

サービス工学におけるシステム応答時間のモデル化

○魚住光成(三菱電機(株)) 浅間一(東京大学)

One idea of a system response model in service engineering

*Mitsunari UOZUMI, MITSUBISHI Electric Corp., Hajime ASAMA, The Univ. of Tokyo

Abstract —It is necessary to evaluate the performance of the computer system that provides service at response time. We thought that the response time is expressible by normal distribution. As a result of the experiment, it is able to be confirmed to show the response time of the computer system by normal distribution in simple service process.

Key Words: Service engineering

1. はじめに

人にサービスを提供するシステム、たとえばATMや券売機、ECサイトなど多種多様なサービスが、コンピュータと通信ネットワークを使って提供されている。また、ロボットや知能メカトロニクスが、人にサービスを提供する機会もこれから増えていくと予想される。

これら機械が人にサービスを提供するとき、システムの負荷状況によって時間性能は変動し、サービス品質の変動として利用者への影響が現れる。

我々は、サービスを提供するシステムの応答時間をモデル化し、利用者への影響の定量化を試みた。

2. サービス工学とサービスメディア

サービス工学では、サービスを提供するシステムをサービスメディアと位置づけ、これがサービスを提供するだけでなくサービスの評価計測を行い、利用者の満足度としてシステムにフィードバックするモデルを提唱している。^[1]

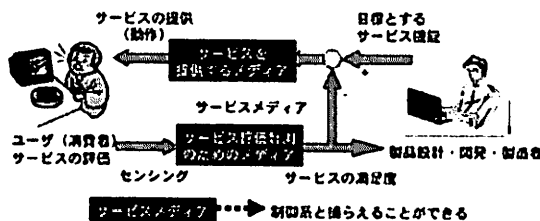


図1 サービス工学のモデル

このモデルにおいて、利用者の満足度は、サービスのコンテンツによる変化と、サービスメディアの品質によって変化すると考えられる^[2]。サービスメディアの品質は、サービスコンテンツを利用者の要求に従って引き出すコンピュータなどの機器の時間性能と、利用者までサービスコンテンツを伝達する通信ネットワークの通信品質によって変動する。

本研究では、このサービスを提供する機器の時間性能のモデル化を試みた。

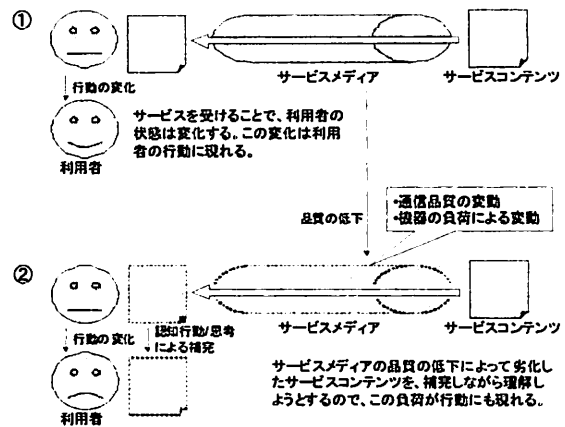


図2 サービスメディアの品質

3. コンピュータの時間性能の変動

コンピュータの処理性能の向上により、1台のコンピュータが複数の利用者に対し、同時にサービスを提供する事は一般的になっている。このときの処理性能は、スループットとして表されることが多い。

スループットは、単位時間当たりの処理件数や出力したデータ量であり、コンピュータの処理能力をわかりやすく表す指標である。システム設計を行う上では、スループットが線形に増加する範囲の多重処理を実用範囲として想定することもある。コンピュータの多重処理では、一定の多重度までは線形にスループットが高くなるが、それ以降はスループットが頭打ちとなる。この領域では、個々の要求の、結果が得られるまでのターンアラウンド時間(応答時間)が、多重度が上がるにつれ低下していく。その為、システム設計においては、スループットが頭打ちとなるより前の多重度を使用上の上限とすることが多い。

4. スループットを指標とした場合の課題

スループットは、コンピュータの処理能力の指標となるが、利用者個々の応答時間を推定することは難しい。多重度が上がるにつれスループットが上昇しても、個々の要求の応答時間については、均一であるとは限らない。

