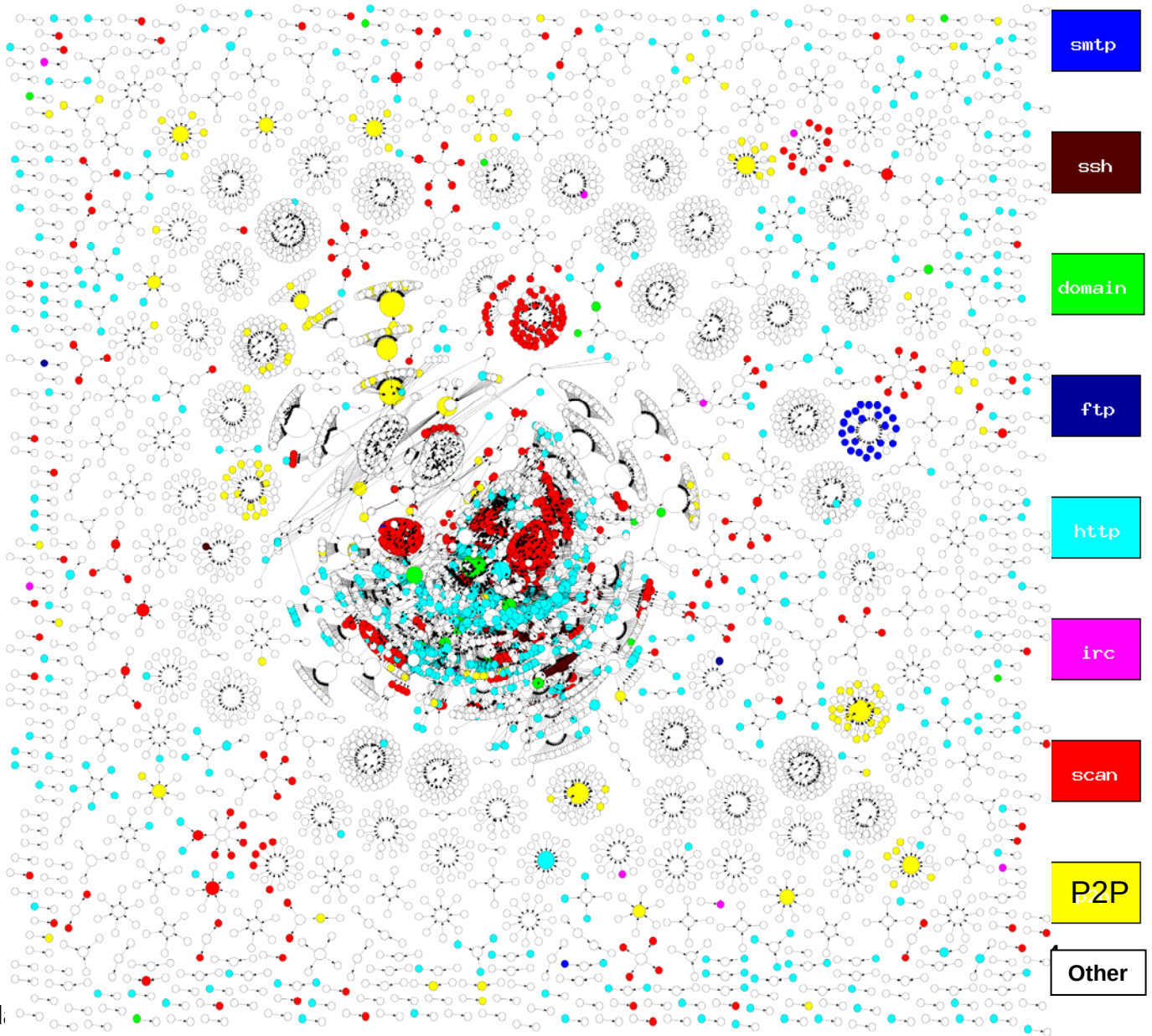
- 1/1000パケットサンプルデータで通信関係をみてみると. . .
- スキャン系はほぼ見えなくなる。
  - –見えても単発としてしか現れない。



