トラフィック観測によるネットワークの診断

工学資源学部情報工学科　**五十嵐　隆　治**

1．はじめに

インターネットはユビキタスIT社会においては必須のインフラとなっていて，現代の生活には欠かせないものとなった，と言っても過言ではないであろう。現在のインターネットの根幹をなすTCP/IPプロトコルIPv4は近い将来IPv6にバージョンアップされ，新しいプロトコル体系のもとでのWeb2.0も多様に且つ大きく進化（変化）して行くであろう。このような発展は，実は明と暗の部分がありインターネットの利用者もまた運用者も，いろいろな問題や事例に適切に対応しつつ進んで行くことが，今後ますます重要になるであろう。

多様に進化しているインターネットの使用形態の例を図1に示す。このようなインターネット上で疎通するトラフィックは正規のもののみならず，悪意あるものや，アプリケーションのタイミングによる異常なものも重畳されている。このような異常トラフィックの疎通を防止しつつセキュリティーを確保し，またアプリケーションのQoS（Quality of Service：品質）を保とうとする場合，当該LANの管理者の負担は膨大なものとなってしまう。管理者の負担を軽減し業務を支援する方法・施策はいろいろと提案されているが，ひとつのシステムで完結させることは至難であり，やはり管理者の経験に基づいた複数の施策を適宜採用するしかないのが現状であると思われる。このような場合，適切なネットワークの管理には適切なネットワーク診断が欠かせない。

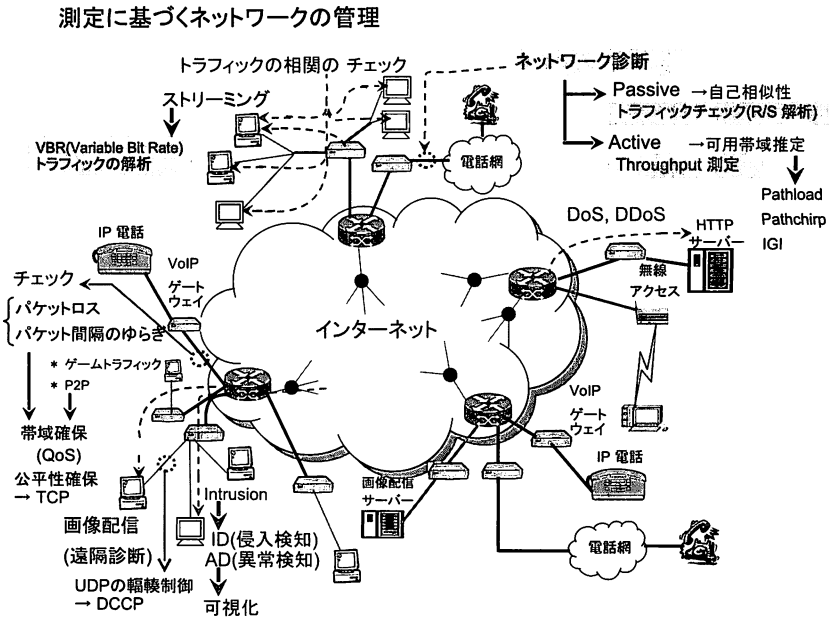

図1　いろいろなアプリケーションが疎通するネットワークの運用と管理

2．ネットワークの診断

インターネット上では，図1に示すように，多様なアプリケーショントラフィックが疎通していて診断も容易ではないが，「診断」という観点からは図2に示すように，パッシブ（受動的）な診断とアクティブ（能動的）な診断がある。受動的な診断は，ひとの診療では聴診に，また能動的な診断は打診にあたると考えていいであろう。両診断法とも，その使い方で長所・短所がある。ネットワークに余分な負荷をかけたくないときには，観測点でトラフィックを「覗く」だけの受動的な診断法が適する。多少の負荷にはなるが，アプリケーションのQoS確保のために，ある程度正確な可用帯域を推定したい場合には，能動的な診断法が好都合であると言える。現在では図2に示したような診断の結果を，その他広範なネットワーク情報と統合したネットワーク管理支援システムなども提案されていて，種々の診断結果はこのような管理システムの「エージェント」のひとつとなる。以下に，我々とセンターとの共同研究内容の一部も含めたネットワーク診断に関し概略を述べる。