サービスメディアの品質は、利用者からみたシステムの品質であるため、スループットではなく、応答時間に基づく指標であるべきである。

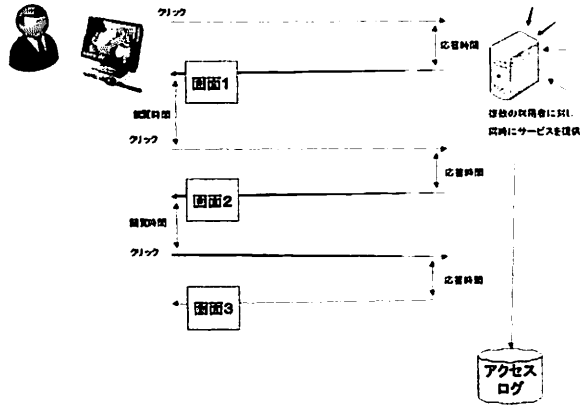


図 3 システムと応答時間

5. ターンアラウンド時間と多重度

コンピュータは、外部からの要求を単純に逐次処理するだけでなく、可能な場合は、複数の処理を同時並行に実行する。この傾向は、CPUのマルチコア化などの影響で強くなってきている。その為、要求を受け付けてから結果がでるまでの時間を正確に表現することは難しい。

しかし、統計量としては多重度が上がるにつれ、応答時間の平均と分散が上昇することが考えられる。

特定の多重度における応答時間は、正規分布となると仮定し、ある多重度における平均を μ_i 、分散を σ_i^2 とすると、

$$N(\mu_i, \sigma_i^2) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left\{-\frac{(y-\mu_i)^2}{2\sigma_i^2}\right\} \quad (5.1)$$

と表すことができる。

6. 実機による計測

今回、あるコンピュータに対し複数の利用者を模して多重度を上げる実機環境を構築し、計測を行なった。

対象としたコンピュータは、複数のサービスコンテンツを有し、利用者もそれらを組み合わせて使用するため、一連のサービスコンテンツのアクセスを連続して自動的に行なうプログラムを利用者を模す疑似端末に実装し、これを多重に実行することで、多重度を上昇させた。

計測は、個々のサービスコンテンツの要求から結果取得までの応答時間を疑似端末に記録した。

7. 計測結果の分析

計測した応答時間のヒストグラムを作成し正規化した。図 3、図 4 は、それぞれ、サービスコンテ

ンツA、サービスコンテンツBのヒストグラムである。便宜上、グラフは多重度毎の折れ線で表している。

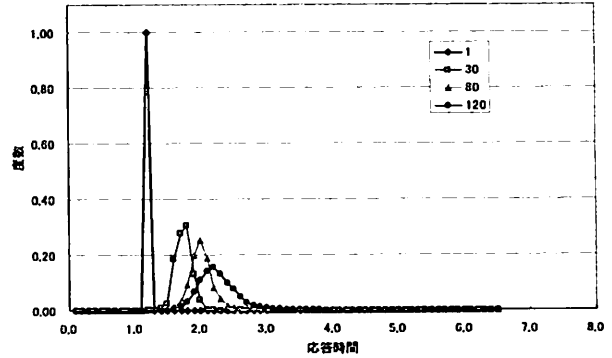


図 4 サービスコンテンツ A の応答時間の分布

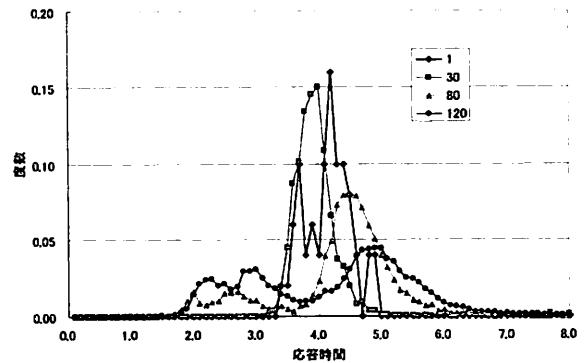


図 5 サービスコンテンツ B の応答時間の分布

サービスコンテンツAは、コンピュータ内で保持しているデータで要求に対する応答が生成できる。こうした処理の場合は、式(5.1)として想定した傾向であるといえる。サービスコンテンツBは、複数回のディスクアクセスを伴う処理である。こうした処理の場合は、多重性能を向上させることを意図したコンピュータ内のアルゴリズムの影響を受け、多重度が上がるにつれ、単純に応答時間の平均と分散が上昇するとは言いえないことがわかる。

8. 結論

サービスを提供するコンピュータの応答時間は、多重度が上昇するにつれ、その平均と分散は上昇する。処理が単純なものであれば、応答時間の分布は正規分布として捉えることができるが、処理が複雑な場合は個別に吟味が必要である。

複雑な処理を伴うサービスの応答時間のモデル化が今後の課題である。

参考文献

[1] 浅間一：サービス工学とシステムインテグレーション、計測と制御 278-283 Vol. 44, 2005.
 [2] 魚住光成, 浅間一：サービス工学における利用者のインタラクション評価、第25回日本ロボット学会学術講演会 1E35, 2007