

インターネットトラフィックの統計解析 ～傾向と対策～

森 達哉† 川原 亮一‡ 内藤 昭三†

† NTT 情報流通プラットフォーム研究所

‡ NTT サービスインテグレーション基盤研究所

2002年 3月26日

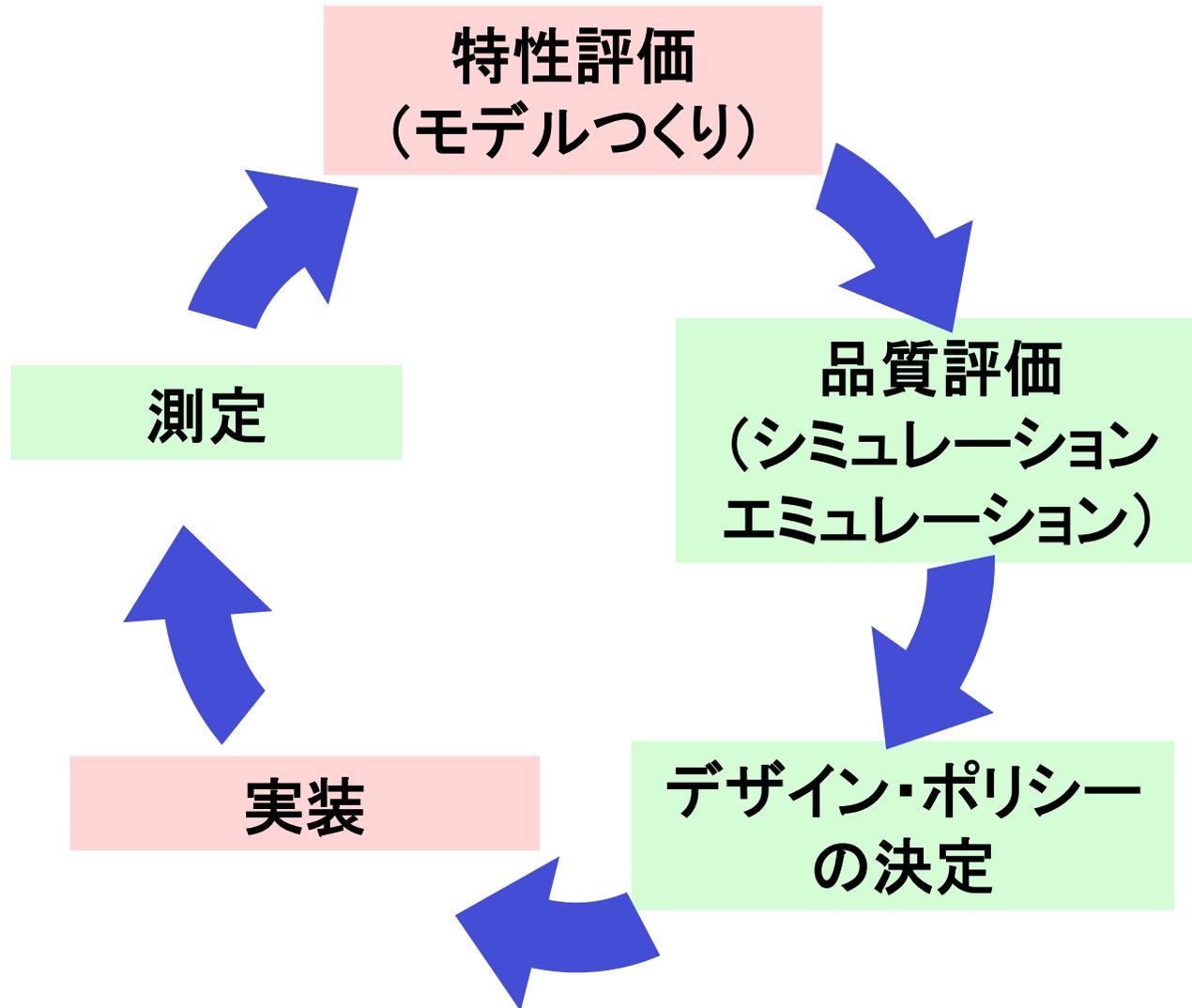
テレコムサービス協会

技術研究会最新動向セミナー

トラフィック流量変動の特性評価

- より高品位なサービス実現にむけて
 - 品質評価
 - 制御
 - 予測(設計)

特性評価の位置付け

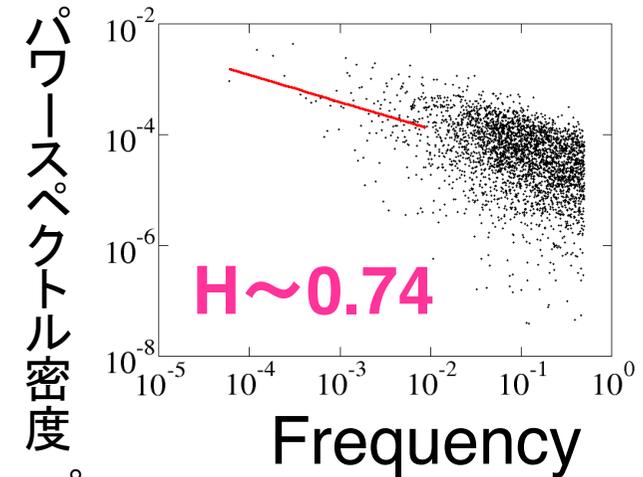
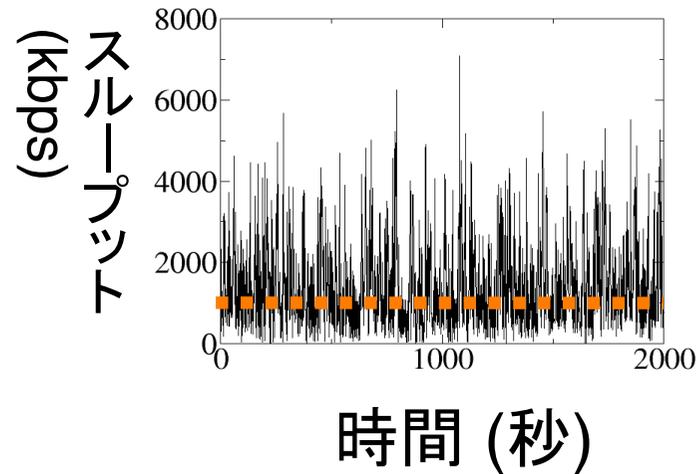


トラフィックの流量変動特性

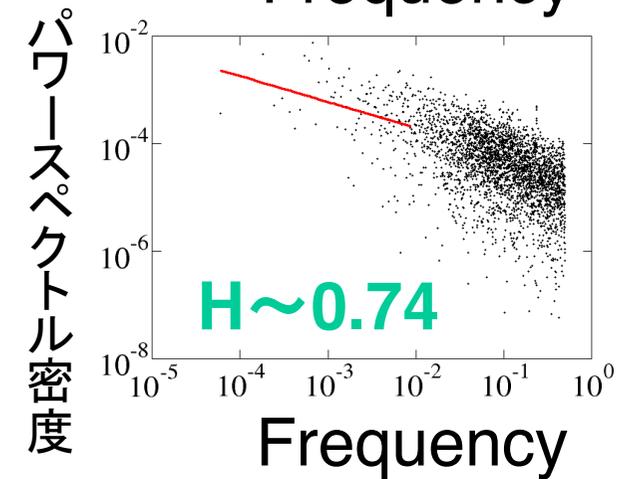
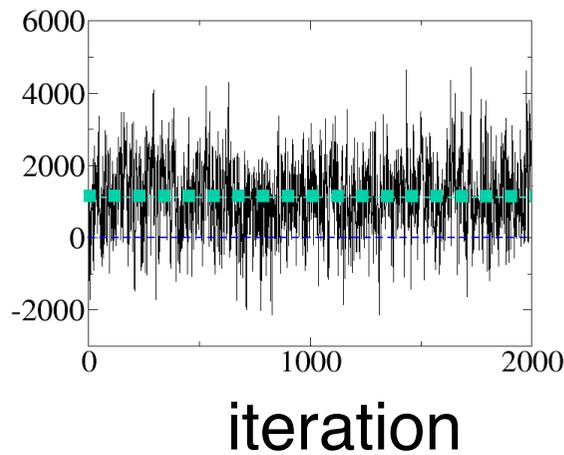
- 時間的構造 (時間軸)
 - 時間相関 (パケット到着間隔)
- 空間的構造 (変動軸)
 - 平均値からのずれ

トラフィックの流量変動

実トラフィック



擬似トラフィック (fGN)



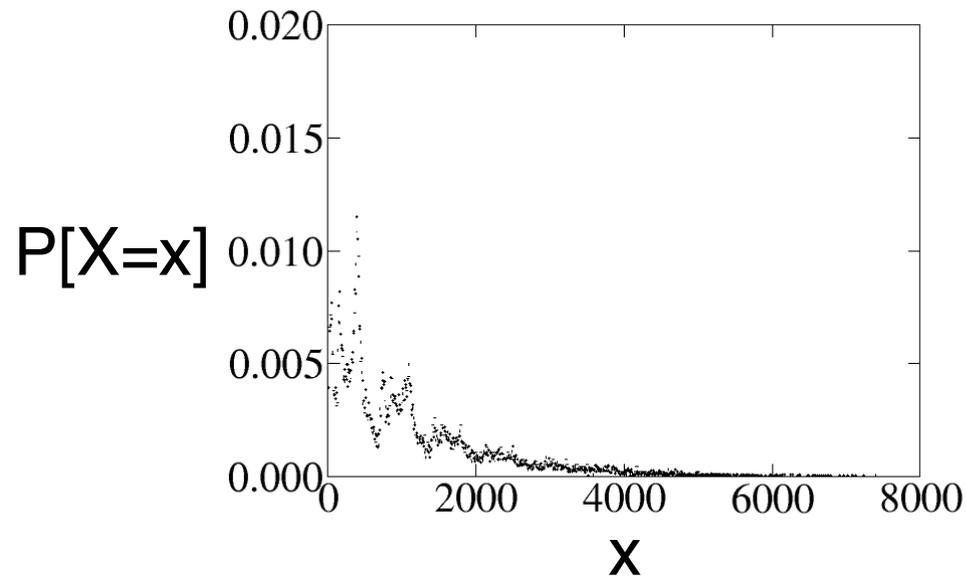
流量変動の時間相関

- 無相関モデル (memory-less)
 - パケットはでたらめに到着
 - トイレに到着する客の時間間隔
 - ポアソンモデル
 - 電話網を中心としたトラヒック理論の基礎
- 長期自己相関モデル (long-memory)
 - パケット到着間隔は履歴をひきずる
 - 比較的混んでいる道における車の車間距離
 - 自己相似モデル
 - 1993年ベルコアの研究者による報告以来脚光を浴びている