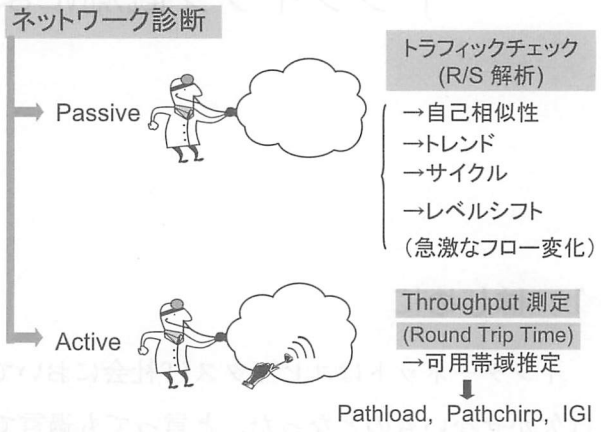

図2　ネットワーク診断：
受動的な診断と能動的な診断

2．1　能動的な診断

能動的な診断法の概略を図3に示す。本例は当該ネットワークの帯域の把握，すなわち当該アプリケーション疎通時に使用可能な帯域（可用帯域）を推定する方法で，プローブパケット（試験パケット）と呼ばれるパケットをネットワーク内に送信して測定する。このため被測定ネットワークには余分な負荷がかけられることになり，輻輳状況が深刻な場合には不都合さを生じさせることもある。TCPパケット通信ではRTT（ラウンドトリップタイム）の測定により再送タイムアウトを制御しているが，余分な負荷をネットワークに掛けたくないときにはRTTを利用し，簡易に混雑具合をチェックするのもひとつの方法であろう。

