

情報通信システムのモデル化手法とOR：マルコフモデルを超えて

01207040 千葉大学 塩田茂雄 SHIODA Shigeo

1 はじめに

情報通信とオペレーションズリサーチとの関係は深い。特に、待ち行列理論は、電話網の設計のために、コペンハーゲンの電話会社に勤める A.K.Erlang が確率論を応用した論文を 1908 年に発表したことに始まるとされる。それ以降も、情報通信技術の進展に刺激を受けながら、待ち行列理論は大きく発展してきた。

近年、インターネットの急速な普及に代表されるように情報通信環境は大きく変貌を遂げたが、その一方で、従来の待ち行列モデルの限界が指摘されている。本稿では、従来の待ち行列モデルの限界と、モデルの限界を克服するための様々な試みについて概説し、モデル化手法に関する私見を述べる。

2 マルコフモデルの成功と限界

2.1 情報通信システムと待ち行列モデル

世の中にはさまざまな情報通信システムがあるが、ここではコンピュータネットワークをとりあげる。コンピュータネットワークとはコンピュータや中継装置が相互に通信回線で接続されたネットワークである。インターネットもコンピュータネットワークの一種である。コンピュータネットワークではパケットと呼ばれるデータのかたまりを中継装置がリレー方式で通信相手のコンピュータに転送する。中継装置はパケットを受け取ると、次の中継装置に通信回線の太さ（帯域）で決まる速度でパケットを転送するが、たまたま、複数のコンピュータ（もしくは中継装置）からパケットを大量に受け取ると、次の中継装置への転送が間に合わず、一時的にパケットの「待ち行列」が形成されてしまう。

一般に、情報通信システムでは、このように一時輻輳による待ち行列がシステム内に頻繁に形成され、ネットワークの性能に大きな影響を及ぼす。これが、情報通信ネットワークにおいて待ち行列理論が使われてきた所以である。

2.2 待ち行列理論とマルコフモデル

待ち行列理論ではケンドールの記法 (Kendall's notation) を用いて待ち行列システムを表す。例えば、「M/M/1」は客がポアソン過程に従って到着し、客の要求するサービス時間が指数分布に従い、窓口が 1 つの待ち行列システムである。ポアソン過程もしくは指数分布を表す「M」という記号は、マルコフ性 (Markovian) もしくは無記憶性 (Memoryless) の頭文字をとったものらしい。ポアソン過程や指数分布はいずれもマルコフ性を持つ確率モデルであり、マルコフ性はある種の無記憶性ということができる。ポアソン過程や指数分布という、必ずしもあまり身近でない確率モデルが

待ち行列理論で重宝されてきたのは、マルコフ性（または無記憶性）という性質がモデルを解析する上で極めて便利であったためである。

待ち行列モデルで厳密な解析が可能なモデルは、全ての「マルコフ性」を持つモデルであると言ってさしつかえない。事実、解けるモデルにはどこかで「M」（もしくは、その親戚の「MAP」や「PH」）という記号が使われている。

コンピュータネットワークでは、客がパケットに相当し、サービス時間はパケット長に相当する。パケットがポアソン過程に従って到着すること、もしくはパケット長が指数分布に従うことが、かつて厳密に検証されてきたかどうかは疑わしいが、少なくとも以前は、「M/M/1」のような簡易な待ち行列モデルが、それなりに使われてきたのである。

2.3 マルコフモデルの限界

最近の通信システムの性能評価の論文では、以前ほどマルコフモデルは使われなくなってきている。これは、マルコフモデルの限界が 1990 年代頃から指摘されてきたためである。

その理由の一つは、モデル同定の複雑さにある。インターネットでは文字・音声・画像など各種のデータがネットワーク上を転送されるが、例えば動画のデータ到着過程は明らかにポアソン過程より複雑である。しかし、ポアソン過程より複雑なマルコフモデルを使おうとすると、モデル同定（モデルのパラメータ決定）が途端に破綻してしまう。

例えば、ポアソン過程の最も単純な拡張として、「2 状態マルコフ変調ポアソン過程」がある。これは到着率の異なる二つのポアソン過程が交互に繰り返されるというモデルである。この単純なモデルであっても、4 つのパラメータ（各ポアソン過程の到着率と平均継続時間）を含み、その決定には煩雑なデータ解析が必要である。ネットワーク設計では 1 年後の需要予測等が必要になるが、需要モデルに複数のパラメータを含むマルコフモデルを使えば、各パラメータ値の予測技法も必要である。確かにマルコフモデルは解くことはできるが、モデルの同定がネックになりやすい。

2 番目の理由は、実システムのマルコフ性に対する疑問であり、これはベル研の W.E. Leland らの研究 [2] 等に端を発している。彼らは 1989 年から 1992 年にわたり、ベルコア内の LAN の膨大なデータを詳細に解析し、コンピュータネットワークのパケット到着過程はもはやマルコフ性を持たないこと（正確には長時間相関性を持つこと）を「証明」してみせた。それ以来、インターネットコミュニティでは、「マルコフ性」を陽に仮定したモデルは、時代遅れのものになってしまった。マルコフモデルを擁護する研究結果（最近のインターネットのトラヒックはマルコフモデルが通用