流量変動の空間構造

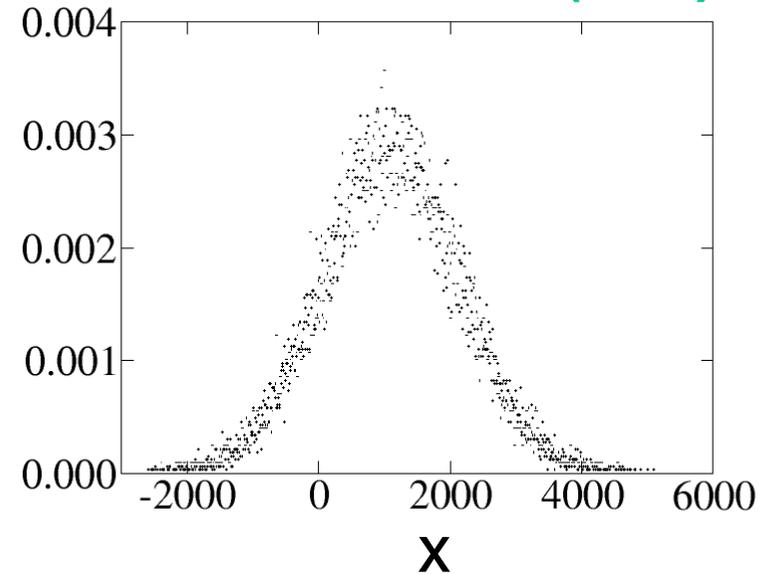
- 平均値周りの変動・・・**周辺分布**

実トラフィック



非正規分布

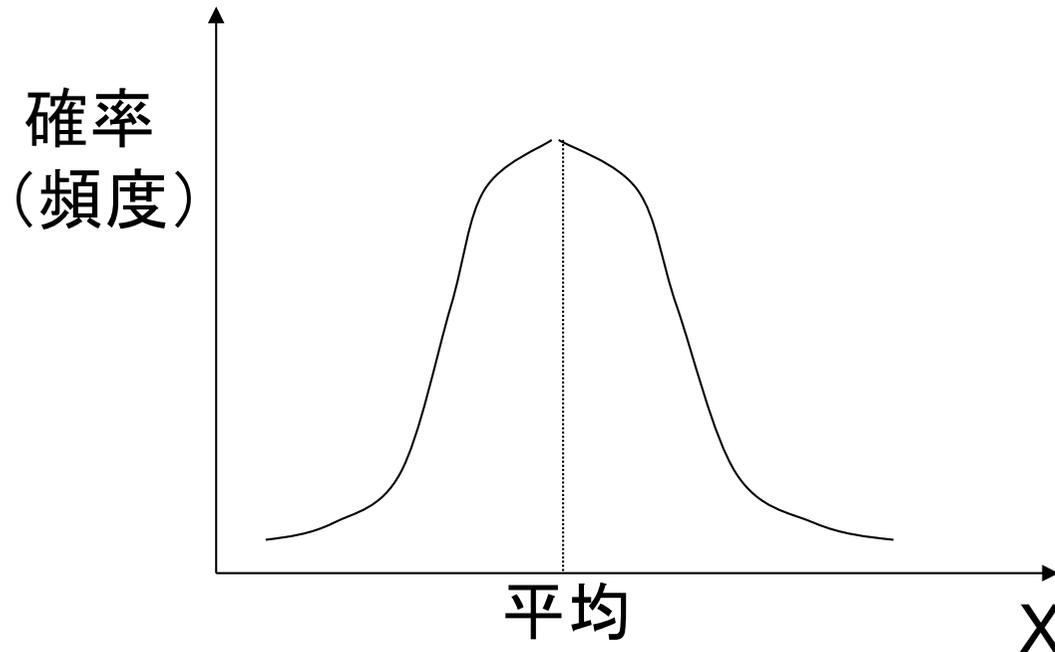
擬似トラフィック(fGn)



正規分布

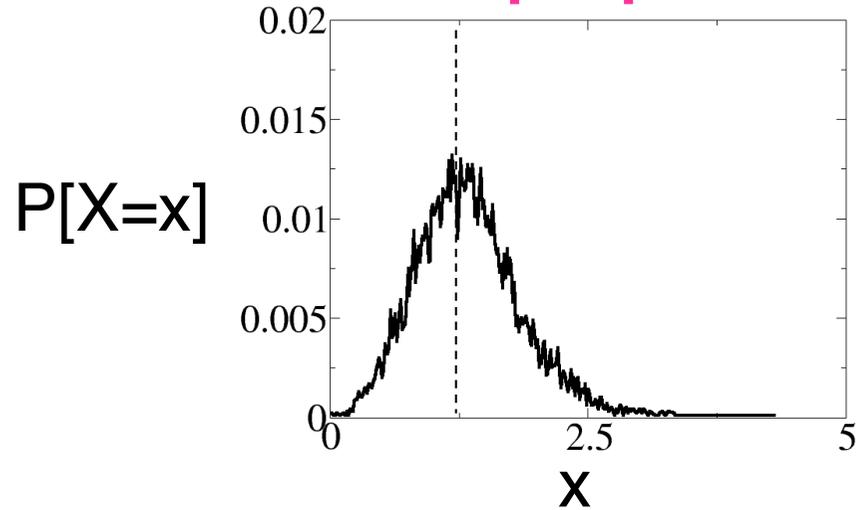
正規分布 (normal distribution)

- 全ての分布の源
 - 中心極限定理により保証
 - でたらめな変数の重ね合わせ → 正規分布に収束

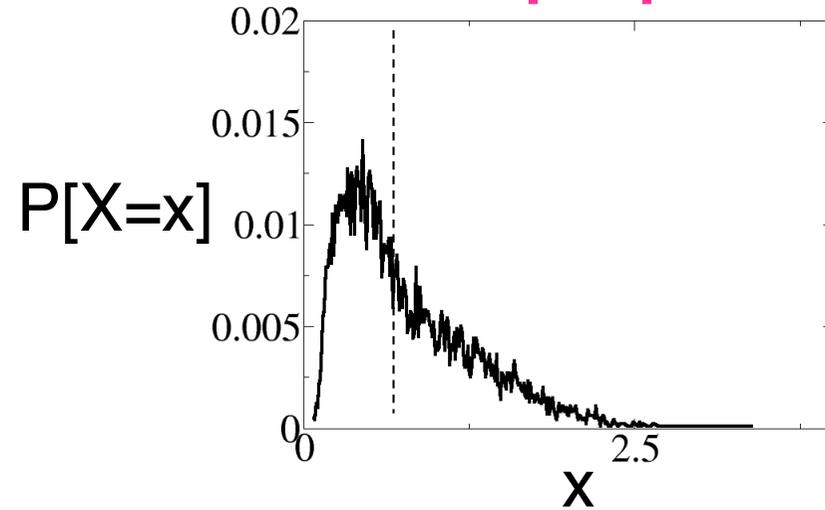


トラフィック変動の示す非正規分布の例:

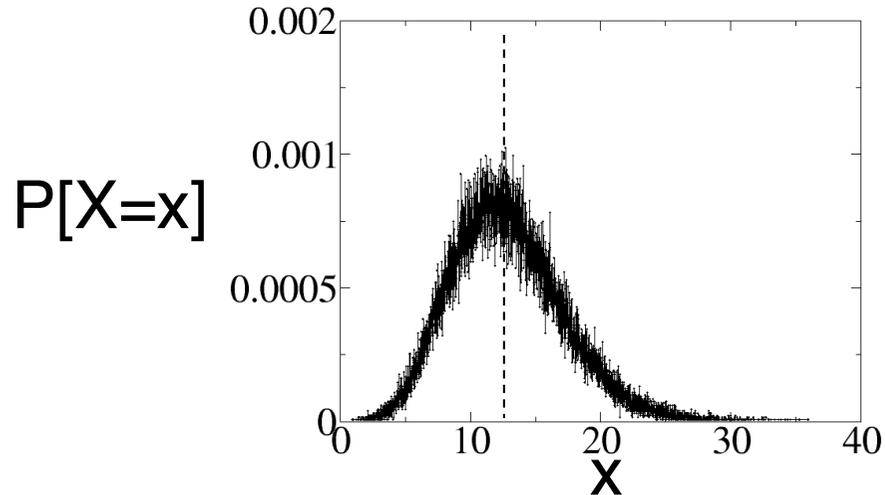
WIDE sample point A-a



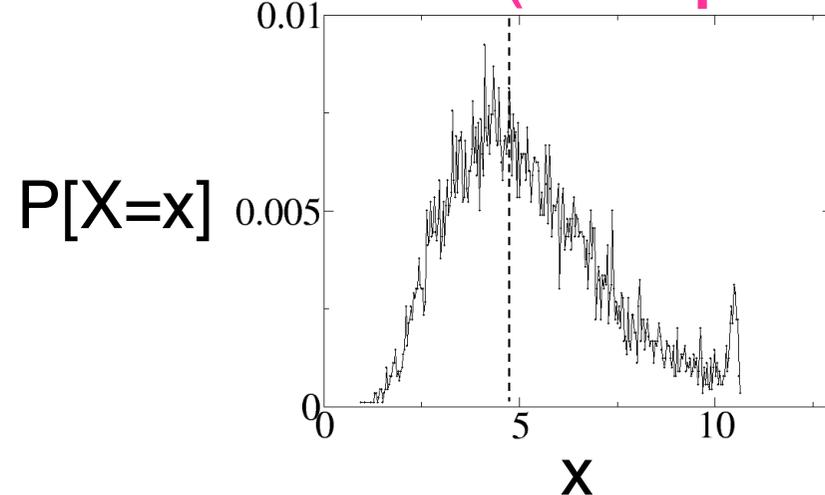
WIDE sample point A-b



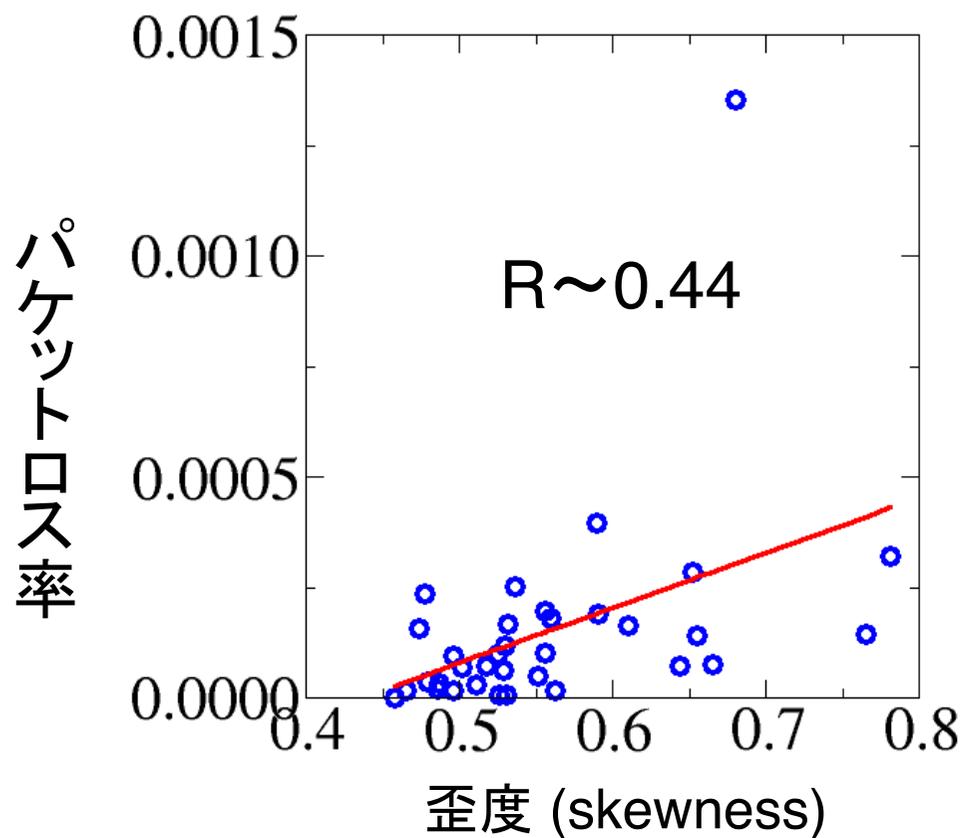
OCN-SINET (OC-3)



ECL-OCN (12Mbps ATM)



非正規性とネットワーク品質



小 ← 非正規性 → 大

[Tatsuya Mori 2001]

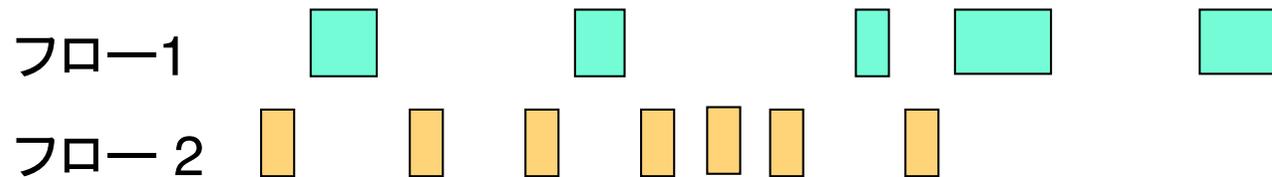
流量変動が示す非正規性の原因

- Normalなフローの重ねあわせではない！
 - インターネットはヘテロ
 - 中心極限定理が成立していない
 - 重ねあわせがさらに増えれば成り立つ？
 - » 現実のトラフィックではそうはなっていない
- 個々のフローの振る舞いに着目する

IPフローの定義

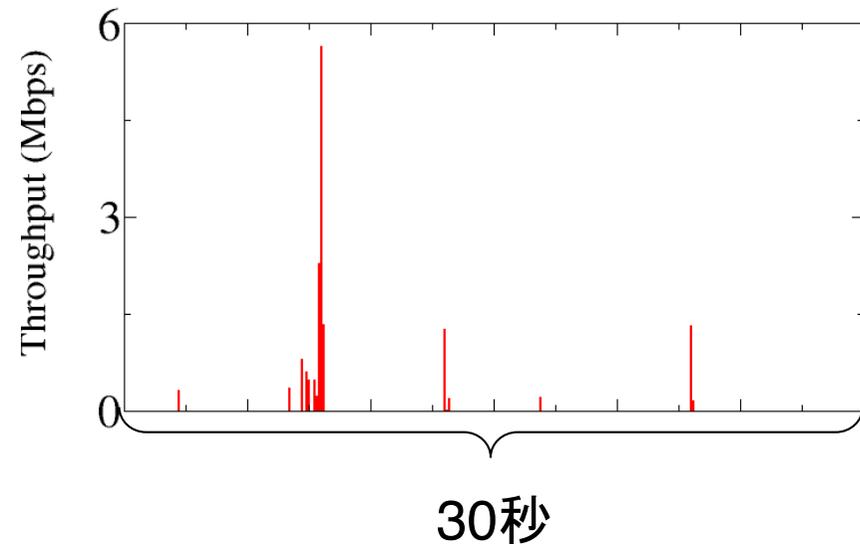
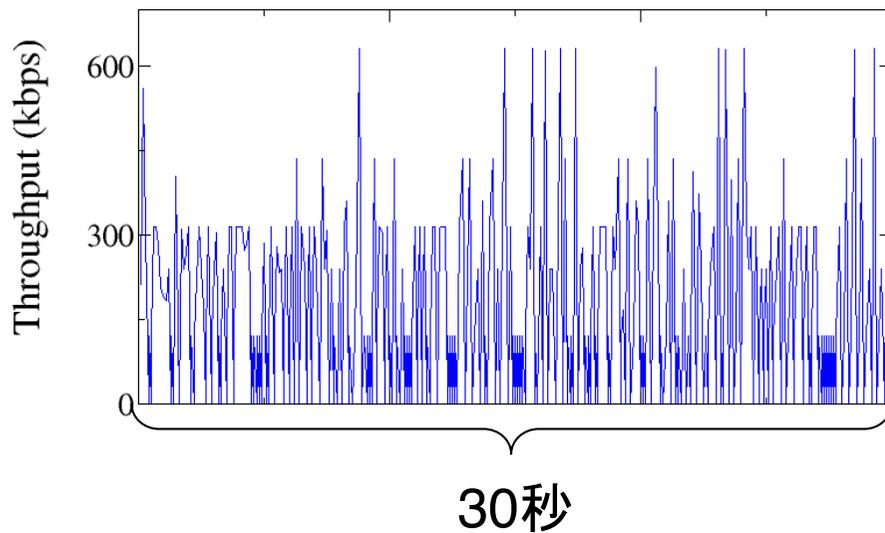
- 以下の5つのキーの同一な組み合わせを持つ
パケット群

