A test of measuring method of satisfaction rating in service engineering

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

Abstract — We tried to use a distribution of the browsing time of the screen which a system displays as a method of satisfaction rating of information service systems. As a result of analyzing using real data, it turned out that users can be classified into three groups. We report on the correlation of the user's classification according to this method and the questionnaire survey result to the user.

Key Words: Service engineering

1 はじめに
ロボットや知識メカトロニクス、IT機器が、人の反応を見ながら適切な動作を行うことや情報提供などのサービスを行うことが期待されている。人がサービスを提供する場合は、相手の反応を見て満足しているかどうか判断し、それに応じた対応を行う。これをロボットなど機械が行う場合は、何をセンサし、どのように判断するかが課題となる。

サービス工学では、人の「満足」は行動に現れると仮定し、これをセンサし判断することで、満足度を向上させるためのシステムの制御サイクルを提案している。我々はサービス工学の視点を取り入れ、IT機器によって実装したWebのサービスの利用者の満足度を図る方法について提案した。

2 サービスシステムとサービス工学
2.1 サービスシステムの現状
人にサービスを提供するシステム、たとえばATMや売券機、ECサイトなど多種多様なサービスが、コンピュータネットワークを使って提供されている。また、ロボットや知識メカトロニクスが、人にサービスを提供する機械等にこれに基づけていくと予想される。これらのシステムの開発者は、画面の表示時間や応答性能、多重性能などに観点を定め、利用者が円滑に使用できるシステムにすべく開発している。

しかし、「使い難い」、「わかり難い」等の苦情は少なくない。こうした利用者の不満はシステムでは検知することができず、アンケート等で顕在化したりする。これは、人に代わって機械がサービスを提供する上での課題であった。

2.2 サービス工学の適用
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。システムが、利用者のサービスに対する評価をセンシングできれば、上記のような課題は解決することになる。

3 利用者のカテゴライゼーション
3.1 満足度測定の難しさ
これまでシステムが利用者のサービスに対する評価、満足度を測定してこなかった背景には、何をセンシングすればよいか明らかでなかったことにある。

3.2 満足は行動に現れる
利用者の満足、不満足は無意識のうちに行動に現れ、これを捉えることができるのではないかと考えた。画面操作に伴うものであれば、画面の触れ方や操作に要する時間などが、感覚を持って熱心にインタラクティブに与えると、それほど興味が無く操作が緩慢になりがちな人で異なり、満足、不満足を反映した傾向を示すのではないかという仮説である。

3.3 Webによるサービスにおける利用者の行動測定
我々はIT機器によって実装したWebを使ったサービスを対象に利用者の行動測定を行った。Webを使ったサービスはマウス操作による要求とその結果の画面の表示といったシンプルなプロトコルで構成されている。画面が表示されてから次のアクションをとるまでの画面の閲覧時間に着目し、この時間から利用者のカテゴライゼーションが可能であるとした。

4 画面閲覧時間によるカテゴライゼーション
4.1 測定対象の概要
対象としたシステムは約70名の登録された利用者が週に何度かアクセスするシステムである。Webサーバでは利用者に画面を表示する時にその時間を書くシステムを製作し、利用者のPCとWeb
サーバはLANで接続されており、画面に情報が表示されるまでの時間は安定して遅延がない。

4.2 閲覧時間の分布
1ヶ月で約700回表示された特定の画面について、表示時間をyとしたときの分布g(y)のヒストグラムをFig.2に示す。70名それぞれがランダムに操作した結果であるが、その分布はFig.2のように正規分布とはならない。

![Fig.2観測された閲覧時間の分布](image)

5 閲覧時間の分布に対する仮説
5.1 3つの群が存在するとする仮説
この利用者は一様な集団ではなく、閲覧時間が異なるいくつかの群からなるのではないかと考えられる。我々は、以下の3つの群から構成されていると仮説を立てた。
・群1：熟心にインタラクティブに使う利用者の群
・群2：どちらでもない利用者の群
・群3：興味が無く操作が緩慢になりがちな利用者の群

また、各群はそれぞれ次の正規分布であると仮定すると、各群の分布は

\[ N(\mu_i, \sigma_i^2) \quad N(\mu_j, \sigma_j^2) \quad N(\mu_k, \sigma_k^2) \]  

