

ECサイトへの訪問キーワードに着目したアクセス動向の可視化手法の提案

久米 亨†

中村 健二†

小柳 滋†

†立命館大学

1 はじめに

インターネットの発達に伴い、ECサイトで商品を購入する機会が増加した。それに伴い、ECサイト間での顧客獲得に向けた競争にも拍車がかかっている。このような背景の中、ECサイトの運営者は、顧客確保のためにECサイトを解析し、その利便性の向上を目指している。ECサイトの分析を支援するための研究としては、アクセスログの分析およびその分析結果を可視化する手法が数多く提案されている。

このような点から、著者らは、アクセスログの分析結果を可視化する上で、どのようなニーズがあるかECサイトの運営者にヒヤリングをおこなった。その結果、「顧客一人一人の行動を把握できること」、「検索キーワードとアクセス先の関連性が把握できること」の2点が重要であるとの回答を得た。そのため、本研究では、上記2点を考慮した可視化手法について提案する。

2 既存研究

Webサイトの状況やアクセス状況を可視化する研究が数多く行われている。Webサイトの状況を可視化する手法 [1][2][3] では、サイト内の構造を可視化する手法が提案されている。これらの手法では、Webサイト全体の構造を把握できるため、サイト管理者がサイトの全体像を把握する際やページ間の関連性の把握などにおいて利用されている。一方、アクセス状況を可視化する手法 [4][5][6] では、Webサイトにおける顧客全体のアクセス動向を可視化する手法が提案されている。これらの手法は、Webサイトにおける顧客のアクセス動向を把握できるため、ページ毎のアクセス数やアクセス傾向の把握などにおいて利用されている。

これらの研究では、サイト全体の可視化を目的としており、サイトの構造やサイト全体のアクセス状況を把握できる。しかし、本研究が目的としているECサイトのアクセスログを可視化する場合には、Webサイト全体のアクセス状況に加えて、上記2点についても詳細に把握する必要がある。そのため、本研究では、ECサイトに特化した可視化手法を提案する。

3 提案手法

3.1 アクセスログの精査

本処理では、アクセスログを顧客のアクセス単位に分割して使用する。分割の流れを次に示す。まず、アクセスログをIPアドレス単位に分割する。次に、各IPアドレスが同じでアクセス間隔が30分以内のものは、

Visualization Method of Customer Trend Based on Access Information to EC Site

†Toru KUME Ritsumeikan University

†Kenji NAKAMURA Ritsumeikan University

†Shigeru OYANAGI Ritsumeikan University

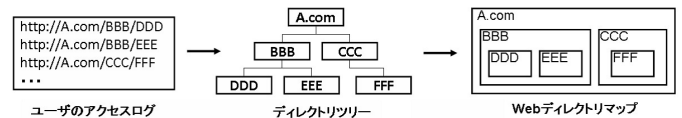


図1: サイト構造の可視化の流れ

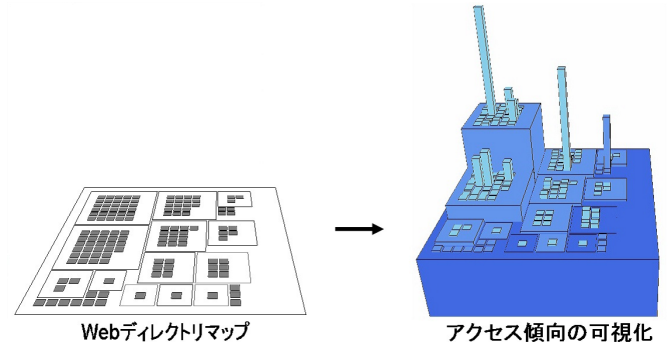


図2: アクセス傾向の可視化

一連のアクセスと見なしセッションを作成する。そして、作成したセッション毎にデータを取得する。最後に、各セッション内で検索エンジンからのアクセスが存在する場合は、その際の検索キーワードを訪問キーワードとして抽出する。

3.2 サイト構造の可視化

本処理では、アクセスログを可視化するための基礎となるWebサイトの構造(Webディレクトリマップ)を作成する。Webディレクトリマップの構築の流れを図1に示す。まず、全アクセスログから、条件に一致するものを取得する。次に、取得したアクセスログに含まれるURLを分割し、包含関係を考慮したディレクトリツリーを構築する。そして、ディレクトリツリーを親が子を含む矩形で構成されるWebディレクトリマップを構築する。

3.3 アクセス傾向の可視化

本処理では、作成したWebディレクトリマップに基づき、アクセスがあったページにアクセス数に応じた高さを与える。アクセス傾向の可視化の流れを図2に示す。各ディレクトリのアクセス傾向を高さと表現することで、アクセスが集中しているページを明確にする。

3.4 アクセスログの可視化手法

3.4.1 顧客IDに着目したアクセス動向の可視化

本可視化手法では、「顧客一人一人の行動を把握できること」を目的として、購入に至った顧客の一連の流れを可視化する。対象となるアクセスログは、購入まで至ったセッションとする。本可視化により、顧客動向やWebディレクトリマップの広さから、商品をどの程度サイト内で探索し購入に至ったのかを把握できる。