<送信IPアドレス> <受信IPアドレス>
<送信ポート番号> <受信ポート番号>
<プロトコル>



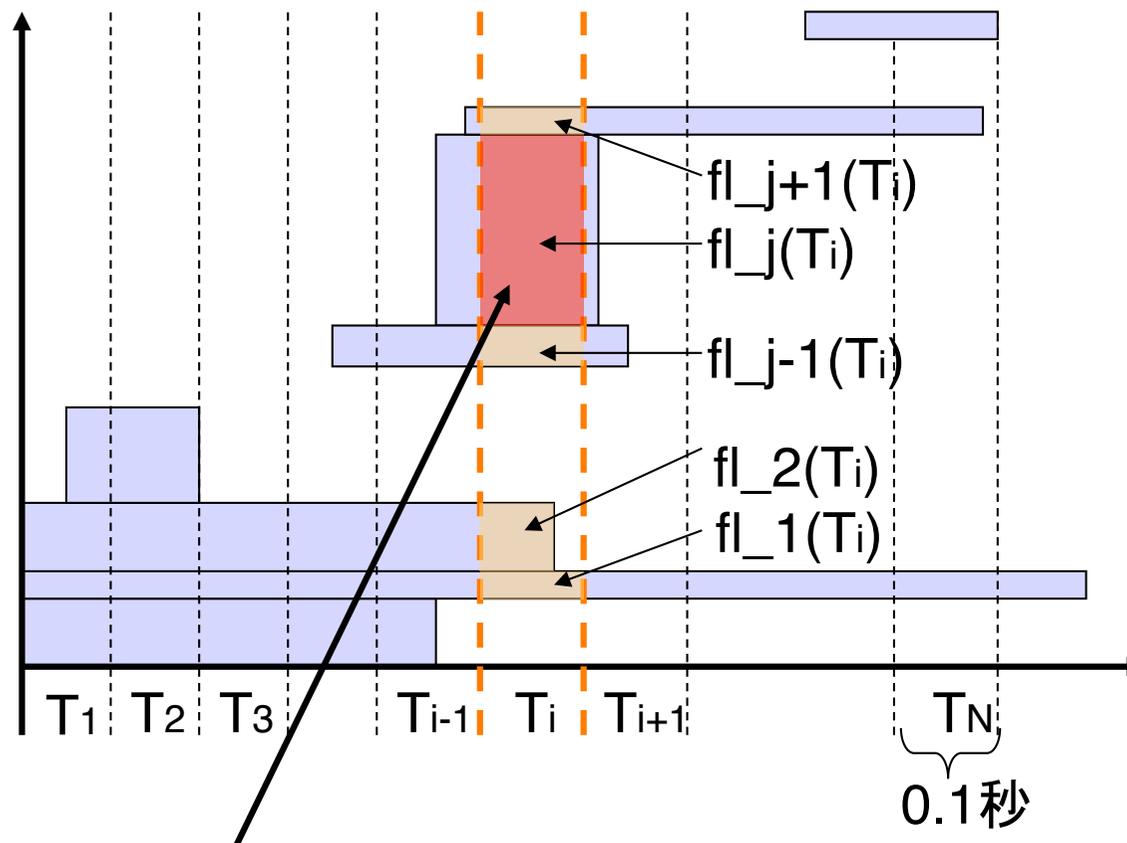
個々のIPフローの特性

- 激しく揺らいでいる
 - 平均レートでは不十分なケースがある
- 間欠的なバースト性
 - 短時間内での振る舞いが重要



Per-time-blockフローの定義

- IPフローの特性を捕らえるために

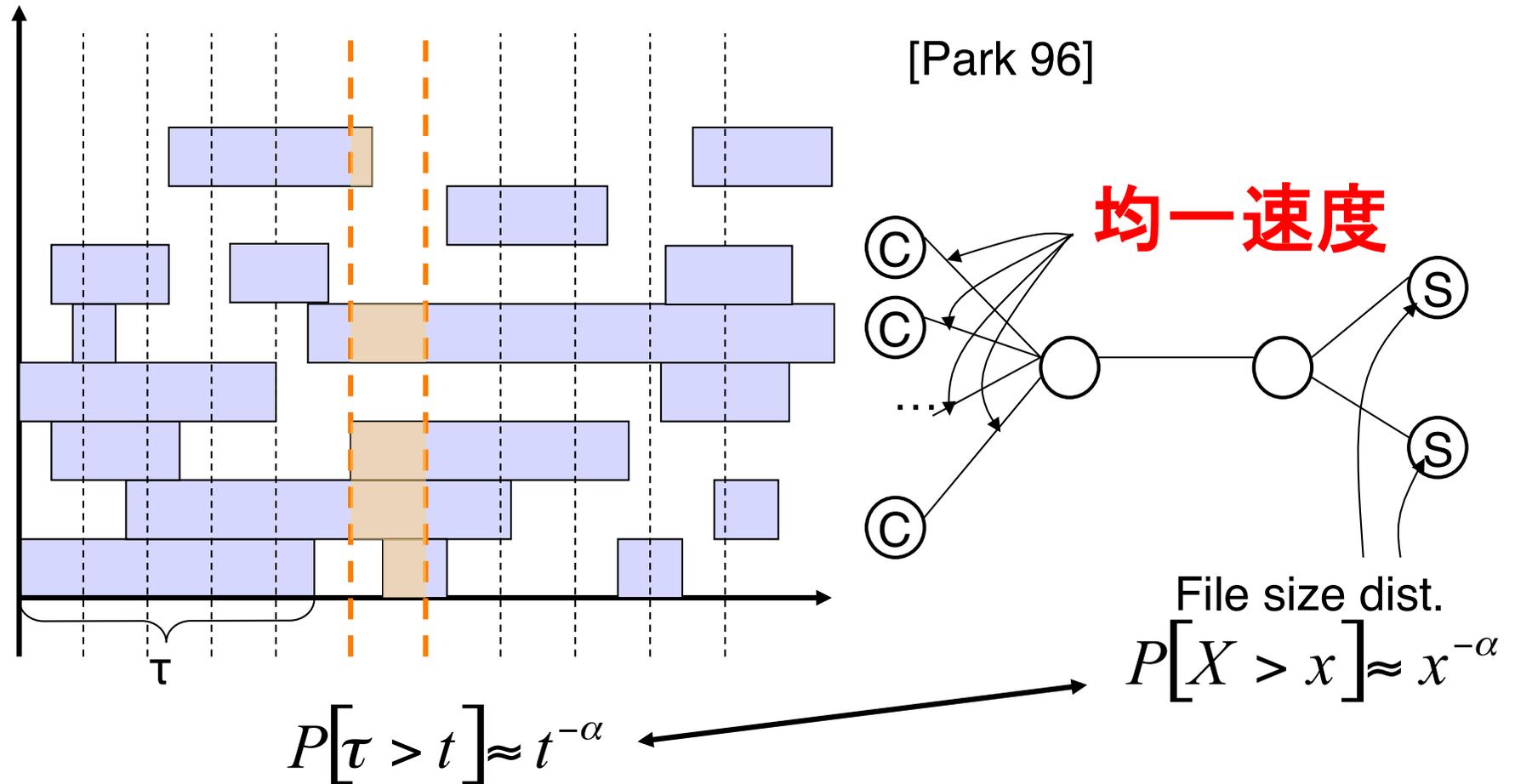


Greedy フロー (20パケット/0.1秒) [Tatsuya Mori 2002]

Greedy フローの特性評価

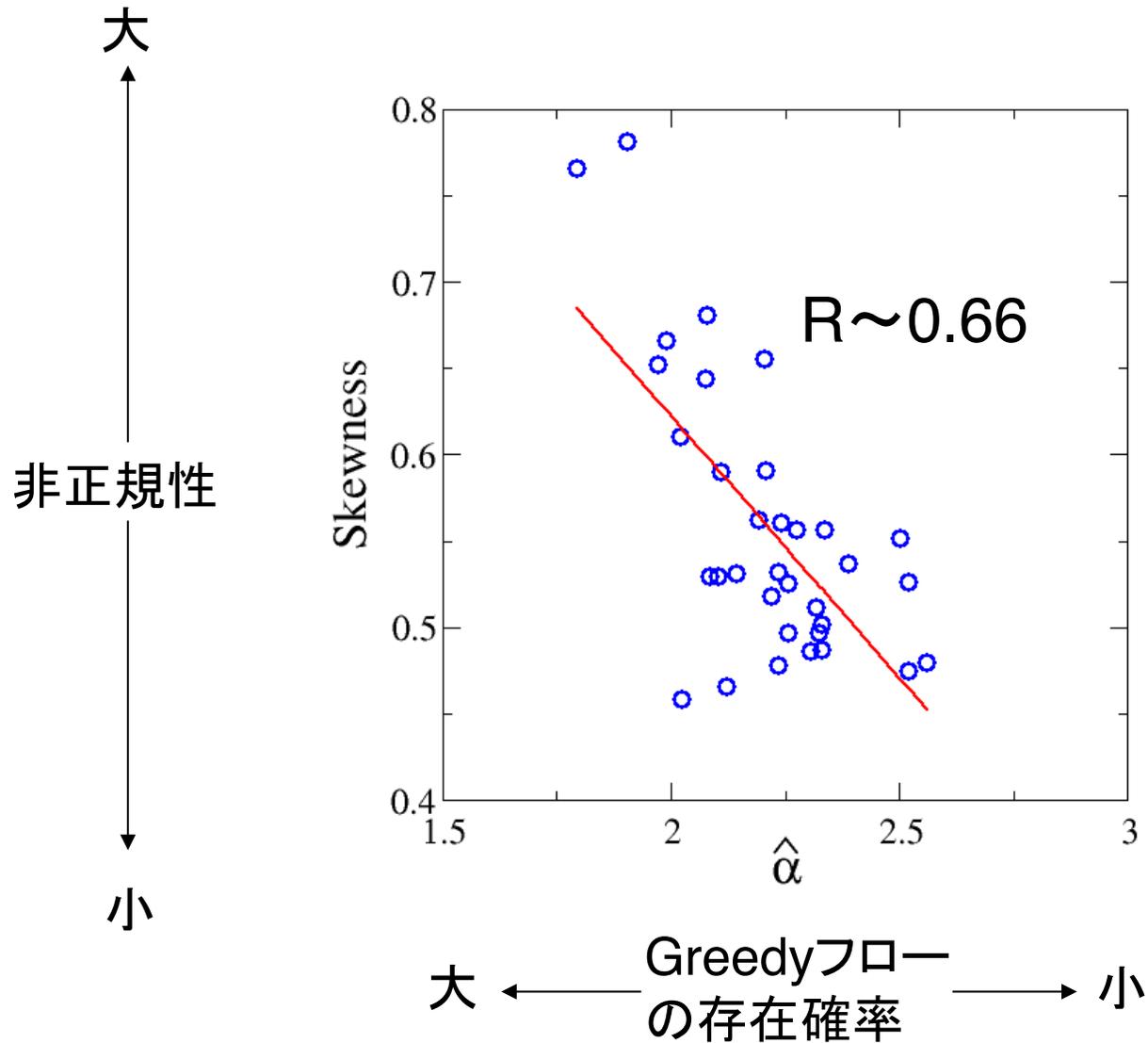
- サイズ分布
 - How much?
- ホップカウント統計
 - Where?
- アプリケーション統計
 - Who?

