

## TCPを用いた通信ネットワークの自己相似性に関する研究 Self-similar behavior of Communication networks using TCP

杉山 吾一 松葉 育雄  
† Goichi Sugiyama † Ikuo Matsuba

### 1. まえがき

近年のインターネットの急激な発達と共に、待ち時間の増加やレスポンスの遅れなどが目立つようになってきた。それらは、従来のポアソン分布のモデルでは説明できない現象であり、これを測定したトラフィックデータより説明しようというのが通信トラフィック理論である。トラフィックの測定はパケットが観測ポイントに到着した時刻と到着したパケットサイズを記録することによってなされる。また、その分析は適当な時間間隔で集計されたパケット数やデータ量の時系列を解析することによってなされる。トラフィックデータを様々な角度から分析したところ、到着パケット数の集計時間単位を大きく変化させても得られる時系列データはほぼ同様で統計的に同じ挙動を持った、激しい変動を示す。これが、従来のポアソン分布では説明できなく、このことを自己相似性という。

さらに、ここ数年の研究でこの自己相似性は必ずしも成立するものではなく、相転移現象として捉えることが自然であるという指摘がなされている[1]。相転移モデルは自明でない臨界点を挟んで非輻輳相と輻輳相に分け、その臨界的な状態でのみ自己相似性が成立するという立場を取る。自己相似性を内包する相転移現象で特徴付けられるネットワークがどのようにして生成されるかを解明することは将来のネットワーク設計において重要な意味を持ち得る。なぜかといえば、臨界点、つまり自己相似性が観測される状態がパケット通信において最も効率の良い状態であるという可能性が示されているからである。

自己相似性を生成する原因としては様々な要素が指摘されている。例えば、ARQの再送効果[2]、確認応答パケットの圧縮[3]、確認応答パケットが他のコネクッションに割り込まれること[1]が自己相似性の原因となっているという指摘もある。本論分では、TCPの高信頼化サービスという角度から自己相似性について考えているが、中でも今回はフロー制御に関わる機能が自己相似的なトラフィックの生成に果たしている役割を考察する。

### 2. シミュレーションおよび解析方法

#### 2.1 シミュレーションモデル

シミュレーションは図1のようなネットワークを用いて行う。このモデルは近年研究されているケリー・トゥリーと呼ばれるノード配置を利用して近代ネットワークに適したモデルを考察する。このノード配置は、ルータ同士が接続されているネットワーク構造に近いため、実際のネットワークに見られる自己相似性をモデル化していくのに適

している。既存の研究より、1次元的なモデルや格子状のモデルより実際のネットワークに近い解析結果が出ていることからこのことは証明されている。

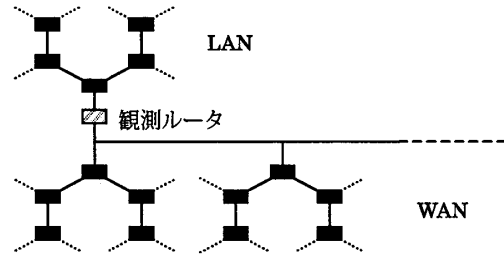


図1 ネットワークモデル

更に、本モデルではケリー・トゥリーで形成されるLANを  $n$  個重ね、大規模ネットワークを形成している。上側のLANと下側のWANに分け、その中継ノードにおいて観測されるトラフィック量を解析データとした。実際のネットワークにこのモデルを当てはめると、上側のLANがLAN内のノード端末に、WANのノードがWWWサーバのようなサーバ類に対応する。

#### 2.2 解析方法

トラフィックの自己相似性の解析には、ハースト解析という方法を用いて、ハースト数を求めることにより行う。ハースト数の計算方法について述べる。

$x_j(i)$  を始点  $j$  から始まる  $i$  番目のデータ  $x_j(i) = x_{j+i-1}$

$$\langle x_j \rangle_a = \frac{1}{a} \sum_{i=1}^a x_j(i) \quad (1)$$

とする。平均値を

と定義し、以下の2つの統計量を準備する。

$$S(a) = \max_{1 \leq k \leq a} \sum_{i=1}^k (x_j(i) - \langle x_j \rangle_a) - \min_{1 \leq k \leq a} \sum_{i=1}^k (x_j(i) - \langle x_j \rangle_a) \quad (2)$$

$$V(a) = \left[ \frac{1}{a} \sum_{i=1}^a (x_j(i) - \langle x_j \rangle_a)^2 \right]^{1/2} \quad (3)$$

具体的な計算手順としては、まず区間  $a$  を定めておき、始点をいろいろ変化させて異なる多くの平均値  $\langle x_j \rangle_a$  を作成する。  $x_j(i) - \langle x_j \rangle_a$  を  $i = 1$  から  $i = k$  まで足し合わせる。  $k$  を1から  $a$  まで変化させ、その和の最大値と最小値の差を計算する。これが、  $S(a)$  を与える。  $S(a)$  は平均値からの離れ具合の大きさを示す。そして  $R(a) = S(a)/V(a)$  を計算する。特に、大きい  $a$  で、

† 千葉大学自然科学研究科知能情報工学専攻

† 千葉大学工学部情報画像工学科

$$R(a) = (a/2)^{H_H} \quad (4)$$

を満たす時、その指数として得られた値  $H_H$  をハースト数  $H$  の推定値とする。一般に、 $H = 0.5$  の時に短期記憶特性で時間相関は無く、 $0.5$  より大きい時は長期記憶特性で時間相関を持つ。