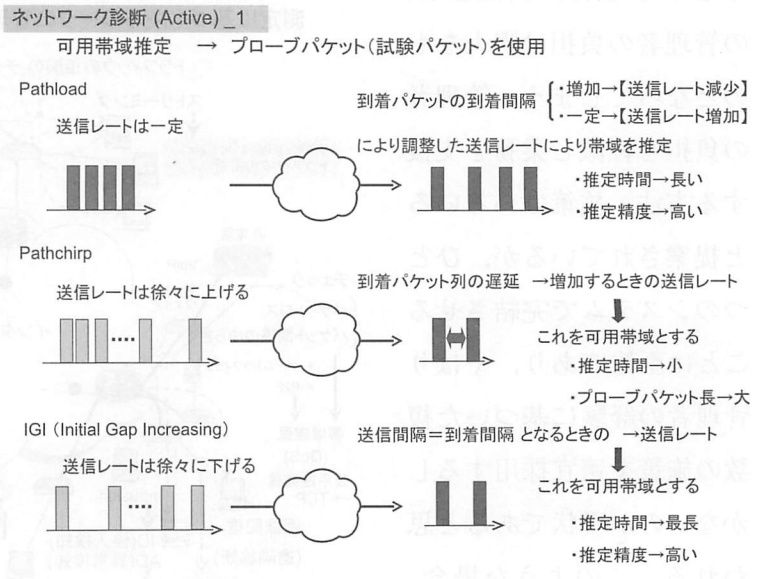

図3　ネットワーク診断：能動的な診断法の代表例

2．2　受動的な診断と測定ツール

受動的な診断においては図2に簡単に示したように，トラフィックを測定することになる。通常は注目しているルータにミラーリングポートを設けて測定するので，プローブパケットを用いる能動的な診断に比べ，ネットワークに余分な負荷が発生することはない。簡易診断には汎用のモニタリングプロトコルであるSNMPを利用するのが適当と思われる。SNMP環境下ではRMON（Remote Monitoring MIB）の利用，またはフリーウェアであるMRTG（Multi Router Traffic Grapher）の利用は監視・診断目的を限定する場合には非常に有用なツールとなる。取得情報の質の向上を意図する場合にはベンダが提供している計測技術を採用する方法もある。よく知られているものにCisco社のNetFlow，　InMon社のsFlowがある。これらNetFlow，sFlow はサンプリングベースの計測技術であるため，計測の負荷は小さくできる。計測の負荷が大きく，また保存データ量が大きくなる不都合はあるものの，汎用計測ツールとして使えるtcpdumpは使用者側でいろいろなデータ整理法を提案することにより，強力な解析ツールとなり得る。我々はこの観点からの研究も進めている。

3．パケットトラフィックの性質

3．1　トラフィックの性質とその意味

インターネット上のパケットトラフィックのふるまいが，従来の通信網で用いられていたポアソン過程によりモデリングできるのであれば問題はさほど複雑ではない。しかしパケットトラフィックは従来のポアソンモデルでは記述が困難であることが明らかにされて以来，ネットワークの設計・構築・運用・管理には新たな対応が求められるようになっている。パケット流は統計理論的では，２次の自己相似過程のうち，その記述パラメータであるハーストパラメータHが $1/2 < H < 1$ となるLRD（Long Range Dependence：長期依存性）過程に従っていると言われている。その特徴的な性質を図4に示す。

図4に示されているトラフィック時系列の大きな特徴は，時間スケールを変化させてもバースティネス（ゆらぎの度合い）が消失しないことである。トラフィック時系列に関して言えば，これは１次元のフラクタルということになる。同図中のReal Trafficが実測トラフィック，　Poissonがシミュレーションで得られたポアソントラフィック，FGN（Fractional Gaussian Noise；フラクショナル・ガウスノイズ）が長期依存過程をシミュレーションしたトラフィックである。この

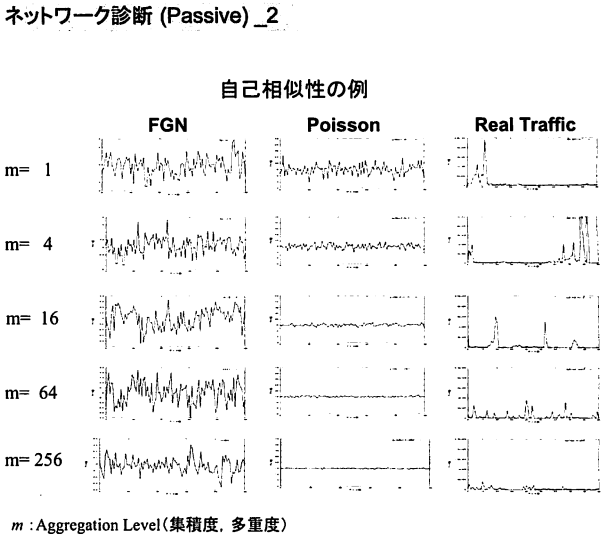

図4　パケットトラフィックの自己相似性

「バースティネスの保存」が興味あり且つ悩ましい問題を提示してくれるのである。ひとつはバッファ長をどのようにしたらいいのか，という問題であるが，バースティネスがどのような時間スケールでも消失しないのであれば，理論的には無限長のバッファが必要である，ということになる。これは現実には容認できないものである。もうひとつはネットワーク運用上の問題で，たとえば，トラフィックの異常フロー検知の閾値設定をどのようにしたらいいのか，という問題と関わってくる。すなわち，自己相似性に従ったバースティネスと異常フローとの識別が困難になり，単にフローに閾値を設定したのみでは異常トラフィックの判断が困難となる，ということである。これを回避するためには知識ベースの判断が必要ということで，いろいろな規模のネットワークにおける，いろいろな条件下でのトラフィックデータの蓄積と解析が必要であることを示唆している。