[cf] (Heavy-tailed) ON/OFFソースモデル

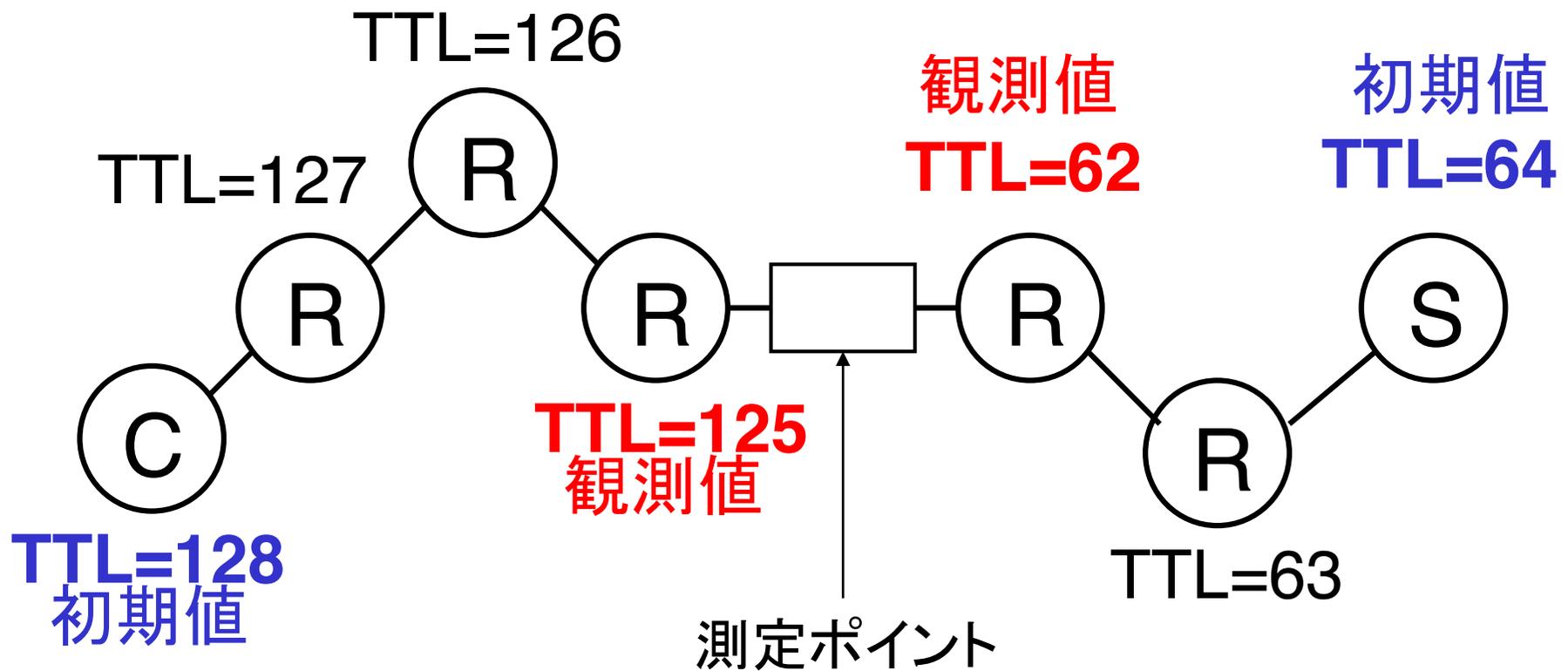


単純なトポロジーモデル → greedy フロー無し

Per-time-blockフローサイズ分布 と周辺分布の関係



ホップカウント推定手法 (1/3)



$$C,S間のホップカウント = (128 - 125) + (64 - 62) + 1 = 6$$

ホップカウント推定手法 (2/3)

- 各種システムによって初期TTL値が異なる
 - BSD, Linux系: 64 (RFC: 1340)
 - Windows系: 32, 128
 - Solaris 2.X系: 255
 - etc.
- Passive OS fingerprint 技術を用いる
 - TCP/IP オプションの実装の差を用いる
 - TTL, WS, DF, MMS, TOS, SackOK, NOP等
 - 全フローに対して約1割以上のフローのホップカウント(i.e., 発着アドレスに対する初期TTL)が特定できた

ホップカウント推定手法 (3/3)

- フロー毎ホップカウント推定の手順
 1. トレースに含まれる全てのパケットについて、発信アドレスに対応するTTL初期値を推定する
 2. 発着信のアドレスに対してTTL初期値が特定できたフローについてホップカウントを算出する

注)

1. 双方向からのパケットが観測されたフローのみホップカウント推定が可能
2. ホスト間経路は往復で同じと仮定

データ

- WIDE sample point-B
 - インターネットバックボーン
 - WIDE プロジェクト 国際接続線 下り(米→日)トラフィック
 - 100 Mbps Ethernet (18 Mbps CAR保証つき契約)
 - 2001年 9～11月 平日 14:00～ 約200万パケット

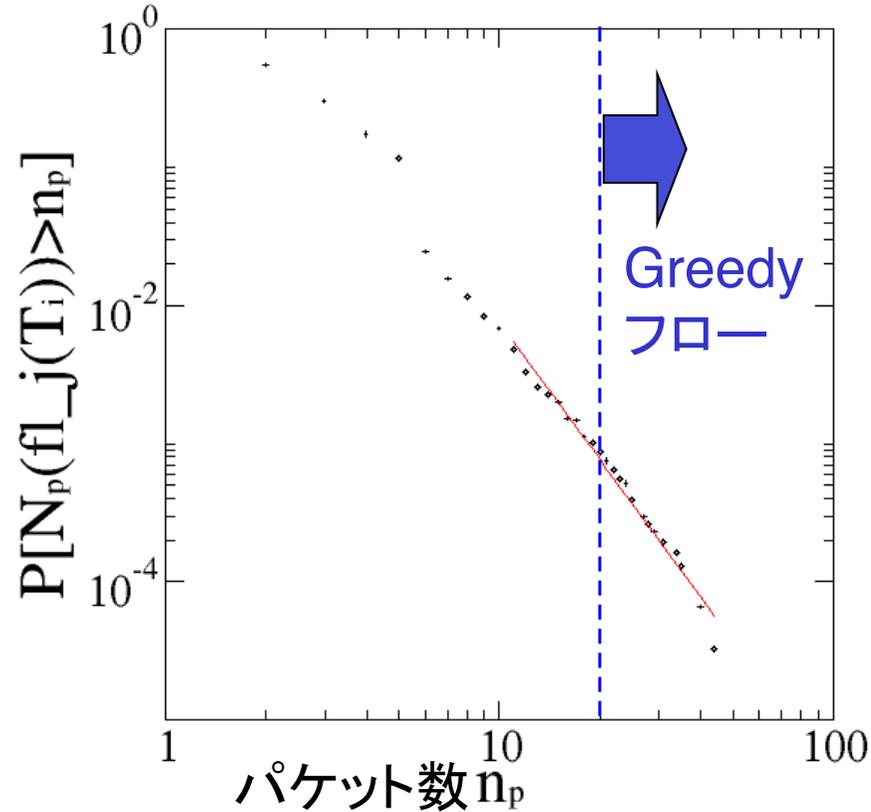
[MAWI]

- ECL
 - インターネットアップリンク
 - NTT研究開発センター 対外接続線 下りトラフィック
 - 12 Mbps ATM link
 - 2001年 7/16～19 14:00～16:00

サイズ分布

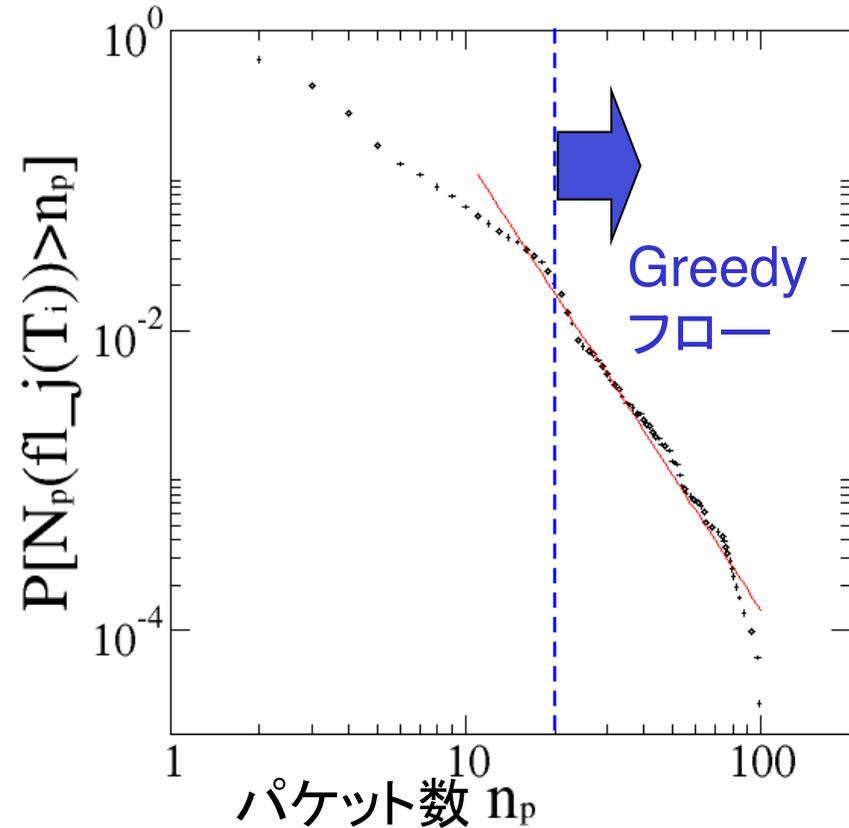
累積分布

WIDE



累積分布

ECL



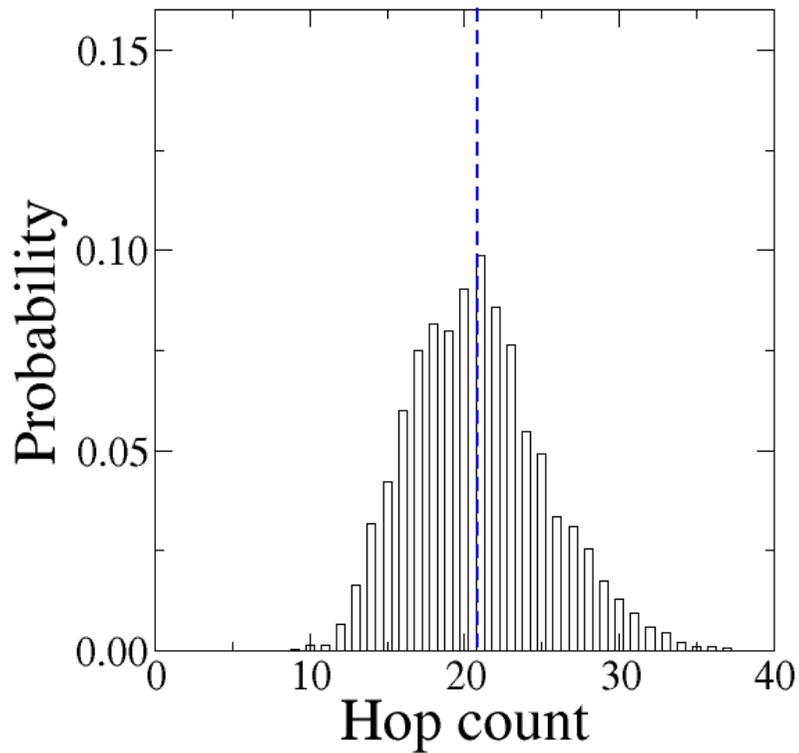
$$P[N_p(fl_j(T_i)) > n_p] \approx n_p^{-3.29}$$

$$P[N_p(fl_j(T_i)) > n_p] \approx n_p^{-3.04}$$

ホップカウント分布: (1/5)