### 3. シミュレーションによる解析

図2にシミュレーションの大まかなプロセスを示す。

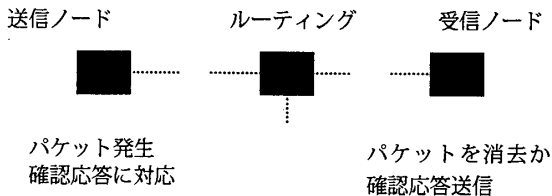


図2 シミュレーションのプロセス

TCP はネットワークの現在の状況に応じて、送信レートを変化させるアルゴリズムを採用している。送信レートを変化させることは、長期的な時間

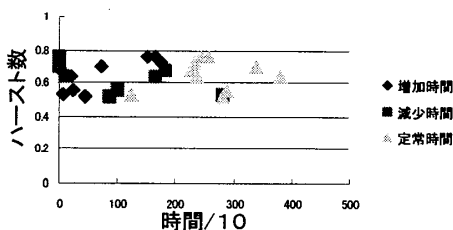


図3 送信レートとはハースト数の関係

相関を生み出し、自己相似的なトラフィックを生成する可能性がある。ここでは、単純な送信レート増減をモデルに組み込んで、その影響を調べる。各ノードでは1プロセスにパケットが確率  $p$  で発生し、宛先は乱数により決定される。パケットは1プロセス1つのノードを移動し、目的地に着いたらパケットは消滅する。また、オーバフローしてしまったパケットも再送せずに単純に消滅する。シミュレーションの状態を決めるパラメータとして  $n = 3$ 、初期状態として  $p = 0.5$  とし、10プロセスごとにネットワークに余裕があれば  $p = 0.7$  に、一定以上のパケット棄却率になったら  $p = 0.3$  とする。このようなプロセスを 5048 回繰り返し、 $t > 1000$  を解析対象とする。

30回のシミュレーションの結果、 $p = 0.5$  に固定した場合と送信レートを変えた場合のハースト数の平均値はそれぞれ、0.52, 0.66 となった。また、送信レートとハースト数の関係も図3に示した。図3から、 $p = 0.7$  に保たれている時間が長いほど大きいハースト数を示していることがわかる。つまり、送信レートを上げてパケットの棄却率が低く抑えられるネットワーク構成が成された(ネットワークのノード配置は毎回変わっている)時に、自己相似的な性質が確認できるということである。

次に、確認応答による再送処理のもたらす影響について調べた。TCP では、送信側があらかじめ送った情報に基づき、受信側がパケットを正しく受信したかどうか確認応答がなされる。もし、データの一部が誤っていたら受信側は送信側にパケットを再送するように要求する。ここでは、データの到達順序などは考慮せず、あくまで全てのパケッ

トを受信したかどうかだけを考える。シミュレーションにより、再送処理が無い

表1 パケット誤り率とハースト数の関係

k	0.001	0.01	0.1
ハースト数	0.58	0.65	0.66
	0.56	0.64	0.86
	0.56	0.61	0.7
	0.57	0.63	0.79
	0.55	0.63	0.72

トを再送する間にパケットに誤りが出る確率を  $k$  とする。データは全てのノードで発生するものとし、そのサイズは 5Kbyte とする。パケットは 1Kbyte に固定する。シミュレーションは  $k$  の値を変え、各5回ずつ行った。確認応答に基づく再送処理がある場合のハースト数は表1のようになった。

表1より、パケット誤り率が高く再送が頻繁に行われる方が高いハースト数が出ていることがわかる。この結果から、確認応答に基づくパケット再送がネットワークの自己相似性に関与しているということが推測できる。

### 4. 結論

本研究では、TCP の高信頼化サービスが通信ネットワークの自己相似性に与える影響について調べた。そして、それらのサービスの一部がパケットの時間相関を強くし、自己相似的な振る舞いを生み出している可能性を示唆した。

シミュレーションの結果から予測されることは、主たる原因となっているのは、TCP のパケットの到着率をある一定の値に保つための送信レートの増減で、確認応答に基づく再送がそれを補助していることである。ただし、それらのサービスは TCP 独自のものに限らず、それに類似したものでも構わないことがわかった。また、それらの前提として、あまりにも頻繁にパケットロスが起こるようなネットワークでは自己相似性は観測されにくいことがわかった。

今後の課題と考えられるのは、実際のネットワークで観測されるハースト数は 0.9 以上で非常に強い時間相関を持っているのにも関わらず今回観測できたのは、0.9 以下にとどまったことから、更に強い時間相関を持ったネットワークの構成要素を探っていくことなどが挙げられる。

### 参考文献

[1] K.Fukuda, M.Takayasu, H.Takayasu, "Self-similar Traffic Originating in the Transport Layer", IC2000, 2000.  
 [2] J.M.Peha. "Protocols Can Make Traffic Appear Self-Similar" In Proceedings of the 1997 IEEE/ACM/SCS Communication Networks and Distributed Systems Modeling and Simulation Conference.IEEE/ACM/SCS, 1997.  
 [3] A.Fwldmann, A.C.Gillbert,P.Huang, and W.Willinger. "Dynamics of IP Traffic : A Study of the Role of Variability and the Impact of Control. In Proceedings of SIGCOMM'99, pp. 301-313, Cambridge, USA, 1999. ACM. "