

Web上の広告活動の分析:成功報酬型広告を題材として

An Analysis of Affiliate Marketing on the Web

石井 聡一^{*1}

Soichi Ishii

芳中 隆幸^{*1}

Takayuki Yoshinaka

福原 知宏^{*2}

Tomohiro Fukuhara

増田 英孝^{*1}

Hidetaka Masuda

中川 裕志^{*3}

Hiroshi Nakagawa

*1 東京電機大学未来科学部

School of Science and Technology for Future Life, Tokyo Denki University

*2 東京大学人工物工学研究センター

Research into Artifacts Center for Engineering, The University of Tokyo

*3 東京大学 情報基盤センター

Information Technology Center, The University of Tokyo

An advertising activity that uses information on the affiliate program on Web is analyzed in the present study. As a result, the type of merchandise and the price, etc. handled the user who is doing an advertising activity on two or more different blog sites and there are understood. It proposes the advertising activity analysis support system, and, then, it reports on the result of obtaining.

1. はじめに

近年, Web 上ではブログやホームページの作成が簡単にできるようになった [1]. このことから, 個人がブログやホームページ上に企業の広告を掲載する成功報酬型広告 (アフィリエイトプログラム) が盛んになっている [2]. アフィリエイトプログラムにより現在, 個人の Web サイトまでも広告媒体として広く認知され, 広告収入を目的としたスパムが増加するという問題が発生している.

Wang らは, アフィリエイト型スパムユーザから広告主までのネットワーク階層を調べ, 5 階層に渡っている事を示し, スパムとの関係が深いドメインの統計分析を行っている [3]. また長谷らは, ブログ記事内のアフィリエイトリンク数を集計し, スパムブログとの統計分析を行い, アフィリエイトブログの半数がスパムブログであるという結果を報告している [4].

本研究では, Web 上の広告情報を用い, 個人のアフィリエイトプログラム利用者 (アフィリエイト) を特定し, 活動分布と掲載商品を調べることによって, 個人の Web サイトが広告掲載媒体としてどのような活動を行っているのか分析を行う. この分析により, スパムアフィリエイトの活動分析が可能である.

本論文の構成は次の通りである. 2. では成功報酬型広告について述べる. 3. では広告情報を用いた分析と結果について述べる. 4. ではスパムユーザと有名ブログの活動分析を行う. 5. では 3. と 4. の結果を踏まえた考察を述べ, 6. で本論文のまとめと今後の課題について述べる.

2. 成功報酬型広告

成功報酬型広告 (アフィリエイトプログラム) とは, Web サイトに広告を掲載し, 閲覧者がその広告をクリックしたり, 商品を購入すると, 広告主から広告掲載主に報酬が支払われる広告手法である. 表 1 にアフィリエイトプログラムの分類を示

表 1: アフィリエイトプログラムの分類

分類	概要	代表サービス例
ASP 型	各種広告主と広告掲載主を ASP が仲介し, 広告主から広告料, 手数料などを得て, 広告掲載主に広告料を支払う.	A8.net, LinkShare, ValueCommerce
独自型	自社が持つショッピングサイトなどを利用し, 広告主となり, アフィリエイトプログラムを提供する.	Amazon アソシエイトプログラム, 楽天アフィリエイト

す. アフィリエイトプログラムは (1)ASP^{*1}型, (2) 独自型の 2 種類に分けられる. ASP とは, 広告主の企業から広告を預り, サイトに広告掲載を希望するアフィリエイトに広告を提供する中間業者である. 図 1 に ASP 型の仕組みを示す. 独自型では, 広告主自身がアフィリエイトプログラムも提供する.

また, 広告料の種類として, (1) 成果型, (2) クリック型, (3) 表示型, (4) 月極め型がある. 成果型では, 広告を経由して商品等が購入されると, 報酬が支払われる. クリック型, 表示型では, 広告のクリック回数, 表示回数に応じて報酬が支払われる. 月極め型では広告掲載料を掲載期間によって決定する.