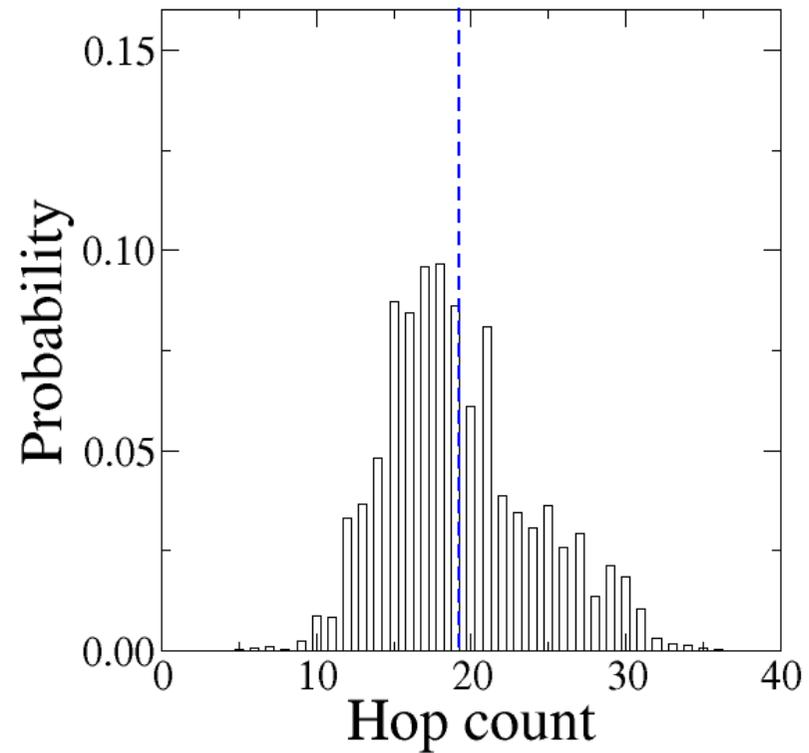
全体

WIDE



平均 = 20.89

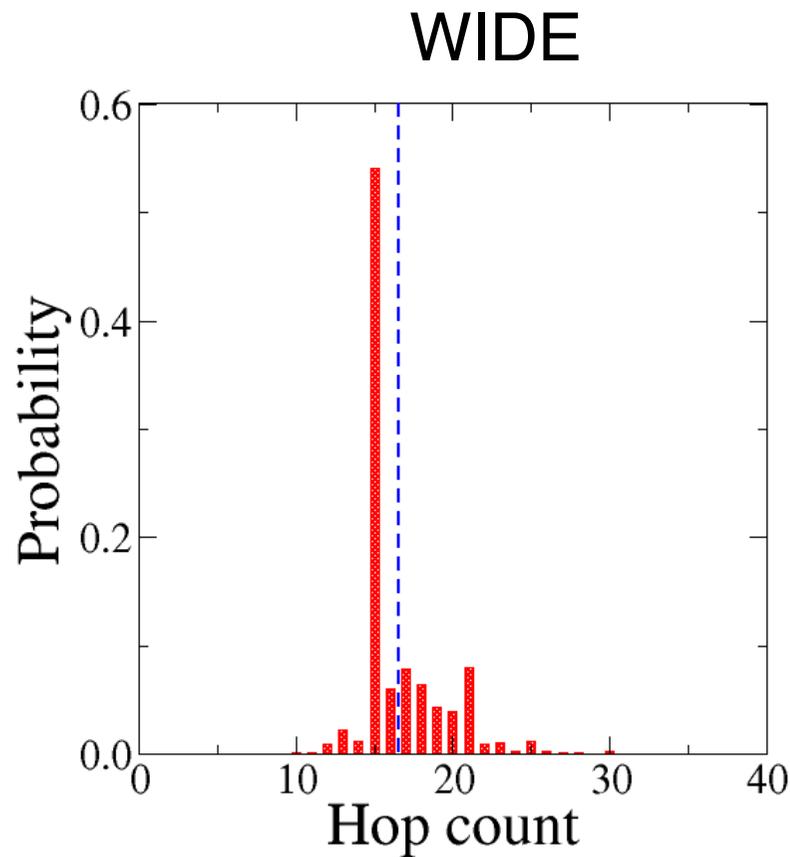
ECL



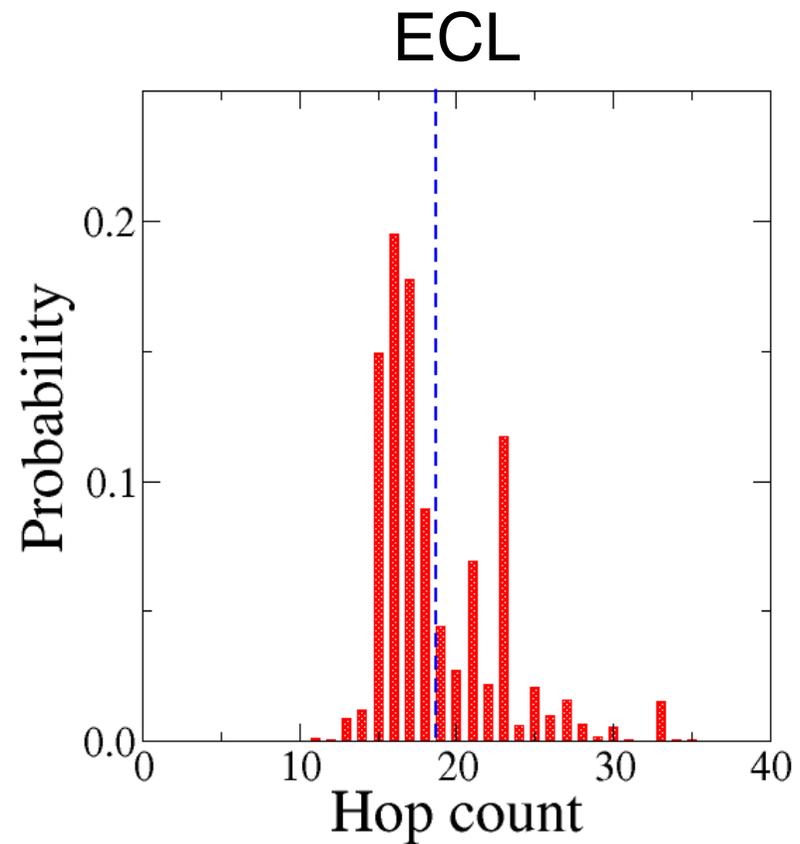
平均 = 19.25

ホップカウント分布: (2/5)