3．2　トラフィックのバースティネス

トラフィック測定により得られる諸量の一例を図5に示す。この例ではトラフィック時系列Xtにフローの急峻な変化が認められるが，他の量たとえば，到着時間間隔分布（IATD：Interarrival Time Distribution）と自己相関関数（ACF：Autocorrelation Function）を見ると，これが直ちに異常トラフィックであるとは言い難いようである。ハーストパラメータの推定法の一つであるR/S解析においてはPox Diagram（ポックスダイアグラム）が用いられるが，この形が図5に示すようになっている場合には当該測定時間帯におけるトラフィックはほぼ自己相似過程に従っている，とみなしてよい。事実，図5にあわせて示した到着時間間隔分布と自己相関関数からもこのことがうかがえる。到着時間間隔分布と自己相関関数の意味は図6に示した説明からも明らかであるが，従来の通信路で用いられていたポアソンモデル，ないしは無相関過程のモデルに従うランダム過程においては，ある時点でのイベントの影響がそれ以降に伝播することはなく，従ってその自己相関関数はδ関数となる。無相関ではない長期依存過程ならびに短期依存過程の自己相関関数はテイル分布（尾を引く関数）となる。またこれらを到着時間間隔分布で見た場合，ポアソン過程ではイベント間の間隔は指数分布となるのでこの log - normal プロットは図6のように直線となる。逆に言えば，無相関過程以外の過程のlog - normal プロットは曲線となることを意味し，図5に示した実測トラ

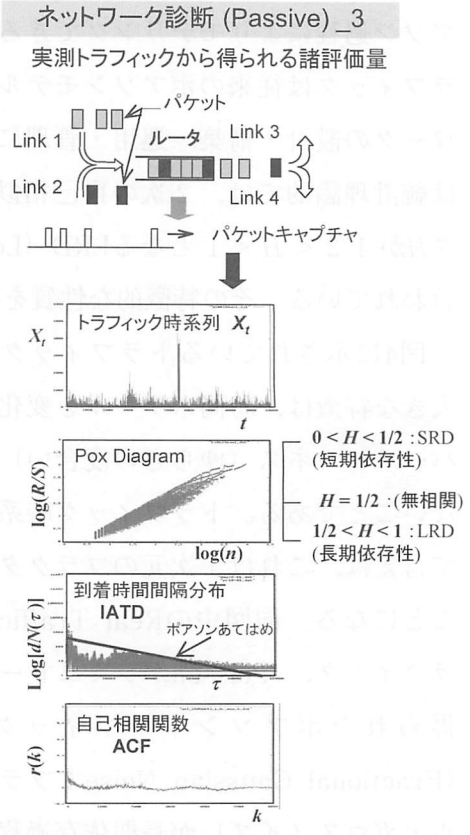

図5 パケットトラフィックの測定とそれから得られる諸特性

フィックは無相関な過程ではないことを示唆している。ただし，パケットトラフィックの研究の初期から指摘されていたことであるが，非定常過程も見かけ上は長期依存過程と似た性質を呈することがあり，実測トラフィックにおいてこれら両者の識別は困難な場合もある。トラフィックが非定常過程に従っているときには，管理者としては異常トラフィックの閾値設定は多少悩ましい問題となってしまう。この辺に関しては今後の検討課題であり，研究面からも大変興味深いテーマである。この非定常過程の問題は「擬似自己相似性」とも関わる問題で，この観点からの検討も今後の大事な課題である。以下にその一例を述べる。

