

ネットワーク携帯端末を用いた通信販売におけるE-businessの特性

筑波大学大学院 *鈴木伸彦 SUZUKI Nobuhiko
筑波大学社会工学系 香田正人 KODA Masato

1. 研究概要

本研究では、携帯端末を用いてオンラインショッピングを行っている企業から提供された携帯Webログデータに対してデータマイニング技術を適用し、ネットワーク携帯端末のアクセス履歴や購買履歴から顧客別に行動(クリック)パターンや嗜好の分析を試みた。

2. 研究目的

- サイト訪問者の商品ページ掲載プロダクトに対する購入意欲を高めるためのサイト構造について研究を行う。
- 有用なEメールキャンペーンの実施を可能にするために、メール送信者のターゲティング(優良顧客の抽出)や送信内容、アクセス時間のカスタマイズについての検討を行う。

3. 携帯端末利用の特性

本稿では、携帯端末特有の結果が得られた、顧客行動時間分析と時系列順パターン検出の結果について述べる。

3.1 顧客行動時間分析

携帯通販会員の行動時間分析を行う。ここでは全会員の過去一年間(2001年8月~2002年7月)のヒット数を時間と曜日でクロス集計し、会員が最も利用する時間帯を判別する。

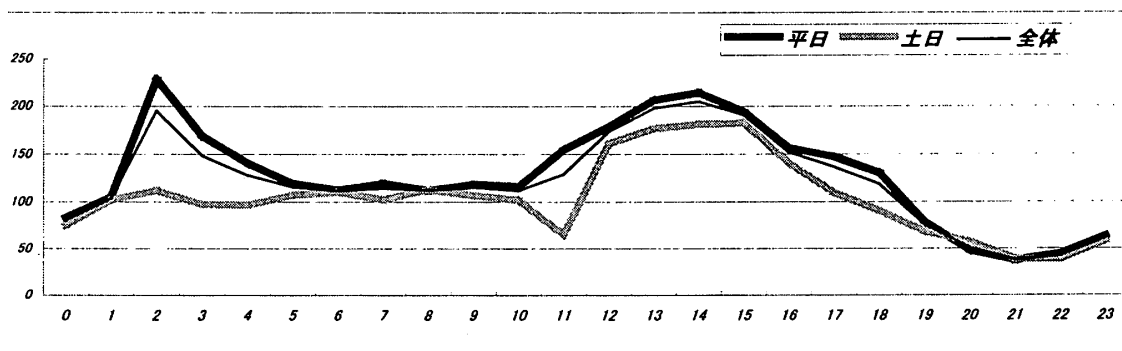


図1 時間別平均ヒット数

図1を見ると、午前2~4時と12~18時という二つのピークが存在する。また、20~24時のヒット数が最も少なく、意外な結果が得られた。一般的には、この時間帯は暇な時間であり、通常はこの時間帯に最もヒット数が増えることが期待される。これは携帯端末からの利用という観点からも推論できる。す

なわち、携帯端末からの利用の最大の特徴は「いつでもどこでも」というユービキタス(ubiquitous)性にあり、暇な時間に何処からでも随時アクセス可能であるという特徴である。しかし、本当に自由な時間帯には携帯通販サイトにアクセスするのではなく、満足度のより高い他の過ごし方を選択する人が多いために、上の推論とは逆に、携帯通販の利用者数が減少するものと推察される。さらに、週末のヒット数は、平日に比べて少ない傾向がはっきりと示されている。こうした逆説的な特徴も考慮して、携帯通販のアクセス時間戦略を練ることが重要であろう。スパムメールを避けるキャンペーン発信時間管理においても同様である。

3.2.1 時系列順パターンにおける確信度

本来、時系列順パターン(Sequential Pattern)分析では相関ルール(Association Rule)で定義されるような確信度(C Confidence)は定義できない。それは、得られた順列に含まれるアイテム間にどれだけの時間経過や、その間に介在した他のアイテムが含まれているかを厳密に表現できないため、確信度に意味がないからである。しかし、ここでは商品ページに至るまでの時系列順パターンを求めることを目的とする。携帯通販のトランザクション単位としての1セッション当りの平均滞在時間は13分強と短時間である。さらに、図1よりアイテムとするページ間にはサイト構造からの強い制約がある。よって、携帯Webログデータにおいては確信度に意味を持たせることが可能になると考えられる。そこで、ここでは時系列順パターンにおける確信度 $conf$ (正確にはConfidence-likeであってConfidenceそのものではない。)を次式で定義することにする。

$$conf(A \rightarrow B) = sup(A \rightarrow B) / sup(A)$$

ただし、上式において $sup(A)$ はアイテムAの支持度 (sup : Support)を表す。

例えば、 $conf(topics \rightarrow prod) = 16.259$ の場合、その解釈は次のようになる。「トピックスのページをヒットした訪問者の16.26%が、その後に商品ページをヒットしている。」

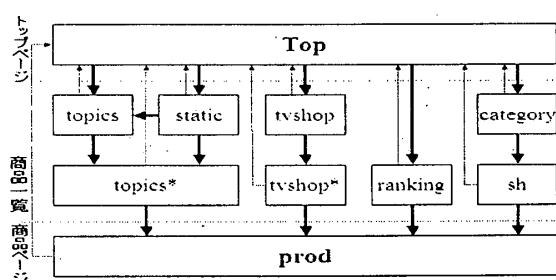


図2 トップページから商品ページへの経路

表1 トップページから商品ページへの経路

トップ		商品一覧	商品ページ
Top	→ topics	→ topics*	→ prod
Top	→ tvshop	→ tvshop*	→ prod
Top		→ ranking	→ prod
Top	→ category	→ sh	→ prod

これにより、ページ推移に関する有益な情報を得ることができる。さらに、携帯端末の場合、ページの上部に設定されたリンク先に移動し易いということが実証された。

謝辞 本研究を行うにあたり、貴重なデータを提供して下さったA社に謝意を表す。