3. ブログに出現する広告情報を用いた Web 上の成功報酬型広告の分析

3.1 分析の目的

本分析の目的は, アフィリエイトの活動分布と掲載商品を調べることに伴い, 国内でのアフィリエイトの実態を把握することである. また, アフィリエイトが活発なブログサービス, 時間帯を得て, 今後のアフィリエイト記事収集の為に役立てる. 本分析によりアフィリエイトの特定が行えることから, アフィ

連絡先: 石井 聡一, 東京電機大学未来科学部情報メディア学科,
東京都千代田区神田錦町 2 丁目 2 番地, 03-5280-3281 ext
2843, 03-5280-3592, ishii@cdl.im.dendai.ac.jp

*1 Affiliate Service Provider

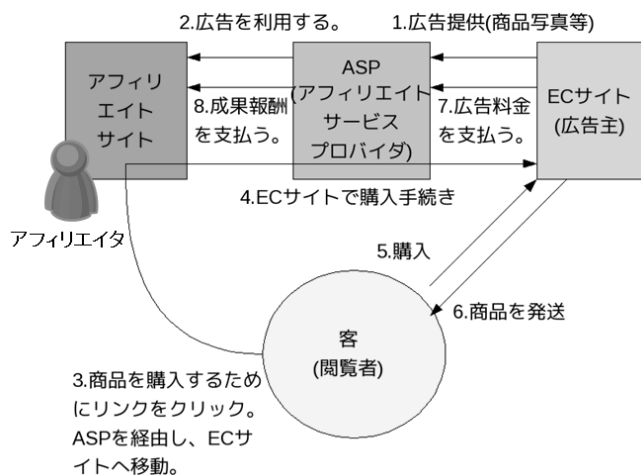


図 1: ASP 型の仕組み

表 2: ブログ記事データ件数

データ	期間	ブログ記事数
A:1 日分	2009 年 1 月 19 日	652,111
B:2 週間分	2009 年 1 月 23 日 ~ 2 月 5 日	578,257

リエイト型スパムの活動実態調査も行う。掲載商品の分析では、商品カテゴリ、価格、などからアフィリエイトの商品傾向を調べる。

3.2 分析データ

分析データとして、8社^{*2}のブログサービス運営会社から取得した記事を利用する。ブログサービスから配信されるRSSを定期的に取得し、そこから新着記事のURLを抽出し、記事を収集した。表2にデータ件数を示す。

分析方法では、HTML内にあるリンク、タグ情報などを用いて、13社^{*3}のアフィリエイトプログラムの記事数を集計し、その中から大手3社 Amazon アソシエイトプログラム(以下 Amazon)、Google AdSense(以下 Google)、楽天アフィリエイト(以下楽天)のアフィリエイトを特定し、アフィリエイトの数や活動状況を調べる。

3.3 活動分布の分析

1 日分での分析

まず記事数の分析では、2009年1月19日に取得した記事652,111件中13社のアフィリエイト記事は119,495(18.3%)であった。この中からAmazon、Google、楽天の3社に絞ったところ54,978件(8.5%)であった。表3に1日分データでの3社の記事数^{*4}、アフィリエイト数を示す。3社のアフィリエイト記事からアフィリエイト(ユーザ数)を特定したところ68,857件となった。そこから同一ユーザを除くと、14,487件(21.0%)に減少した。これをユニークユーザとして、記事数をユニーク

*2 Ameba,fc2,JUGEM,livedoor,Seesaa,Yahoo!, ココロログ, ウェブリブログ

*3 A8.net,AccessTrade,Amazon アソシエイト,DMM,Google AdSense,infocart,infostore,infostyle,LinkShare,Store-mix,TrafficGate,ValueCommerce, 楽天アフィリエイトの13社である。

*4 表3,4では、楽天はユーザが特定出来る記事だけを集計した。

表 3: 1 日分データにおけるサービス毎の記事数とアフィリエイト数

サービス名	記事数	ユーザ数	ユニークユーザ数	投稿数
Amazon	30,368	36,568	8,209	3.70
Google	24,457	27,962	5,366	4.56
楽天	3,418	4,327	912	3.75
合計	54,978	68,857	14,487	3.79

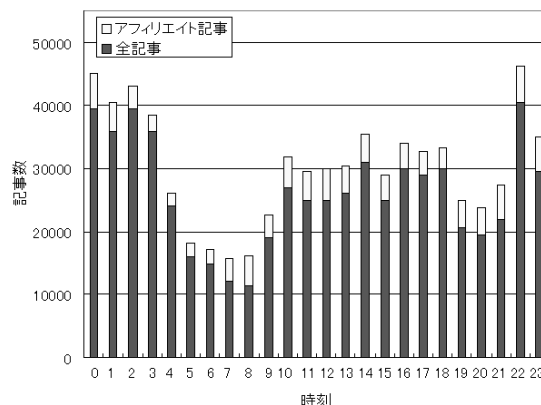


図 2: 1 日分データの全記事数とアフィリエイト記事数比較 (積み上げグラフ)