する、等)も発表されているものの、「マルコフモデル」が復権する気配はあまり無い。実は、「マルコフ性」を使わなくとも待ち行列モデルを解析する道具が準備されてきたのがその理由の一つである。これについては次に述べる。

3 マルコフモデルを超えて

情報通信の分野では、1990年代以降、マルコフモデルを用いない幾つかのモデル化技法が研究されてきた。以下、これらについて簡単に紹介する。

3.1 Deterministic model

インターネットコミュニティでは、1990年代前半から、End-to-Endでのキューイング遅延の最大値を制御する研究が流行し、そのためのモデル化・解析技法が編み出されてきた。これを「deterministic model」と呼ぶ。待ち行列では、通常、待ち行列システムへの入力量を確率モデルで表現するが、deterministic modelでは入力量の上限を確定的に与える。例えば、

$$t \text{ 時間内の入力量} \leq \min\{pt, at + \sigma\}.$$

ここで、 p はピーク速度、 a は平均速度に相当し、 σ はバーストサイズと呼ばれる。このように入力量の上限を確定的に決めることは、実際に理にかなってない。というのは、1台のコンピュータからの入力量は、通信回線の帯域やウィンドウサイズ、符号化技術などの要因により確定的にその上限が決まるからである。また、このモデルはCACやトラヒックレギュレーションなど、実際のネットワーク制御との相性も良い。さらに、deterministic modelにより、キューイング遅延の最大値だけでなく、分布も評価できることがわかっている。deterministic modelは、インターネットコミュニティでは、現在、標準的な性能評価技法として認識されつつある[1]。

3.2 極限定理の利用

モデルを「厳密」に解くことにはさほど実用的な意味はない。なぜならば、前提とするモデルに、すでにモデル化誤差が含まれているからである。厳密性を放棄することにより、マルコフモデルの壁は容易に超えることができる。

確率論における極限定理(中心極限定理など)は、厳密性にこだわらなければ極めて強力な解析手段を提供する。例えば、多数のコンピュータからのデータ入力のある待ち行列を考える。コンピュータの台数が十分に大きい極限(バッファや処理速度もコンピュータ台数に合わせてスケールさせる)では、一台のコンピュータからの入力データの確率特性に関わらず、待ち行列の確率的な振る舞いは、ある単純な法則に従うことが極限定理により保障される。極限定理のポイントは、確率モデルの詳細に関わらず成立することであって、従ってマルコフモデルに従わなくとも良い。

極限定理には様々な種類があり、中心極限定理は従来から良く用いられてきた「道具」の一つであるが、1990年代以降、情報通信システムの性能評価の分野で最も成功を収

めたのは「大偏差原理」であろう。パケットの廃棄確率などを 10^{-9} といった極めて小さい値に制御しようとするとき、大偏差原理はその効力を発揮する[4]。

3.3 ノンパラメトリックアプローチ

2.3章で述べたように、マルコフモデルの欠点の一つはモデル固有の煩雑さにある。それなら、いっそのこと観測できる量からダイレクトに(確率モデルを介さずに)システムを評価したらどうであろうか。つまり、背後にある確率モデルの存在を信用せずに、直接見える現象のみを信用してモデルを解析しようという試みである。(これは、量子力学の「観測」の考え方に相通ずるところがあるかもしれない!)

例えば、1ミリ秒間に待ち行列に入力されるデータ量をシーケンシャルに測定し、その分布(相対頻度)を求める。普通ならば、その分布に適合する確率モデルを見つける(同定する)という手続きが入るが、この分布情報から直接的に待ち行列長やロス確率などを評価するわけである。実は、このようなアプローチは、ORの領域ではこれまで殆ど検討されてこなかった(と思われる)が、情報通信の分野では実際に検討例があり、興味深い結果が報告されている[3]。

4 現場では

以上は全て「研究」上の話であって、実際にネットワークの設計・管理業務を行う現場で、紹介した手法が使われているかどうかは別である。筆者は、NTT東日本において、電話網の設備保守業務を行う、いわゆる現場サイドの部署で働いた経験がある。このような部署の担当者はもちろんORという言葉を知らない。そもそも、研究所を除くと、本社組織にしろ「M/M/1」の意味を解する担当者がある程度である。また、実システムは大変複雑であって、エレガントなORの解法はなかなか適用できない。実システムに矛盾しないモデル化を模索するほど、高度な数学が必要になるというジレンマもある。一方で、情報関係の資格試験では相変わらず(実際にはあまり役に立たない)「M/M/1」モデルの設問が出題されている。ORと現場の溝は深い。

参考文献

- [1] C. Chang. *Performance Guarantees in Communication Networks*. Springer-Verlag, 2000.
- [2] W. Leland, S. Taqqu, W. Willinger, and D. Wilson. On the self-similar nature of Ethernet traffic. *IEEE/ACM Trans. Networking*, 2:1-15, 1994.
- [3] H. Saito. Call admission control using upper bound of cell loss probability. *IEEE Trans. Commun.*, 40(9):1512-1521, 1992.
- [4] A. Shwartz and A. Weiss. *Large Deviations for Performance Analysis*. Chapman & Hall, 1995.