と表すことができる。ただし、

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

とし、\( \mu_i, \mu_j, \mu_k \)および\( \sigma_i^2, \sigma_j^2, \sigma_k^2 \)は、群1, 2, 3の利用者の分布の平均と分散、\( \alpha_i, \alpha_j, \alpha_k \)は、それぞれの群の占める比率で、\( \sum \alpha_i = 1 \)とする。

すべての利用者の閲覧時間の密度関数は、その重ね合わせとなり,

\[ f(y) = \sum_i N(\mu_i, \sigma_i^2) \]  

となる。

5.2 合成例
次に、観測した度数g(y)とモデル式(5.3)による度数f(y)から、\( \chi^2 \)式(5.4)が最小となるパラメータを求めた。

\[ \chi^2 = \sum \frac{(g(y_i) - f(y_i))^2}{f(y_i)} \]  

このパラメータでモデル式(5.3)から求めた度数をグラフに表すとFig.3のようになる。比較的短い閲覧時間で次のページに移る群1、閲覧時間に時間を要する群3、閲覧時間に特長のない群2に分かれていることがわかる。[1]

![Fig.3 つの正規分布の重ね合わせ](image)

6 利用者カテゴライゼーションの検証
6.1 アンケート調査
4章で取り上げたシステムの利用者に対して、アンケート調査を実施した。このアンケートにおいて、利用者カテゴライゼーションの結果がユーザ満足度と相関があるか調査するために、以下の設問を設定した。

<table>
<thead>
<tr>
<th>Table 2 アンケート調査の設問</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 統計を使ってみて、役立つと思いましたか？</td>
</tr>
<tr>
<td>2 ホームページが提供する内容・サービスは満足していますか？</td>
</tr>
<tr>
<td>3 ホームページの使い方はわかりやすいですか？</td>
</tr>
<tr>
<td>4 ホームページの文字やグラフなどは見やすいですか？</td>
</tr>
<tr>
<td>5 内容を容易に確認できますか？</td>
</tr>
<tr>
<td>6 日々のデータ入力は容易ですか？</td>
</tr>
<tr>
<td>7 自分の情報が正しく入力できていることを容易に確認できますか？</td>
</tr>
<tr>
<td>8 システムを使いたい時、システムはきちんと動作していませんか？</td>
</tr>
<tr>
<td>9 システムが表示する情報は信用できますか？</td>
</tr>
<tr>
<td>10 情報は安全に守られていると思いますか？</td>
</tr>
</tbody>
</table>


この回答は、SD法に基づいた「どちらともいえない」を中心とした5段階の評価（+2から-2）の近似度の選択することで行っている。43名から回答を得、集計結果はTable.3のようにになった。

Table.3 アンケート回答結果の集計

<table>
<thead>
<tr>
<th>項番</th>
<th>+2</th>
<th>+1</th>
<th>0</th>
<th>-1</th>
<th>-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>31</td>
<td>4</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>20</td>
<td>17</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>21</td>
<td>15</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>22</td>
<td>11</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>18</td>
<td>18</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>8</td>
<td>9</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>16</td>
<td>15</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>19</td>
<td>5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>21</td>
<td>13</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>13</td>
<td>17</td>
<td>13</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

6.2 閲覧時間の分布とアンケート結果の相関の仮説
アンケート実施月の1ヶ月間にその3つの画面の閲覧時間のデータを使用してユーザ・カテゴライゼーションを行った。画面は、「日々の記録を表示」（画面140）、「測定データのグラフ表示」（画面120）、「初期画面」（画面99）であり、各画面は、それぞれ11件、1103件、52件である。
3つの画面の閲覧時間によるカテゴライゼーション結果をTable.4に示す。

Table.4 閲覧時間のカテゴライゼーション結果

<table>
<thead>
<tr>
<th>群1</th>
<th>群2</th>
<th>群3</th>
</tr>
</thead>
<tbody>
<tr>
<td>画面</td>
<td>平均</td>
<td>分数</td>
</tr>
<tr>
<td>140</td>
<td>11.5</td>
<td>4.6</td>
</tr>
<tr>
<td>120</td>
<td>5.4</td>
<td>2.4</td>
</tr>
<tr>
<td>99</td>
<td>12.4</td>
<td>5.2</td>
</tr>
</tbody>
</table>