表 4: 2 週間分データにおけるサービス毎の記事数とアフィリエイト数

サービス名	記事数	ユーザ数	ユニークユーザ数	投稿数
Amazon	18,236	22,182	6,398	2.86
Google	12,158	16,516	4,323	2.81
楽天	2,065	2,435	906	2.28
合計	34,796	41,103	11,627	2.99

ユーザ数で割り、1 ユーザ当たりの記事投稿数を調べたところ 3.79 となった。

図2に1日分データにおける記事数とアフィリエイト記事数の比較を示す。全記事、アフィリエイト記事共に0時、12時、22時前後に記事が集中していることがわかる。また7,8時台では、アフィリエイト記事の割合が多い結果となった。

2 週間分での分析

前項の結果から、2週間分データとして、2009年1月23日から2月5日までの0時、13時、23時に記事を取得した。取得した578,257件中13社のアフィリエイト記事は67,512件(11.7%)であった。この中から3社に絞ったところ34,796件(6.0%)であった。表4に2週間分データでの3社の記事数、アフィリエイト数を示す。アフィリエイトを特定したところ41,103件となった。そこから同一ユーザを除くと、11,668件(28.4%)に減少し、記事投稿数を調べたところ2.99となった。

図3に2週間分データにおける記事数とアフィリエイト記事数の比較を示す。1月30日(金)、31日、2月1日の全記事数が少なく、アフィリエイト記事の割合は順に、15.8%、13.2%、14.3%であった。

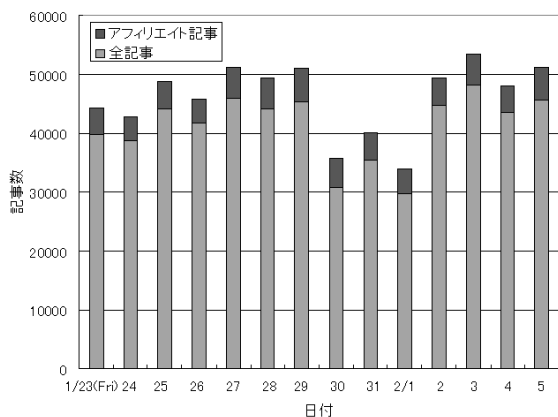


図 3: 2 週間分データの全記事数とアフィリエイト記事数比較 (積み上げグラフ)

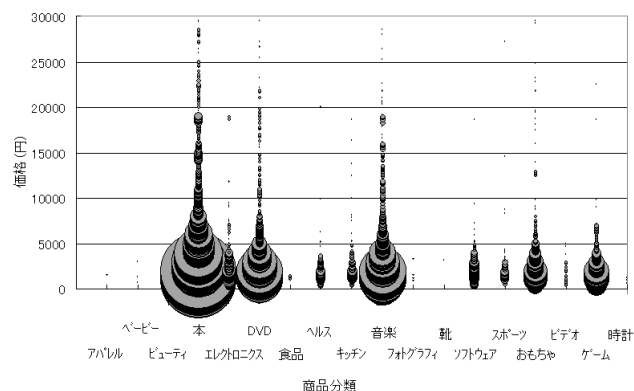


図 4: 2009 年 1 月 19 日における掲載商品の価格帯と商品カテゴリ

3.4 掲載商品の分析

本節では、Amazon の掲載商品の分析を行う。記事データから Amazon 独自の商品識別番号 ASIN*5 を抽出し、AWS*6 を用いて商品情報を取得した。

1 日分での分析

抽出した ASIN は 112,118 件であった。そこから同一商品を除くと 16,215 件 (14.4%) となり、その内、価格が取得できたものは 12,818 件 (79%) であった。商品分析の結果、18 の商品カテゴリを取得し、価格は 158 円から 354,113 円であった。3 万円以上の商品は 87 件 (価格取得件数中 0.67%) となり、高額商品の掲載はごく一部である。図 4 に 12,818 件中 3 万円以下の商品分類と価格を示す。商品分類に着目すると、本 (50.4%)、音楽 (18.6%)、DVD (17.0%) の掲載が多く、価格では、1 万円以下の商品が 12,320 件 (96.1%) となっている。

2 週間分での分析

抽出した ASIN は 67,399 件であった。同一商品を除くと 30,859 件 (45.8%) となり、価格が取得できたものは 23,676 件 (76.7%) であった。商品分析の結果、19 の商品カテゴリを取得し、価格は 193 円から 448,000 円であった。図 5 に 1 日分と 2 週間分データでの価格帯と商品件数を、図 6 に 1 日分と