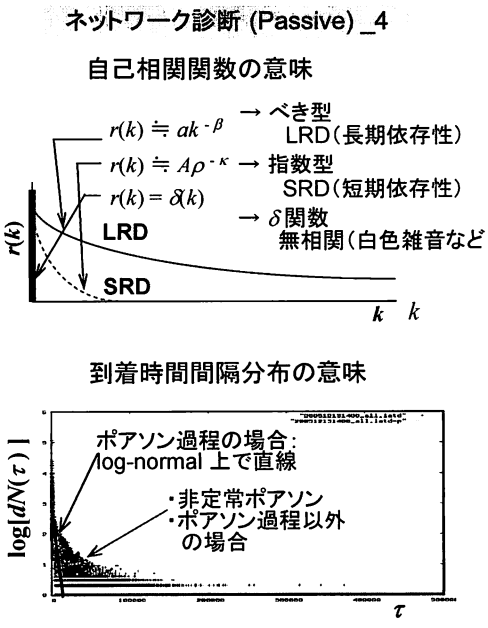

図6　パケットトラフィック時系列における自己相関関数と到着時間間隔分布の意味

3．3　擬似自己相似性

図7に教育研究機関におけるトラフィック時系列の例を示す。いろいろな特徴が認められるが顕著な例のひとつは，長期間（この例では12日間）で観測されるトラフィックには明確なDiurnal Variation（日間変化）が認められることである。この特徴はPox Diagramと自己相関関数に認められる。Pox Diagramには，FGN過程に認められる特徴から推移した傾向が認められ，自己相関関数には明らかな日周期が認められる。また1日間のトラフィックでもフロー量は時間によって異なっていて，フロー量が小（Low），中程度（Normal），大（Busy）であるときのPox Diagram，到着時間間隔分布，自己相関関数には各々相違が認められる。この例ではトラフィック時系列がポアソン過程にも，また理論的な長期依存過程にも従っていないであろうことが推測される。これよりトラフィックのモデリングないしは解析は大変難しそうである，ということがよくわかる。図7に示した諸量の特徴はモデリングの困難さは提供するものの，TE（トラフィックエンジニアリング）の視点からはひとつの有用な知見を提供している，と考えられる。すなわち各グラフの理論形状からの推移は，実トラフィックの何らかの状態を表

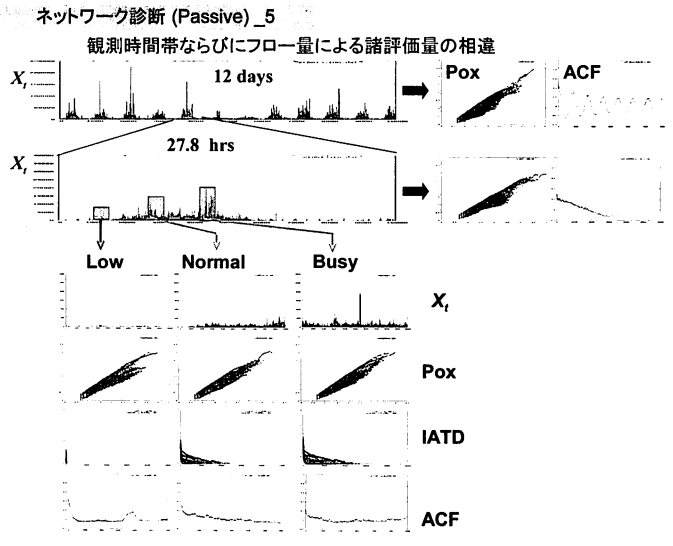

図7　長時間に渡るパケットトラフィックの測定例とトラフィック時系列から得られるPox Diagram，到着時間間隔分布，および自己相関関数．12 日間に渡る測定結果からはトラフィックにDiurnal Variation（日間変化）が認められることがわかる．

していて，この理論値からの推移部分と当該トラフィック特性の関係を明らかにすることで，ネットワークの適切な診断と運用に十分寄与し得るモジュールの開発が可能となる。この際，トラフィックのどのような特徴が図7のグラフのような傾向を与えるのか，を把握しておくことは非常に有益である。図8にこの関係を把握するための試みの一例を示す。