Greedy フロー

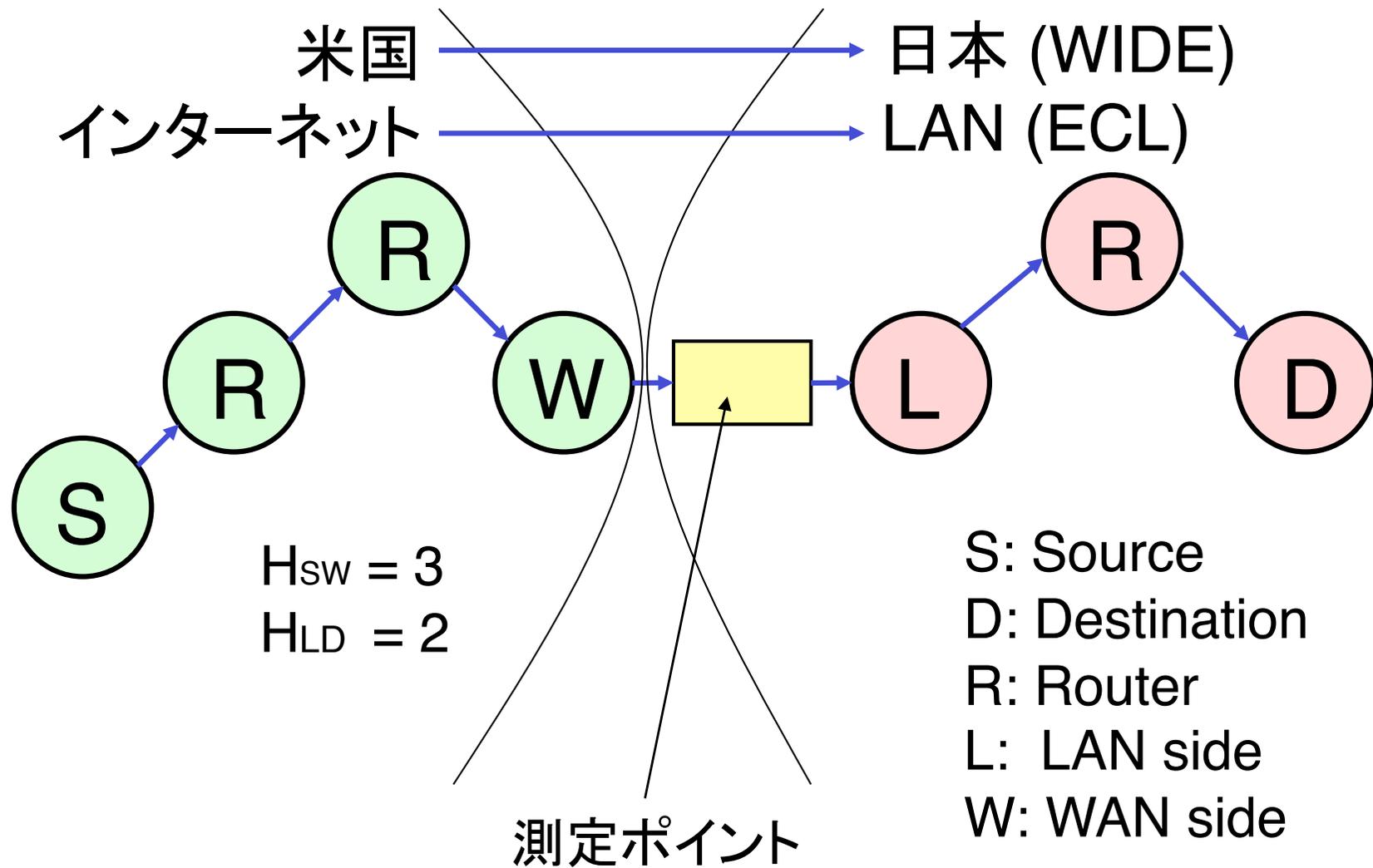


平均 = 18.69



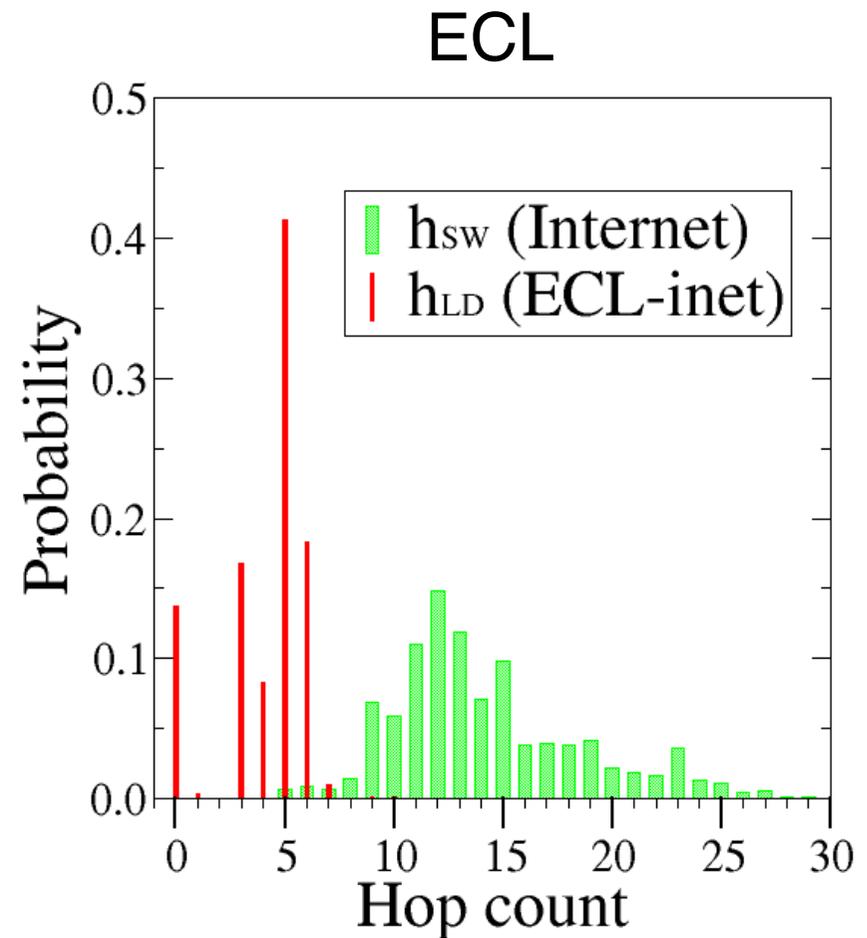
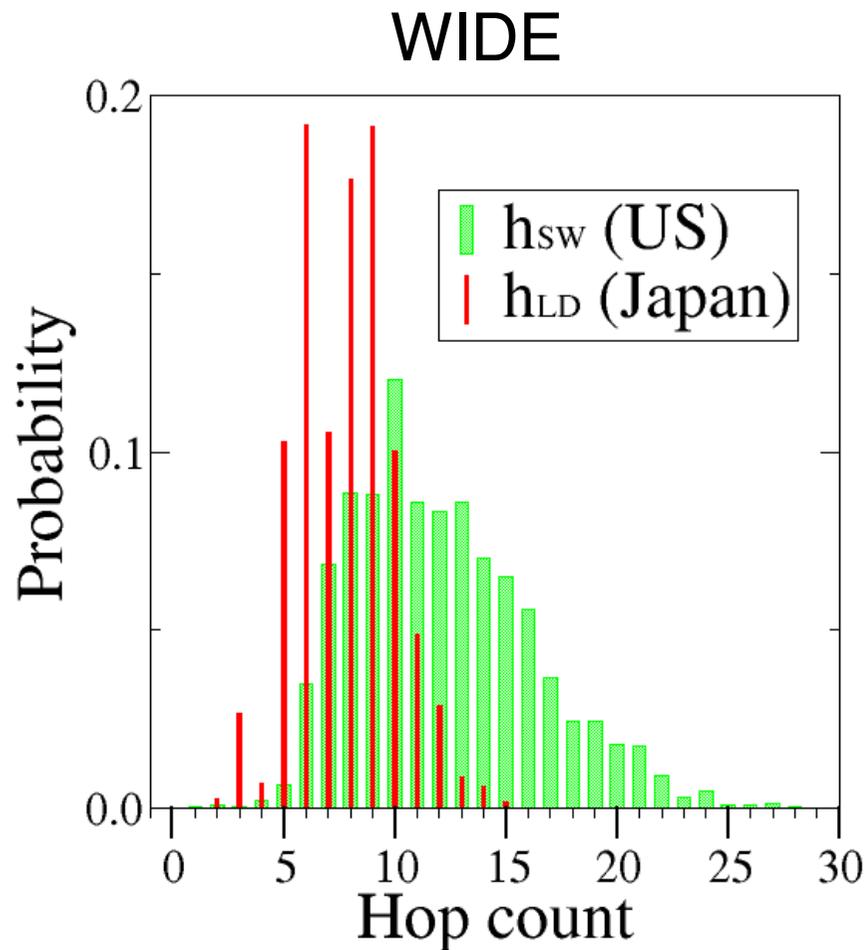
平均 = 16.58

ホップカウント分布: (3/5)



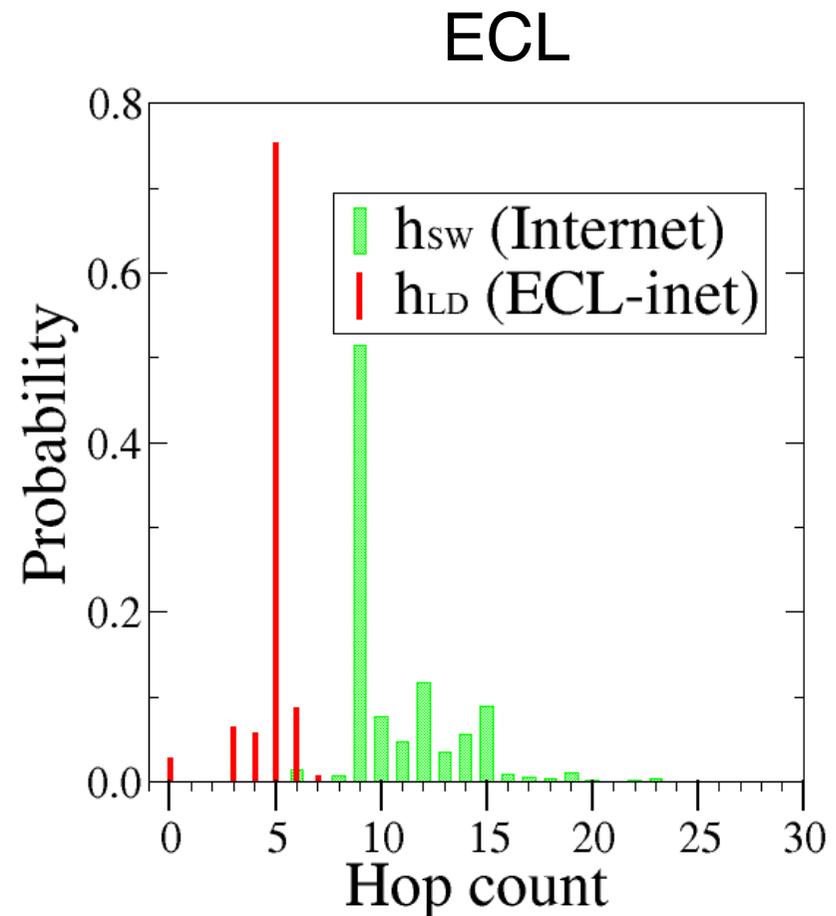
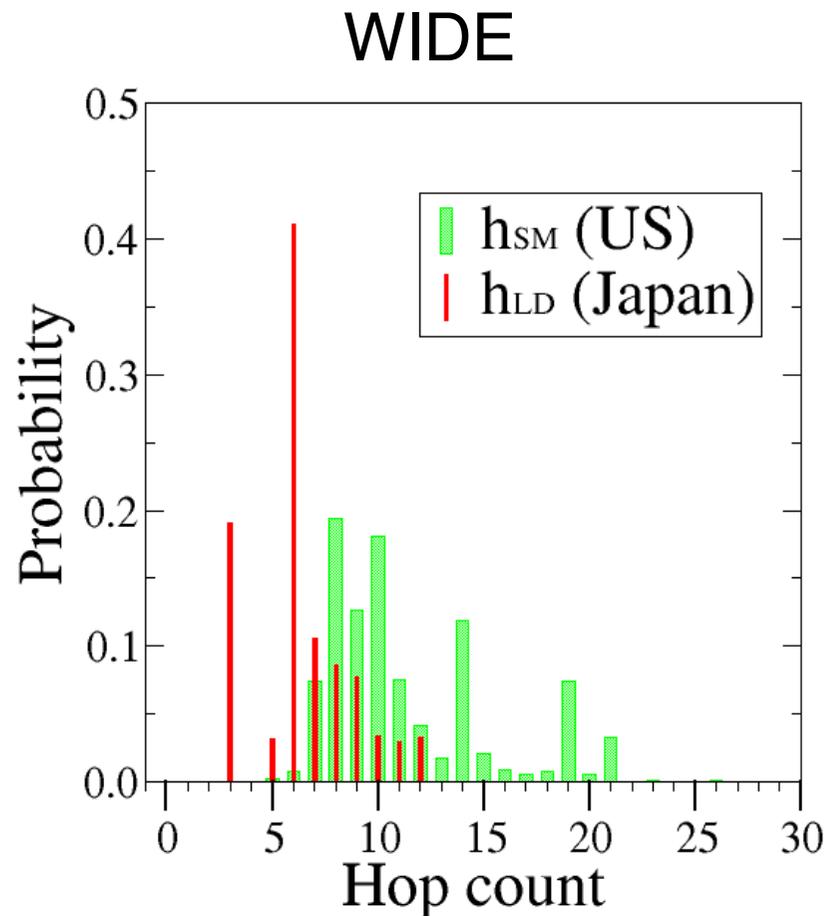
ホップカウント分布: (4/5)