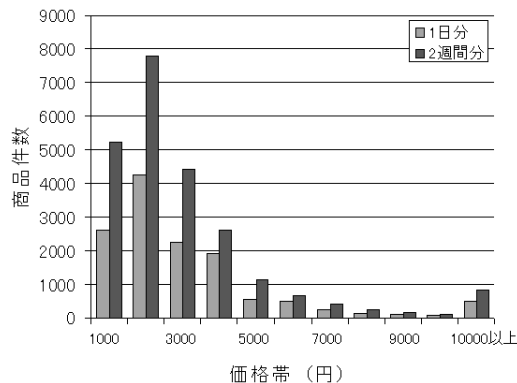


図 5: 1 日分と 2 週間分データでの価格帯と商品件数

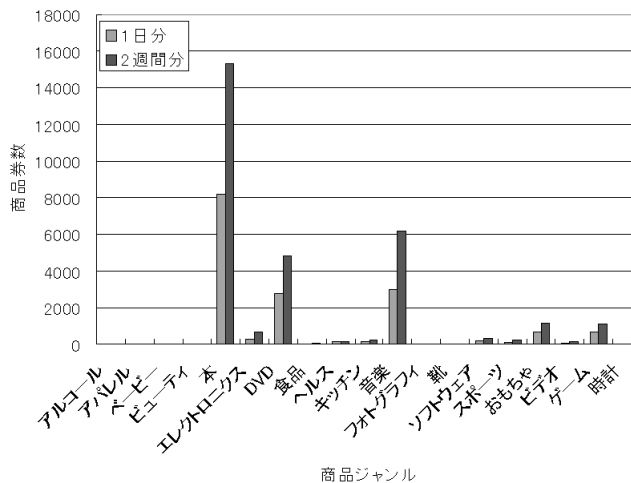


図 6: 1 日分と 2 週間分データでの商品カテゴリと商品件数

2 週間分データでの商品カテゴリと商品件数を示す。1 日分の分析と同じく、本 (49.7%)、音楽 (20.1%)、DVD (15.6%) の掲載が多く、1 万円以下の商品が 22,840 件 (96.5%) であった。

4. スпамユーザと有名プロガの活動分析

本分析により、アフィリエイトの特定が行えることから、1 日分の分析データからスパムユーザを特定し、Amazon の掲載商品のカテゴリ分析を行う。また、比較として有名プロガ (アルファプロガ) の分析も行う。

4.1 スпамユーザの特定

収集したアフィリエイト記事からユニークユーザを特定し、ブログサービスが利用するユーザを除去した後に、ユーザの投稿回数を集計し、上位 20 ユーザをスパムユーザ候補とし、目視でスパムの判定を行った。20 ユーザ中全てがスパムユーザであった。

4.2 活動分析

全ユーザの活動を確認してみたところ、複数のブログサービスにまたがって記事を投稿しているスパムユーザが 3 人存在した。図 7 にスパムユーザが利用するブログサービスの割合を示す。今回の分析では、ブログサービス A が突出してスパムユーザが多い結果となった。

*5 Amazon Standard Identification Number

*6 Amazon Web Services

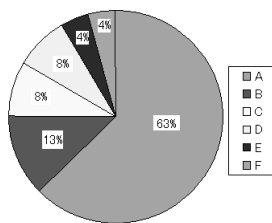


図 7: スпамユーザが利用するブログサービス

表 5: スпамユーザと有名ブログの掲載商品 (一部)

スパムユーザ	有名ブログ
「ストウブ」でじんわりほっこり幸せなレシピ?シェフに愛されるフランスの人気鍋 (主婦の友生活シリーズ)	なぜ世界は不況に陥ったのか 集中講義・金融危機と経済学
福山雅治 Photo&Episode The One & Only Boy (RECO BOOKS)	ハイエク 知識社会の自由主義 (PHP 新書)
相撲 2009 年 01 月号 [雑誌]	市場の変相
赤い糸 上	電波利権 (新潮新書)
改訂新版病気の原因は背骨にある「小林整体」の魅力?即効!ダイナミック矯正で痛みと病根を断つ	過剰と破壊の経済学 「ムーアの法則」で何がわかるのか? (アスキー新書 042) (アスキー新書)
ヘルシー耳ツボダイエットでやせて魅せる?ツボ療法と最新栄養学を組み合わせた科学的ダイエット法	Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism
やせない理由(わけ)?これなら瞬時にやせられるヘルシー耳ツボダイエット法	なぜ、アメリカ経済は崩壊に向かうのか?信用バブルという怪物
ロザリオとパンパイア 10 (10) (ジャンプコミックス)	致命的な思いあがり (ハイエク全集 第 2 期)
植物はなぜ 5000 年も生きるのが?寿命からみた動物と植物のちがいが (ブルーバックス)	最底辺の 10 億人 最も貧しい国々のために本当になすべきことは何か?
花粉エキスの力 (パワー)?更年期を乗りきる	世界経済危機 日本の罪と罰