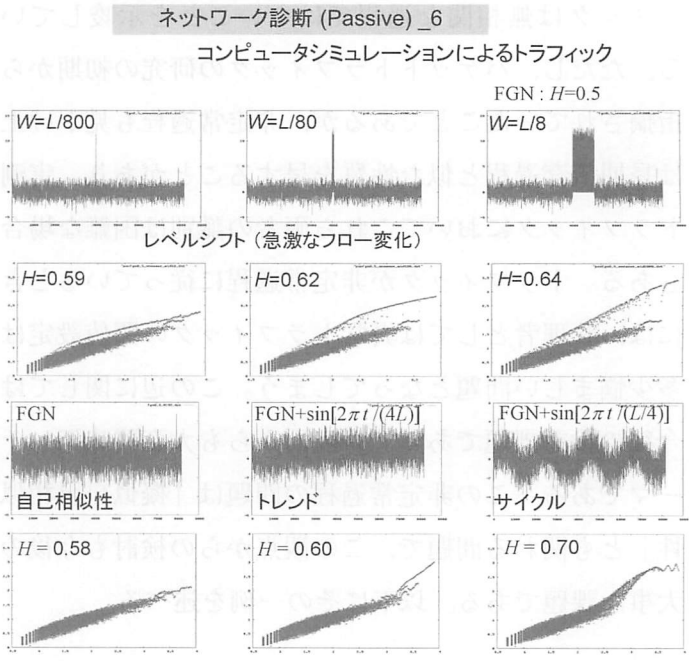

図8　コンピュータシミュレーションによるパケットトラフィックの自己相似性からの推移の検討例

4．シミュレーション

図8は理論的な長期依存過程であるFGNに，代表的な非定常性を呈するレベルシフト（急激なトラフィック変化に対応，たとえばFlood攻撃など）とトレンド（たとえば図7に示したトラフィックの1日内の変化など），ならびにサイクル（たとえば図7に示した長期間トラフィックの日間変化など）を重畳させた時のPox Diagramの変化傾向を示す。Pox Diagramに特徴的な変化が認められていることが明らかである。レベルシフト重畳においてはPox Diagramの分散傾向が顕著となっている。それでは実際のネットワークにおいてもこのような変化が観測されるのであろうか？　この疑問に答えるために実ネットワークを用い，図9に示すようなシミュレーション実験を試みた。実ネットワークとはいっても実際に稼動しているネットワークで擬似攻撃を試みることはできないので，研究室に設置しているネットワークを用いて，図7に示したレベルシフトに対応するFlood攻撃を実施してみた。この結果図7ないしは図8に示した例に類似したPox Diagramの分散が認められた。なお，この実験においてはPox Diagramを高速に求めるアルゴリズムを新たに開発し，リアルタイムで図9に示したようなグラフを描画した。図9の結果はひとつの有益な示唆である。実トラフィックのPox Diagramならびにハーストパラメータをリアルタイムで導出できれば，これをモジュールとして，たとえば管理者支援用のエージェントのひとつとして，ネットワークの運用・管理システムへ実装することが可能となる。

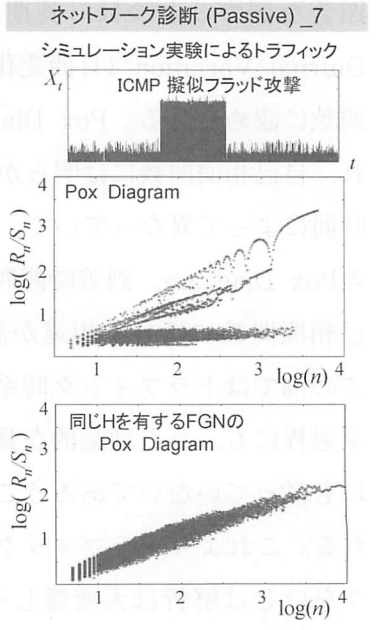

図9　シミュレーション実験によるパケットトラフィックの自己相似性からの推移の検討例