全体



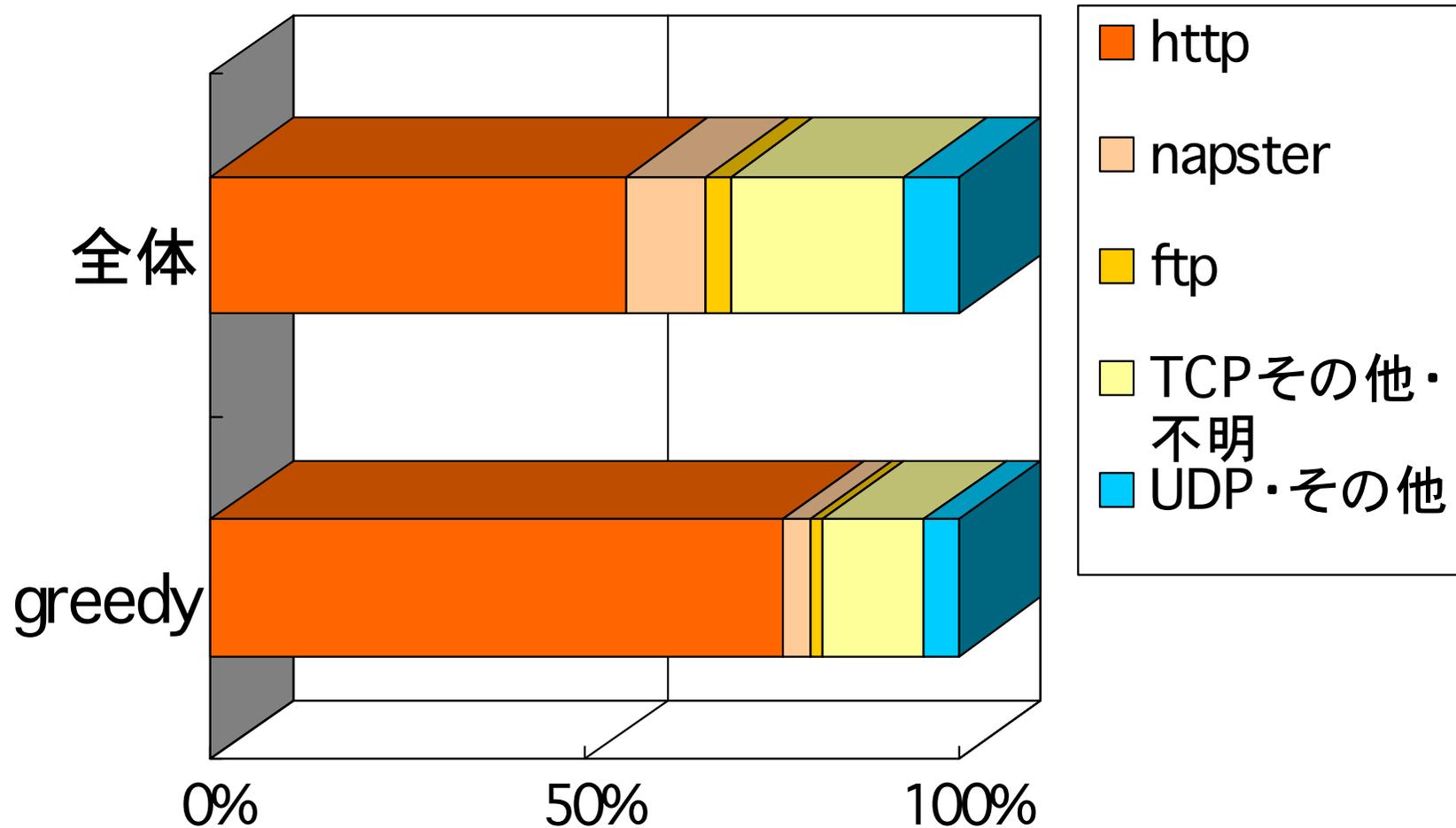
ホップカウント分布: (5/5)

Greedy フロー



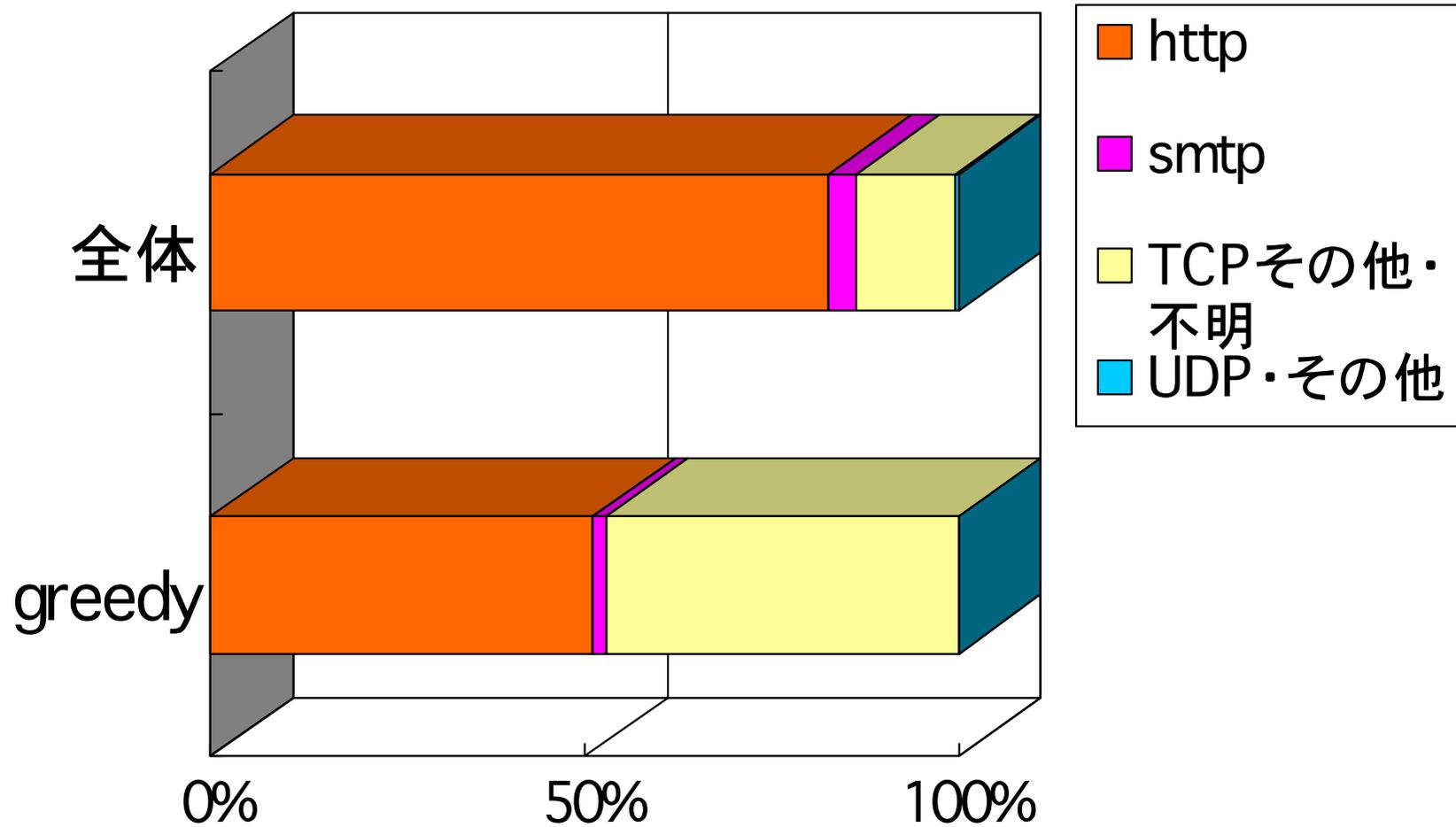
アプリケーション統計: (1/2)

WIDE



アプリケーション統計: (2/2)

ECL



まとめ

- Per-time-blockフロー統計

- サイズ分布

- べき則を示す (greedyフローの存在)
 - 遍在する普遍的な性質か？ → 更なるデータの解析が必要

- ホップカウント統計

- Greedyフロー・・・ホップカウント小の傾向
 - バックボーンでは外部、内部ともにあてはまる
 - アップリンクでは必ずしもそうではない(内部側)

- アプリケーション統計

- WIDEではhttpフローがgreedyに
 - ECLではその他・不明がgreedyに
 - proxyサーバ性能
 - ダイレクトアクセスホストのP2Pアプリ?
 - DDOS?

まとめ

- 今後の指針

- 短期間での変動特性をフローレベルでとらえた

- 平均レートで決定されるフローの重畳モデルの限界
 - Greedyフローをいかに抑えるか？ (制御)
 - Greedyフローを考慮に入れた帯域設計をする

- トラフィックモデルの提案

- 時間的に変動する発信源が、べき則に従うパケット数を単位時間あたりに確率的に送出するモデル

- トポロジー(ホップカウント)と流量変動の関係を調べた

- トポロジーを考慮したトラフィックモデル、プロビジョニング
 - Ex) ブロードバンドアクセス網に対してコンテンツを配信するネットワーク

ご清聴ありがとうございました

- Any Question?

参考文献

- [Polly Huang 2001] Polly, Huang, **Internet as a Dynamic System -- Research For Network Engineers And Control Theorists** , <http://www.tik.ee.ethz.ch/~huang/talk/>, 2001.
- [Tatsuya Mori 2001] Tatsuya Mori, Ryoichi Kawahara, A study on the marginal distribution of network traffic. IEICE Technical Report, IN 2001-107, pp. 1-7, 2001.
- [Tatsuya Mori 2002] 森達哉, 川原亮一, 内藤昭三, Per-time-block フロー統計に着目したネットワークトラフィック解析, 信学会, 信学技法, NS2001-224, pp. 1-8, 2002.
- [MAWI] <http://www.wide.ad.jp/wg/mawi/>