顧客Aのアクセス動向

【顧客Aの検索キーワード】回転数 計測器

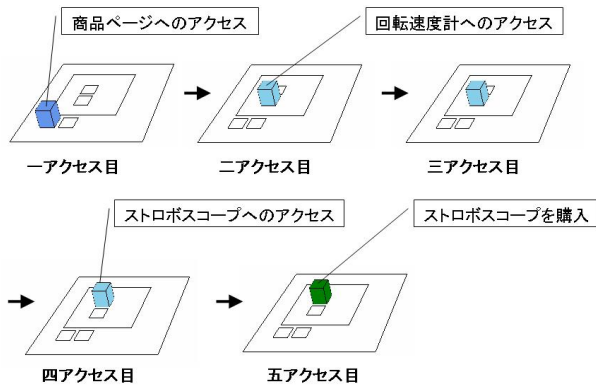


図 3: 可視化手法

3.4.2 検索キーワードに注目したアクセス傾向の可視化

本可視化では、「検索キーワードとアクセス先の関連性が把握できること」を目的として、検索キーワード毎のアクセス傾向を可視化する。対象となるアクセスログは、同一の訪問キーワードを持つセッションとする。本可視化により、ディレクトリの高さから、対象の訪問キーワードでアクセスした顧客がどのページに集中しているかを把握できる。

4 評価実験

4.1 実験環境

本実験では、EC サイト『テックジャム (<http://www.tech-jam.com>)』の 2008 年 4 月のアクセスログを用いて可視化を実現している。アクセスログは、IP アドレス、日時、検索ワード、アクセス先、購入履歴等の情報が保存されており、データ件数は約 20 万件である。

4.2 顧客 ID に注目したアクセス動向の可視化

本実験では、顧客 A がストロボスコープを購入に至るまでの流れを可視化し、その結果を図 3 に示す。

図 3 より、顧客 A は 5 アクセス目で商品を購入しているため、当初から購入する予定だった事が分かる。また、購入品と検索単語が違うが、それに類似した商品を購入していることが分かる。

4.3 検索キーワードに注目したアクセス動向の可視化

本実験では、「フラスコ」「ピーカー」という検索キーワードからのアクセス傾向を可視化し、その結果を図 4 に示す。

図 4 より、「フラスコ」という検索単語に対しては適切なアクセスがあったといえる。また、「ピーカー」という検索単語に対しては、アクセス先が二分されている。これにより、ピーカーを購入する場合の経路が二パターンあることが分かる。

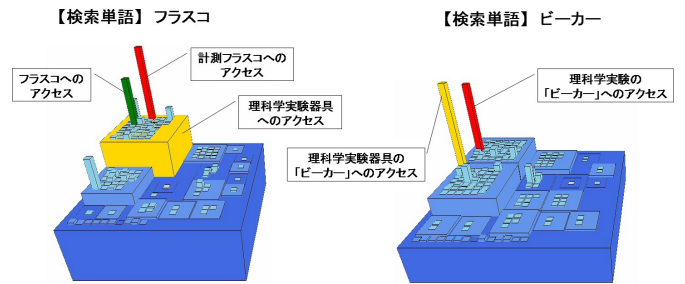


図 4: 可視化手法

5 おわりに

本研究では、EC サイト運営者にヒアリングした結果に基づき、EC サイトで利用可能な 2 つの可視化手法を提案した。EC サイトのアクセスログを用いて実験したところ、顧客動向および検索キーワード毎のアクセス傾向を確認することができた。

今後は、各顧客と同様の動きを行っている顧客の検索や EC サイトの導線設計が不適切である部分の提案など、マーケティングや EC サイトの改善に役立つ情報を自動的に提案する手法を検討する予定である。

謝辞

本論文を遂行するにあたり (株) テックジャムの各氏には貴重なご意見・示唆を賜った。ここに記して感謝の意を表する。

参考文献

- [1] 塩野秀和, 西山晴彦, 松下温: [納豆ビュー] の対話的な情報視覚化における位置づけ, 情報処理学会論文誌, Vol.38, No.11, pp.2331-2342 (1997) .
- [2] Robertson, G., Mackinlay, J. and Card, S.: Cone Trees: Animated 3D Visualizations of Hierarchical Information. In Proceedings of ACM CHI 91, ACM, pp.189-194 (1991).
- [3] Rekimoto, J.: The Information Cube: Using Transparency in 3D Information Visualization, Third Annual Workshop on information Technologies & Systems, pp.125-132 (1993).
- [4] 山口 裕美, 伊藤 貴之, 池畑 裕子, 梶永 泰正: 階層型データ視覚化手法「データ宝石箱」とウェブサイトの視覚化, 画像電子学会誌, 画像電子学会, Vol.32, No. 4, pp. 407-417 (2003).
- [5] 伊藤貴之, 梶永泰正, 池端裕子: データ宝石箱: 大規模階層型データのグラフィックスショーケース, 情報処理学会グラフィクス&CAD 研究会, Vol.2001, No.81, pp.65-70 (2001).
- [6] Itoh, T., Muelder, C., Ma, K. and Sese, J.: A Hybrid Space-Filling and Force-Directed Layout Method for Visualizing Multiple-Category Graphs, 2009 IEEE Pacific Visualization Symposium, IEEE, pp.121-128 (2009) .