WIDE MAWI公開トラフィックデータを利用

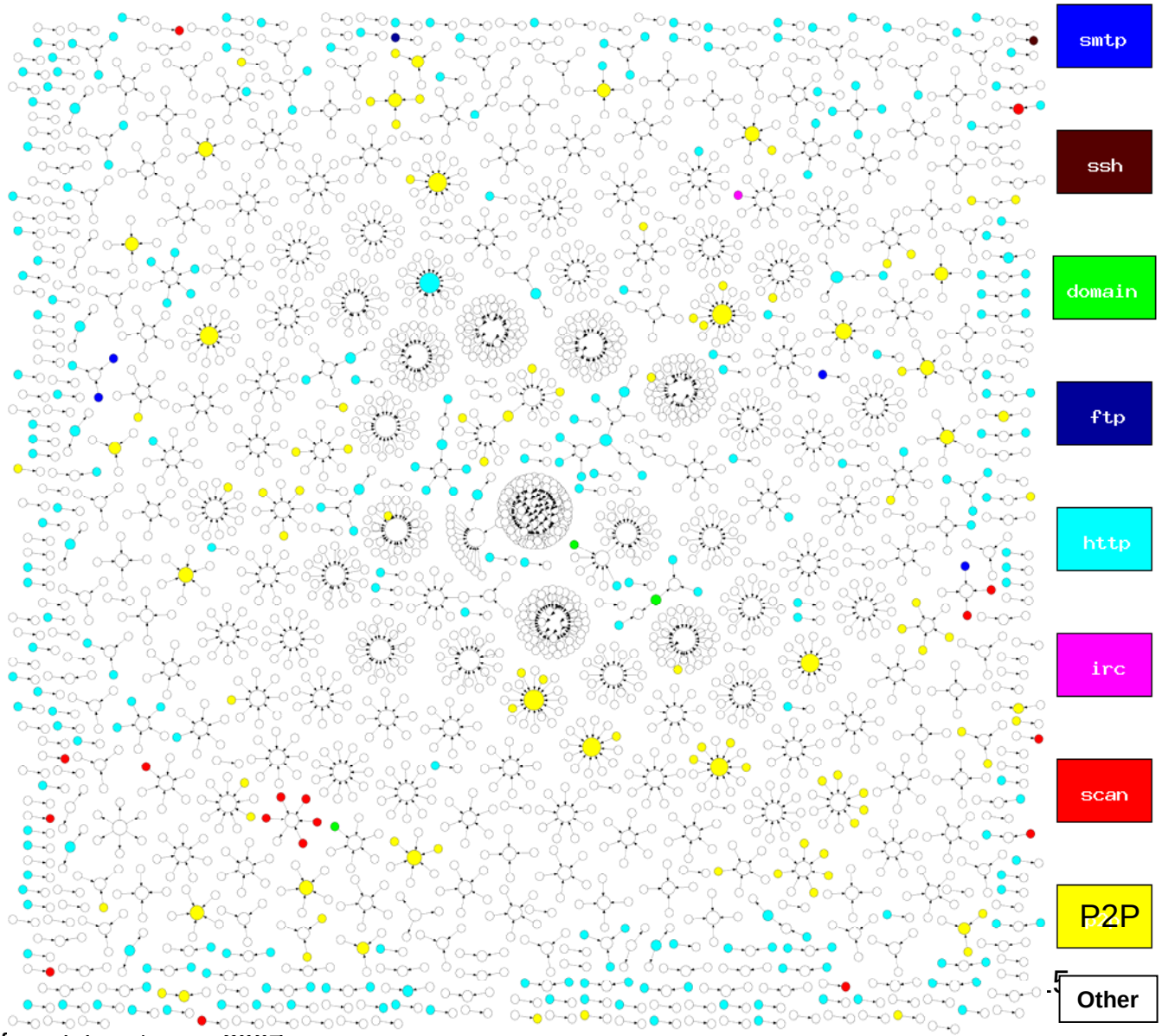
ホスト間通信関係その2：サンプル無し

- 別のデータ
- スキャンのクラスタが多数
- Webのクラスタもいくつか

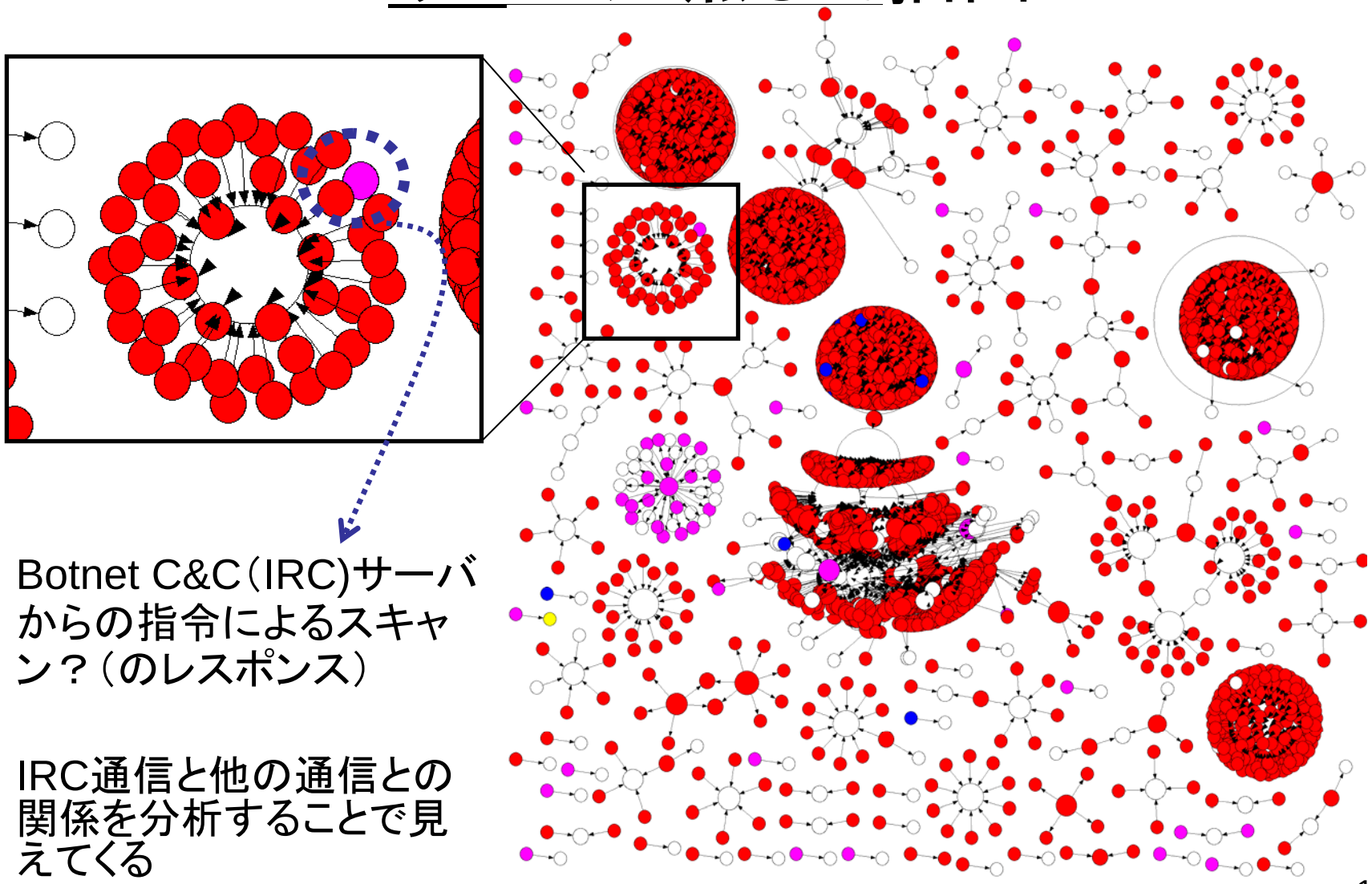


ホスト間通信関係その2:サンプル有り

- 1/1000パケットサンプルデータで通信関係をみてみると. . .
- クラスタ構造が見えなくなる
  - リンクが切れるため
- スキャン, IRC, メール(SMTP)も見えなくなる
- 見えるのは、パケット数が多いP2P, Webばかり



IRC, scan系やっているホストの通信だけ サンプル無しで抽出



- Botnet C&C（IRC)サーバからの指令によるスキャン？（のレスポンス）
- IRC通信と他の通信との関係を分析することで見えてくる

まとめ

- 今日
  - サンプルなし計測の応用例として
    - 品質計測
    - 通信関係構造分析
  - の実例を紹介
  - 攻撃の影響分析、原因分析に適用可能か？

- 今後
  - 通信関係構造の変化検出による攻撃（予兆）検出，ネットワーク利用形態の変化検出ができるとうれしいかも.

おまけ1：M/M/1

- サーバAトラフィック量とレスポンス時間の関係をM/M/1モデルで近似してみた.
- （レスポンス時間⇔待ち客数の変換は必要だが），傾向はほぼ一致.
  - リンク帯域ボトルネックの時のキューイング遅延は近似できる.
  - サーバCPU処理速度ボトルネックとかだと，多分こんなに合わない.

  ⇒きれいに合う/合わないでどっちがボトルネックかとかいえないか…