掲載商品の分析では、商品カテゴリを 3.4 で最も取得件数が多かった本に絞って、複数のブログサービスにまたがって記事を投稿していたスパムユーザの中で、本のみを紹介していたスパムユーザ 1 人を対象とした。有名ブログでは、アルファプログラマー*9を参考に本のみを紹介していた 1 人を対象とした。表 5 にスパムユーザと有名ブログの掲載商を示す。スパムユーザの商品ジャンルでは、ばらつきが見られ、有名ブログでは、ジャンルに均一性が見られた。

5. 考察

3.3 では、時刻、曜日による記事数の変化とユニークユーザ数を調べた。1 日分では、全記事 652,111 件中アフィリエイト

記事が 119,495 件 (18.3%)。2 週間分では、全記事 578,257 件中アフィリエイト記事が 67,512 件 (11.7%) であった。1 日分データでは、7, 8 時にアフィリエイト記事の割合が多く、2 週間分の分析では、1 月 30 日 (金)、31 日 (土)、2 月 1 日 (日) にアフィリエイト記事の割合が多かった。今回の分析では、データ量が十分とは言えず、また取得期間も短い為、今後のさらなる分析結果から判断する。この一部のデータを除けば、アフィリエイト記事は時間、曜日に問わず一定量存在していることが言える。ユニークユーザ数では、Amazon、Google、楽天に絞り、分析を行った。1 日分データでは、3 社合計のユニークユーザ数が 14,487 件 (21.0%) となり、2 週間分データでは、11,668 件 (28.4%) となった。このことからアフィリエイトは同じ ID で複数の記事を投稿していることがわかる。今後、ユニークユーザのスパム、非スパム判定を行い、データを蓄積しておけば、3 社のアフィリエイト型スパムのフィルタリングが可能である。

3.4 では、Amazon での掲載商品のカテゴリ、価格帯を調べた。その結果、1 日分、2 週間分データ共に商品カテゴリでは、本、音楽、DVD の順で掲載が多く、本に至っては全掲載商品の半分を占めていた。このことから価格帯も 5,000 円以下だけで、1 日分 86.3%、2 週間分 89.6%となっていた。Amazon は成果報酬型なので、購入されやすい本に集中したと考えられる。

4.2 では、スパムユーザの特定を行い、スパムユーザと有名ブログの活動分析を行った。スパムユーザの活動分析では、ユーザの投稿回数を参考に、スパムユーザ候補 20 人を抽出したところ全てスパムであった。この内、複数のブログサービスにまたがり、スパムを投稿しているユーザが 3 人存在し、ブログサービスによる掲載商品の变化は見られなかった。今後の分析では、スパムユーザと非スパムユーザにおいて掲載商品に特徴的な差が見られるのか、定量的な分析を行っていく。

6. おわりに

本論文では、Web 上の広告情報を用いた広告活動の分析を行った。Web 上の広告情報を利用することにより、ユニークユーザ数、掲載商品、価格、ユーザの商品傾向を得られることがわかった。

今後の課題として、ブログ記事の本文とサイドバーを分けた分析、分析データを増やした定量的かつ継続的な分析を行う。

参考文献

- [1] ブログの実態に関する調査研究の結果. 総務省情報通信政策研究所調査研究部. 2008.
- [2] アフィリエイトサービス市場動向に関する調査結果 2008 年版. 矢野経済研究所. 2008.
- [3] Y.M. Wang, M. Ma, Y. Niu, and H. Chen. Spam double-funnel: Connecting web spammers with advertisers. In Proceedings of the 16th WWW Conference, pp.291-300, 2007.
- [4] 長谷巧, 山本匠, 原正憲, 山田明, 西垣正勝. アフィリエイトに着目したスパムブログ評価方式に関する検討. 情報処理学会研究報告, 2009-CSEC-44, pp.97-102, 2009.

*9 <http://alphabloggers.com/alphabloggers>