上記カテゴライゼーションにおける$X^2$はいずれも118.5未満であり、分布表から、10%の有意水準で実測データとカテゴライゼーションによるデータは一致するという仮説は棄却されない。各画面のヒストグラムは、Fig.5, Fig6, Fig7の通りである。
Table.3に示したアンケート結果の集計と、Table.4のカテゴライゼーション結果の各群の比率に関係があることが考えられる。アンケート結果の集計は各観点からの利用者の満足度の表示であると考えられる。それと画面の閲覧時間からカテゴライゼーションした各群の比率に相関がある。このカテゴライゼーションの結果を満足度の判断に使用できる可能性を示すことになる。

6.3 実データによる検証
Table.3をグラフ化すると、Fig.7のようになる。Table.4をグラフ化すると、Fig.8のようになる。
アンケート結果は5段階の評価であり、カテゴライゼーションは3段階の評価である。アンケート結果の階級段階の和を計算して、アンケート結果のデータとしてカテゴライゼーション結果との比較を試みる。

Fig.5 実測した$g(y)$とモデル式$f(y)$（画面140）

Fig.6 実測した$g(y)$とモデル式$f(y)$（画面120）

Fig.6 実測した$g(y)$とモデル式$f(y)$（画面99）
(1)「日々の記録を表示」(140)
「日々の記録を表示」(画面140)のカテゴリライゼーションと類似的アンケート結果項目として4、10が挙げられる。

項観察において、「+2」を群1に、「+1」を群2に、「0」「-1」「-2」を群3に割り当て、X^2を求める3.0となる。この比較検定の自由度は2であり、分布表から、10％の有意水準でカテゴリライゼーションの結果とアンケート結果は一致すると仮説は棄却される。

項観察において、「+2」を群1に、「+1」を群2に、「0」「-1」「-2」を群3に割り当て、X^2を求めると1.1となる。10％の有意水準でカテゴリライゼーションの結果とアンケート結果は一致すると仮説は棄却される。

(2)「測定データのグラフ表示」(120)
「測定データのグラフ表示」(画面120)のカテゴリライゼーションと類似的アンケート結果項目として6、7が挙げられる。

項観察において、「+2」「+1」を群1に、「0」を群2に、「-1」「-2」を群3に割り当て、X^2を求めると0.7となる。10％の有意水準でカテゴリライゼーションの結果とアンケート結果は一致すると仮説は棄却される。

項観察において、「+2」を群1に、「+1」を群2に、「0」「-1」「-2」を群3に割り当て、X^2を求めると9.8となる。10％の有意水準でカテゴリライゼーションの結果とアンケート結果は一致すると仮説は棄却される。

(3)「初期画面」(99)
「初期画面」(画面99)のカテゴリライゼーションと類似的アンケート結果項目として8、10が挙げられる。

項観察において、「+2」を群1に、「+1」を群2に、「0」「-1」「-2」を群3に割り当て、X^2を求めると3.8となる。10％の有意水準でカテゴリライゼーションの結果とアンケート結果は一致すると仮説は棄却される。

項観察において、「+2」を群1に、「+1」を群2に、「0」「-1」「-2」を群3に割り当て、X^2を求めると3.8となる。10％の有意水準でカテゴリライゼーションの結果とアンケート結果は一致すると仮説は棄却される。

以上により、カテゴリライゼーション結果とアンケート結果は相関があるといえる。

7 結論
Webシステムの利用者のWebの閲覧時間を測定すると、その数値の利用者のヒストグラムは、正規分布にはならない。

利用者には3つの群が存在しそれぞれが正規分布であると仮定すると、その重ね合わせのモデル式と実測値のヒストグラムに10％の有意水準で有意な相関がある。モデル式は3つの群それぞれの正規分布の平均、分散、比率をパラメータとして持つ。このそれぞれの群は「満足」、「不満足」、「いずれでもない」を示す群であると考えられる。

ユーザ・カテゴリライゼーションによって抽出した利用者の「満足」、「不満足」、「いずれでもない」群の分布と、利用者のアンケートによる満足度評価の間には10％の有意水準で有意な相関があることが3つの画面を対象に確認することができた。

新たな観測値（利用者）に対し、「満足」、「不満足」、「いずれでもない」のいずれの群に属する確率が高いか、ユーザ・カテゴリライゼーションによって得たウエブマージを事前確認したユーザ満足度判断を行うことは妥当である。

参考文献
[1] 竹原孝成、村田篤、浅間一：サービス工学における満足度のコンサインメント方法の一案、第6回計測自動制御学会SI部門講演会SM2-6、2005。
[3] 浅間一：サービスメディアのための人の行動計測・分析、第9回人工物工学コロキウム、2004。